FACTORS INFLUENCING LESION DETECTION IN MEDICAL IMAGING

by

Jannick Paule Yvette Rolland

A Dissertation Submitted to the Faculty of the

COMMITTEE ON OPTICAL SCIENCES (GRADUATE)

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1990

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Jannick Paule Yvette Rolland

entitled Factors Influencing Lesion Detection in Medical Imaging

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

H. H. Barrett	May 3, 1990 Date
R. Shucland	May 3, 1990 Date
George W. Seeley	May 3, 1990 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.

H. H. Barrett — Dissertation Director

May 17, 1990 — Date

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 that accurate acknowledgment 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 judgment 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: _____

To my Mam and Dad,

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to all who have given me guidance and support towards my education and especially for this research. I would like to give a sincere thanks to my advisor, Harry Barrett, for giving me the opportunity to work by his side and for creating such a humanistic and stimulating research environment to work in. I would also like to thank George Seeley and Mary Ker for their collaboration with the psychophysical studies and especially George Seeley for his warm encouragement. I will never have enough thanks to send to all the observers for their cooperation and patience in running the psychophysical studies, and especially Jie Yao, Kelly Rehm, and Ted Gooley. I would like to extend my appreciation to the other members of the radiology research group for their cooperation and friendship. Special thanks go to Bill Dallas, Hans Roehrig, Jack Denny, John Aarsvold, Rich Lamoreaux, and Michel Rogulski. I would like also to thank Arthur Burgess from the University of British Columbia, and Kyle Myers and Robert Wagner from the Center for Devices and Radiological Health for some very helpful discussions and encouragement during this research. Special thanks also go to all my friends, especially Mary and Rudy Vasquez, Preston Baggette, John Hynes, Jamie Lantz, Rex Stuesst, Sue Nidermaier, Nicolas Bilder, Beth Jordan, Lee Stanley, and Bruce Scher for their sincere friendship and support over the years. Finally, I want to deeply thank my parents who have given me the joy of learning, and whose love and encouragement, transcending the many miles between us, have been a true driving force for completing this research.

This work was supported by NCI grant PO1 CA23417.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

An important goal in medical imaging is the assessment of image quality in a way that relates to clinical efficacy. An objective approach to this goal is to evaluate the performance of diagnosis for specific tasks, using receiver-operating-characteristic (ROC) analysis. We shall concentrate in this work on classification tasks. Many factors may confine the performance achieved for these tasks, but we shall concentrate on two main limiting factors : imaging bias due to the form of the point spread function (PSF) and object variability.

Psychophysical studies followed by ROC analysis are widely used for system assessment, but it is of great practical interest to be able to predict the outcome of psychophysical studies, especially for system design and optimization. The ideal observer is often chosen as a standard of comparison for the human observer since, at least for simple tasks, its performance can be readily calculated using statistical decision theory. We already know, however, of cases reported in the literature where the human observer performs far below ideal, and one of the purposes of this dissertation is to determine whether there are other practical circumstances where human and ideal performances diverge. Moreover, when the complexity of the task increases, the ideal observer becomes quickly intractable, and other observers such as the Hotelling and the nonprewhitening (npw) ideal observers may be used in the process of looking for good predictors of human performance.

A practical problem where our intuition tells us that the ideal observer may fail to predict human performance occurs with imaging devices that are characterized by a PSF having long spatial tails. The investigation of the impact of long-tailed PSFs on detection is of great interest since they are commonly encountered in medical imaging and even more

generally in image science. We shall show that the ideal observer is a poor predictor of human performance for a simple two-hypothesis detection task and that linear filtering of the images does indeed help the human observer.

Another practical problem of high interest is the effect of background nonuniformity on detectability. This study is of considerable interest because it is one more step towards assessing image quality for real clinical images. When the background is known exactly, the Hotelling and the npw ideal observers predict that detection is optimal for an infinite aperture; a spatially varying background results in an optimum aperture size that matches very closely the size of the signal. Moreover, given a fixed aperture size and a background known exactly, an increase in exposure time is highly beneficial for both observers. For spatially varying backgrounds, on the other hand, the Hotelling observer benefits from an increases in exposure time, while the performance of the npw ideal observer quickly saturates as the exposure time increases. In terms of human performance, results show a good agreement with the Hotelling-observer predictions, while the performance disagrees strongly with the npw ideal observer.

CHAPTER 1

INTRODUCTION

Medical imaging is the art of using the interaction of radiation with the natural elements within the body (soft tissue, bones, air, trace elements) to produce a functional or morphological visualisation of its different parameters. The basic clinical imaging techniques are the measurement of x-ray and ultrasound attenuation through the body, the number of gamma rays emitted by radioactive tracers which have been injected in the body, and the spatial density of spins using nuclear magnetic resonance imaging (MRI). Any imaging technique may be conceived of as a two-stage process: a data-acquisition stage and a processing and display stage. The acquisition stage usually involves a radiation source, an object to be imaged, an imaging system, and a detection device. The radiation source may be external to the object as in x-ray, ultrasound imaging, and MRI, or internal to the object as in nuclear medicine where a pharmaceutical, tagged with a radioactive label, is injected into the patient to bind with or accumulate at certain sites within the body. The quality of the images depends on the design and level of optimization of the imaging device and on the complexity of the object to be imaged.

Although we shall focus in this work on problems encountered in radionuclide imaging or most commonly referred to as nuclear medicine, the concepts developed towards assessing image quality can be applied to other imaging modalities as well. A typical imaging system used in conventional nuclear medicine is an Anger camera, which consists of a collimator, a scintillation crystal, and some electronics to estimate and record the position of the detected gamma rays. To yield high resolution, the collimator must be very selective, which usually implies very low collection efficiency. Because of the low number of counts

recorded at the image plane, the images present a grainy appearance due to quantum or Poisson noise. To lessen this effect, we may increase the collimator bore diameter or decrease its length, but those measures lead to a loss in resolution. To capture more photons, we may also think about increasing the dose in the patient or the counting time, but both of these solutions are unacceptable to the patients, since increased dosage may only be at the expense of tissue damage, while longer exposure times cannot be used for most patients, not only because of their critical states of disease, but also because the acquisition mechanism is not immune to patient motion. Patient motion is surely more likely to occur with increased exposure times and would cause blurring of the images. Because of these restrictions in dosage and exposure time and the necessity to work at high resolutions, nuclear medicine images are always strongly corrupted by Poisson noise.

Medical images are always taken to fulfill a precise purpose or task, and the quality of the images is a reflection of how well this task is performed. One example is the simple classification task, where the physician is asked to classify images as normal or abnormal and, within the abnormal images, the type of the abnormality (for example, pneumonitis, emphysema or chronic bronchitis for lung diseases). To make a diagnosis, the physician uses different sources of information, such as previous studies performed on the patient as well as the personal and family history of the patient. Even then, due to limitations inherent in the imaging process or inherent to the object itself, viz., Poisson noise, blur due to the imaging aperture and the detector, and object variability, to name the most common ones, it is sometimes difficult for the physician to make a correct diagnosis. The purpose of medical imaging is therefore to increase the probability of making a correct diagnosis.

Before exploring ways to increase the performance of diagnosis, we need to define a meaningful evaluation process. Receiver-operating-characteristic (ROC) analysis has been

widely accepted in the medical community as being the way to evaluate accuracy of diagnosis (Swets 1986). Many of the concepts on ROC analysis can be found in a special issue of Seminars in Nuclear Medicine edited by Patton (1978), and some of them are pointed out here. If we assume that, to classify each image of a set, we can compute a number that we shall refer to as the test statistic, the discrimination between normals and abnormals is done by comparing the value of this number to a threshold. A key point in ROC analysis is that the threshold value is made a variable. We define the ratio of the number of abnormal images for which the test statistic is greater than the threshold to the total number of abnormal images to be the true positive fraction (TPF) at the given threshold, and similarly, the ratio of the number of normal images for which the test statistic is greater than the threshold to the total number of normal images, the false positive fraction (FPF) at the given threshold. The ROC curve is then a plot of the TPF versus the FPF as the threshold is varied from a strict to a lax value. Some parameters, such as the area under the ROC curve or the detectability index, can then be estimated from the ROC curve to get a single number as a measure of diagnostic performance. ROC analysis is commonly performed to find out which imaging device yields the best performance of diagnosis, in other words to assess image quality in the clinic. Two approaches may be taken in this respect: either carry out psychophysical studies and analyze the data using ROC or define and compute a test statistic that is useful to the assessment of image quality in the clinic before performing ROC analysis.

Psychophysical studies, very useful in an early stage of image quality assessment, are also very tedious and become largely impractical when dealing with system design and optimization since the performance must be evaluated for each value of the parameters that are being optimized for. This is why an alternative solution is investigated here, and we shall try to find, for each specific task, a useful figure of merit to assess image quality. To

be a useful metric, a figure of merit should be a scalar such that all the tradeoffs between the different parameters of the imaging process are accounted for automatically. It should, of course, be calculable, and, because we direct our efforts in system assessment towards assisting physicians in their diagnosis, the figure of merit should correlate well with human performance. In some particular cases, where there is, for example, a lack of physican experts to diagnose the images and where the cost of diagnosis becomes quite high, excluding an important part of the population from being tested, one can imagine ignoring this correlation and designing an expert system, calibrated using some training sets, that will diagnose the images automatically (Bibbo et al. 1983). This is a totally different problem from what we are looking at, but worth mentioning since one might ask why not use only mathematical observers if their ability to perform is higher than the human. Our research is directed towards working side by side with physicians and assisting them in making their diagnoses by providing them with better imaging systems. The correlation between the theoretical observer and the human observer is, therefore, of primary interest to us. Moreover, we intend to use mathematical observers not to substitute for physicians but as a means to optimize the imaging system during the design stage.

Assuming the use of ROC analysis to evaluate performance of diagnosis, we shall now discuss ways to increase the performance . Given an object and an imaging modality, the first step to undertake is to optimize the imaging process, such as the aperture geometry, the distance from the object to the aperture, and the distance from the aperture to the detector. Besides the limitations due to the imaging modality, the presence of overlapping structures in the images is often a significant limiting factor as well. Historically, the recognition of overlapping structures as a limiting factor has been the motivation for the development of computed tomography, where one may reduce overlapping structures by imaging slices of a 3D object instead of the object as a whole. We propose here to study the

effect of background nonuniformity on detection and discrimination and investigate their impact on the optimization process. Another way of increasing the performance of diagnosis is to explore the benefit of image enhancement on image quality, as a way to extract some specific information from the images and make it visually accessible to the observer. The enhancement may include, among other things, greylevel manipulation, high-pass filtering and noise smoothing. An important point about image enhancement is that, while an image may subjectively look more pleasing to our eyes, it may not help the physician in making a better diagnosis. Therefore, image enhancement should always be thought of in terms of the task to be performed. Since we are interested in the objective assessment of image quality, psychophysical studies are carried out to ascertain the usefulness of the enhancement technique applied, while the performances of statistical observers are computed to look for a substitute for psychophysical studies.

The goal of this dissertation is to make new contributions to the assessment of image quality in medical imaging, and our approach here is to study the impact of some imaging factors such as resolution and time exposure on the ability of human observers to perform well specified tasks that present unusual complexities. In the process of this study, we shall compare the performance of the human observer with several mathematical observers and point out several discrepancies that we found. Chapter 2 presents a review of the assessment of image quality in medical imaging. We review the use of statistical decision theory as a means to define computable figures of merit to be compared with human performance evaluated through psychophysical studies. Chapter 3 first discusses the general importance of long-tailed PSFs. Long-tailed PSFs are described as the sum of a sharp central response that characterizes the intrinsic resolution of the system and a broad response that models, for example, scatter in the patient's body or septal penetration in the collimator; the performance of several statistical observers using long-tailed PSFs are then derived.

Experimental techniques and psychophysical studies, to compare human performance with theoretical predictions using long-tailed PSFs, are presented in chapter 4, where problems involving the display of the computer-simulated images and their impact on performance are investigated. Chapter 5 allows for the study of randomness in the background of the object to be imaged. The effects of the size and the shape of the aperture and the effect of the exposure time on the performance of two statistical observers are presented. The methods and results of two psychophysical studies carried out using "lumpy" backgrounds to study the correlation of the human observer performance with either the Hotelling or the npw ideal observers are given in Chapter 6. Chapter 7 concludes this work with a discussion of the impact of more realistic models of imaging systems on human observer performance and some suggestions for further investigations.

CHAPTER 2

ASSESSMENT OF IMAGE QUALITY IN MEDICAL IMAGING

The ability of medical experts to detect lesions depends to a great extent on their degree of expertise, but even the skilled expert will fail to perform accurately if the images available for observation are of poor quality. This chapter reviews different ways to assess image quality in medical imaging and the convergence of those paths to a common goal, finding a figure of merit that may be used not only to choose the best imaging modality but also to optimize it properly -- in other words, to yield images of higher quality to the expert eye.

General Considerations

Evaluation of medical imaging systems faces a long-standing problem of how to assess image quality. While contemplating a Picasso, a group of people may express very different opinions ranging from great appreciation to total rejection. This type of quality assessment is purely subjective, a strong function of the observers involved, and has a fuzzy correlation with the object under observation. In medical imaging we are interested in the objective assessment of image quality, since the acquired images can be very pleasing to the eye but not show some of the features which are important to the physicians who are making the diagnosis. Despite the fact that image quality has been an elusive point to define, one thing is certain : it must be defined in terms of the task to be performed.

The purpose of medical imaging is to assist physicians in making a correct diagnosis on the state of disease of a patient, and how well this purpose is fulfilled is a measure of the quality achieved. Generally, the task to be performed may be conceived of as either a classification or an estimation task, and often a combination of the two (Barrett

1990). Classical examples of classification are detection of, say, lesions as in brain imaging, nodules in chest radiographs, stenosis as in blood flow studies, and abnormal-tissue texture as in lung imaging or liver scans. Examples of estimation may be the determination of the size of a lesion, the degree of stenosis for abnormal coronary blood flow, or the ejection fraction as in cardiac imaging. The specification of the task is therefore the first step to undertake when trying to define image quality. We shall focus in this dissertation on one form of classification task, the two-hypothesis or binary discrimination task, of which the detection task is a simple but important case. All the concepts and results formulated for the binary problem may, however, be generalized to an M-class problem by dividing the M-class problem into a sequence of two-class problems. In a binary discrimination task, one must decide between the presence of a signal s_1 under hypothesis H_1 or the presence of a signal s_2 under hypothesis H_2. In the case of a binary detection task, the hypothesis that there is no signal present is called the null hypothesis and is often denoted as H_0. To standardize the notations, however, we shall denote as H_1 the hypothesis that no signal is present and as H_2 the hypothesis that there is a signal present.

The next step in assessing image quality is to find a means to rank imaging systems. Physical measurements such as the modulation transfer function, noise power spectrum, signal-to-noise ratio (SNR) and resolution have been commonly used as indices of image quality to assess imaging systems (Shosa and Kaufman 1981). The problem with physical measurements is that it is quite unclear, in general, what their relationship to clinical efficacy is. The efficacy of diagnosis is defined here as the extent to which an observer can differentiate among several states of disease and health using an imaging procedure. A solution to this problem is to define a scalar figure of merit that is calculated in such a way that all the trade-offs of image quality, resulting from an increase of the values of some physical measurements while others decrease, are automatically taken into

account (Wagner et al. 1985). Moreover, to be at the same time meaningful to the clinic, this figure of merit must correlate well with the performance of the human observer, viz., the physician. The most common approaches are detailed clinical evaluation or the use of physical phantoms and each, with its own advantages and disadvantages, have been investigated in the medical community to try to solve this problem. All these approaches stress a common important point, however, that image quality of an isolated image has no meaning in itself because of the random nature of the imaging process. Thus, we must always, while addressing the problem of image quality, consider an ensemble of images.

Detailed clinical evaluation is the first approach that comes to mind. This approach consists of collecting a set of similar and well verified clinical cases that must be carefully matched for any comparison of two systems (Ker et al. 1988). The advantage of such a procedure is that it directly relates to clinical efficacy. The pitfalls are the fact that the procedure must be repeated for each clinical case and for each set of physical parameters of the modality under optimization. Moreover, matching cases must be found if two modalities are to be compared together. Clinical trials are therefore quite difficult in terms of assessing image quality and totally impractical if dealing with system design and optimization.

A more flexible approach, which has become very popular, is to use physical phantoms. Most phantoms are either very crude simulations of a real object or not even a pretense of one. The advantage of using even crude phantoms is that we now have full knowledge of the object to be imaged, and certain diseases may be simulated by adding, for example, tumors at different locations in the object. It now becomes possible to generate images for different parameters of the imaging modality in a reasonable amount of time and also to eliminate the nontrivial matching procedures. The main disadvantage is that the equivalence with clinical images is again often quite unclear, a model of a physical process

being only an approximation to reality.

One of the difficulties in assessing image quality in a way that relates to clinical efficacy comes from the fact that the final element in the analysis of the imaging system is the human eye-brain system, which is still too poorly understood to be accurately modelled. This leads to the necessity of carrying out psychophysical studies to take into account the human observer. This is quite an obstacle, however, when dealing with system design and optimization because of the length of time required to evaluate the systems over the whole range of values that the varying parameters are allowed to take. Because of this inconvenience, we need to explore some alternatives to psychophysical studies.

One alternative has been to consider the visual system as an optimum or suboptimum detector and to use statistical decision theory in combination with some physical measurements of the imaging modality to compute the performance of several mathematical observers (Wagner 1977, 1985). These mathematical observers can be used as standards of comparison for human observers, if they are found to be good predictors of human performance evaluated through psychophysical studies. In the process of comparing mathematical and human observers, an important question to be addressed is whether or not, for the task at hand, any information is being lost or altered at the interface between the human observers and the cathode ray tube (CRT) monitor used as the display device (Biberman 1973; Giger 1985, 1986).

Image Formation

Radionuclide imaging is a technique that uses the property of gamma rays to penetrate the body to image regions of interest within the body. The basic imaging process in nuclear medicine is the projection of a radioactive 3D object on a 2D detector plane by the use of an aperture between the object and the detector. Gamma rays cannot be focused

since the index of refraction of any material, which is a function of the energy of the incoming radiation, becomes 1 for highly energetic radiation. We shall assume in the following discussion that we are only interested in a thin slice of the object, and therefore that the object is also a planar radioactive emitter. The images, thus formed from a planar object to a planar detector, are termed planar images. With the assumption of planar imaging, we also make the assumption that the imaging systems we shall consider are linear and shift-invariant. By linearity we mean that the principle of superposition applies in that an image of a linear sum of objects is equivalent to the sum of the individual images of the objects. This assumption is valid in nuclear medicine because of the small number of photons or counts that are collected to form an image, this small number being a consequence, of course, of the small doses of radiation that are allowed to be injected in the patient at one time. A factor that would invalidate the linearity assumption would be the saturation of the detector that may occur if the imaging time is too large or the emissive source too strong, but this is generally not the case in nuclear medicine. The shift-invariance property is generally a valid assumption in planar imaging for the type of imaging aperture we envisage using. In order to image the object with high resolution, we shall use either a pinhole aperture or a set of long cylindrical holes as the imaging aperture. In the case of a pinhole aperture, Barrett and Swindell (1981) show that the shape of the response function to a point source is invariant with the lateral position of the point source as long as the aperture plane is parallel to the detector plane. They also extend this result to the parallel-hole collimator. In the case of planar objects as assumed, such imaging sytems are shift-invariant. One factor that would affect the shift-invariance property of an imaging system is the object depth, since the size and shape of the response function depends strongly upon the distance of the object to the aperture. This factor is not an issue in planar imaging compared to single-photon emission computed-tomography (SPECT)

imaging where slices of the object are imaged and the dependence of the response function with the location in depth of the slice becomes a predominant factor for large objects.

We describe the emissive object as a function $f(\mathbf{r})$, where $\mathbf{r}$ is a two-dimensional (2D) position vector in the object space. If we denote by $d^2\mathbf{r}$ the area of a source element located at point $\mathbf{r}$, $\overline{f(\mathbf{r})}\, d^2\mathbf{r}$ is the mean number of photons emitted from the elemental area $d^2\mathbf{r}$ per unit time into all space. If we now observe the system for a time T, $T\, \overline{f(\mathbf{r})}\, d^2\mathbf{r}$ represents the mean number of photons emitted from $d^2\mathbf{r}$ during T, if $\overline{f(\mathbf{r})}$ is independent of time. In the same way we described the area of a source element within the object space, we define an area element $d^2\mathbf{r}''$ in the image plane, where $\mathbf{r}''$ is a 2D position vector in the image plane. If we denote by $d\Omega$ the solid angle subtended by the detector element from a point in the object space and if there were no absorbing material in the space between the object and the image plane, the detector element would intercept a fraction $d\Omega/4\pi$ of the radiation emitted from any source element since the source emits into 4π steradians. The mean number of counts emitted by $d^2\mathbf{r}$ and intercepted by $d^2\mathbf{r}''$ during the time T is then given by

$$T\, \overline{f(\mathbf{r})}\, d^2\mathbf{r}\, \frac{d\Omega}{4\pi} = T\, \overline{f(\mathbf{r})}\, d^2\mathbf{r}\, \frac{\cos^3(\theta(\mathbf{r},\mathbf{r}''))}{4\pi(d_1+d_2)^2}\, d^2\mathbf{r}'' \tag{2.1}$$

where d_1 and d_2 are the distances from source to aperture and from aperture to detector, respectively, and $\theta(\mathbf{r},\mathbf{r}'')$ is the angle between the normal to the detector surface and the line of sight from source to detector (see Figure 2.1). If we now take into account the transmitting aperture located between the object and the detector, we shall have a complete description of our image formation system. The principle of image formation is that a photon emitted by the emissive object at location $\mathbf{r}$ in the object space and detected at location $\mathbf{r}''$ in the image space must have passed through the aperture plane at location $\mathbf{r}'$. If

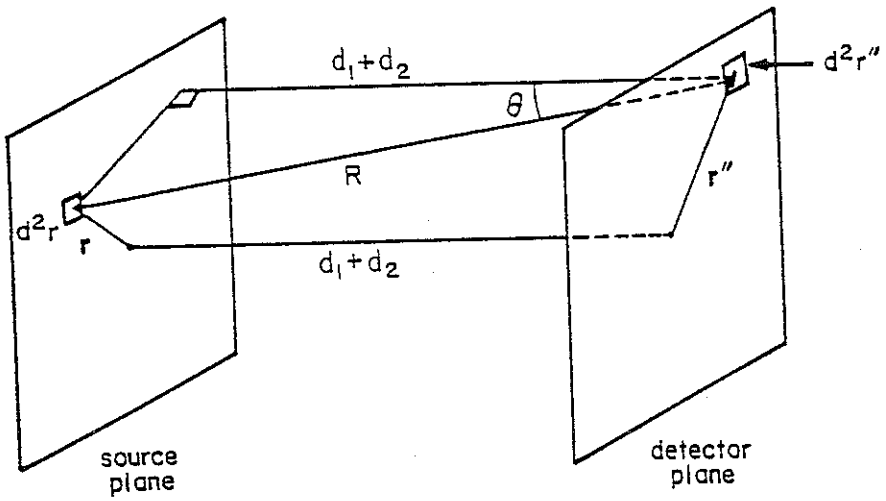


Fig. 2.1. Diagram for solid-angle calculation. The aperture located between the source plane and the detector plane is omitted for clarity.

we denote by $t(\mathbf{r}')$ the transmittance of the aperture at location $\mathbf{r}'$, the fraction of the incident photons that passes through the aperture is $t(\mathbf{r}')$ itself. Moreover, the mean number of photons $\overline{g}(\mathbf{r}'')\, d^2\mathbf{r}''$, intercepted by $d^2\mathbf{r}''$ in the image plane, may come from any location in the object plane and, therefore, an integration over the object space is required. The expression for $\overline{g}(\mathbf{r}'')\, d^2\mathbf{r}''$ is then given by

$$\overline{g}(\mathbf{r}'')\, d^2\mathbf{r}'' = \frac{T}{4\pi(d_1+d_2)^2}\, d^2\mathbf{r}'' \int d^2\mathbf{r}\, \cos^3(\theta(\mathbf{r},\mathbf{r}''))\, \overline{\overline{f}}(\mathbf{r})\, t(\mathbf{r}')\ . \tag{2.2}$$

We shall make the additional assumption that the distance $d_1 + d_2$ is large compared to $|\mathbf{r}|$ and $|\mathbf{r}''|$, so that the angle $\theta(\mathbf{r},\mathbf{r}'')$ may be approximated by 0. The vector $\mathbf{r}'$ can also be expressed as a function of $\mathbf{r}$ and $\mathbf{r}''$, and equation (2.2) reduces to (see Barrett and Swindell 1981)

$$\overline{g}(\mathbf{r}'')\, d^2\mathbf{r}'' = \frac{T}{4\pi(d_1+d_2)^2}\, d^2\mathbf{r}'' \int d^2\mathbf{r}\, \overline{f}(\mathbf{r})\, t(a\mathbf{r}''+b\mathbf{r}) \quad , \tag{2.3}$$

where a and b are relative spacings defined by

$$a = \frac{d_1}{d_1+d_2} \quad \text{and} \quad b = \frac{d_2}{d_1+d_2} \quad . \tag{2.4}$$

Equation (2.3) can now be reduced to a convolution integral by replacing $\overline{f}(\mathbf{r})$ and $t(a\mathbf{r}''+b\mathbf{r})$ by their projections on the image plane. If we denote by $p(\mathbf{r}'')$ the scaled version of $t(\mathbf{r}')$, and if we assume that $d_1 = d_2 = d$ then, $b/a = 1$, and $f(\mathbf{r}'')$ is also the scaled version of $f(\mathbf{r})$. The expression for $\overline{g}(\mathbf{r}'')$ then becomes

$$\overline{g}(\mathbf{r}'') = \frac{T}{4\pi(d_1+d_2)^2}\ \overline{f}(\mathbf{r}'') * p(\mathbf{r}'') = \kappa\, T\, \overline{f}(\mathbf{r}'') * p(\mathbf{r}'') \quad , \tag{2.5}$$

with

$$\kappa = 1/(16\pi d^2) \tag{2.6}$$

The parameter κ will often be referred to as an efficiency factor since it depends only upon the geometry of the imaging system. The variable $\mathbf{r}''$ in equation (2.5) is a dummy variable, and we shall substitute $\mathbf{r}''$ by $\mathbf{r}$ to describe the spatial variation of any function.

An actual measured image $g(\mathbf{r})$ can always be written as $\overline{g}(\mathbf{r}) + n(\mathbf{r})$, $n(\mathbf{r})$ being the noise of the system and $\overline{g}(\mathbf{r})$ the mean count density. Over an ensemble of images, however, we shall distinguish, in general, between two types of noise : uncorrelated Poisson noise, $n(\mathbf{r})$, that is inherent to the quantum nature of the emitted and detected photon stream, and background noise that is a function of the statistical nature of the background inhomogeneity and is responsible for the statistical nature of the object $f(\mathbf{r})$. If we consider a two-hypotheses detection task, for example, we shall call background noise or stuctured noise all the anatomical features in the image that are not the signal to be sought. The term structured noise was given to background noise by Revesz et al. (1974) to differentiate it clearly from quantum noise.

Since we are working with digital images, let us define $g(i) = g(\mathbf{r}_i)\ \epsilon^2$, $\mathbf{r}_i$ being the location of the i^{th} pixel and ϵ^2 its area. The set of, let us say, N numbers g(i) defines an N-dimensional data vector, $\mathbf{g}$, by lexicographically ordering the elements. Boldface type will be used to denote matrices, a vector being a particular case, while normal type will represent scalars. In our simulations the object is a discrete representation of a 2D function $f(\mathbf{r})$, whose pixel values are lexicographically ordered to form an Mx1 column vector $\mathbf{f}$. We can

represent the imaging system by an NxM linear operator **H**, where M is the number of sample points in the object plane and N is the number of detector elements. We can therefore write the imaging equation as

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n} \quad . \tag{2.7}$$

In component form, the imaging equation becomes

$$g(i) = \sum_{j=1}^{M} H(i, j)\, f(j) + n(i) \quad . \tag{2.8}$$

The system operator **H** contains the aperture transmission function $p(\mathbf{r})$, the efficiency factor κ, and the exposure time T.

Statistical Decision Theory

Choosing statistical decision theory to assess image quality is a way to recognize that objective decisions must be based on numbers, not on personal opinions. In a binary decision problem, the decision reduces to the classification of the data at hand into two regions within a so-called feature space.

The first features we may think of are the pixel values of the images themselves, and in that case the feature space is simply the image data space. More generally, however, some other features may be selected to map an N-dimensional data space into a lower dimensional space, allowing for a reduction the dimensionality of the working space. This reduction of the dimensionality of the working space offers, among others, the advantages of faster computations and smaller storage requirements. Feature selection is most often

used, however, to show class separability more effectively. Class separability, however, is not only a function of the selected features, but also of the classifier itself. To resolve this ambiguity, one can decide first of all on a classifier to be used, and then proceed to the selection of the features, given the classifier. In practice, feature selection and classifier design are not so different operations since the classifier may also be conceived of as a feature selector that maps the selected features to a single feature. This single feature will always be referred to as the test statistic and denoted as λ. A summary of the state-of-the-art in discriminant analysis and clustering is given in a Panel on Discriminant Analysis, Classification, and Clustering (1989).

In a binary decision task, we may visualize the two regions to be defined as separated by a boundary such that $\lambda = \lambda_c$, where λ_c is a constant, is the equation of the boundary between the two regions of interest. Thus, the decision process consists of calculating a test statistic λ for each image of the observation set and comparing it to a threshold λ_c. If hypothesis H_2 being true yields, on average, greater values for λ than H_1, then $\mathbf{g}$ is said to belong to class 2 if $\lambda > \lambda_c$ and to class 1 otherwise. We shall now look at how to form the test statistic and how to set the threshold λ_c to define the two regions.

The test statistic is a function of the data $\mathbf{g}$ and of the object and imaging parameters, as well as the noise and object statistics. To form the test statistic, we make the assumption that λ is a linear function of the data. One reason for considering linear test statistics is that they are easy to implement. Another reason is that linear operations on the data may be a good approximation to what the human can do (Fiete et al. 1987), though this assertion has not been, by any means, proven. The general form of the test statistic is thus

$$\lambda = \mathbf{a}^t \mathbf{g} \ , \tag{2.9}$$

where t denotes the transpose operation and **a** is the feature operator. Equation (2.9) shows that λ is a function of several random variables, and therefore it is itself a random variable. In the common case where λ is normally distributed, a general figure of merit is the so-called index of detectability d_a (Simpson and Fitter 1973). The index d_a is defined as

$$[d_a]^2 = \frac{[\overline{\lambda_1} - \overline{\lambda_2}]^2}{P_1 [\sigma_1]^2 + P_2 [\sigma_2]^2} \quad , \tag{2.10}$$

where $\overline{\lambda_k} = \langle \lambda | H_k \rangle$, $[\sigma_k]^2 = \langle (\lambda - \overline{\lambda_k})^2 \rangle$, with the angle brackets indicating the ensemble average over the elements of the k^{th} class (k=1,2), and P_1 and P_2 are the a priori probabilities of occurrence of class 1 and 2, respectively, satisfying $P_1 + P_2 = 1$. In the particular case where $\sigma_1 = \sigma_2 = \sigma_\lambda$, d_a reduces to the index of detectability known as d' (Tanner and Birdsall 1958; and Goodenough et al. 1973) given by

$$[d']^2 = \frac{[\overline{\lambda_1} - \overline{\lambda_2}]^2}{[\sigma_\lambda]^2} \quad . \tag{2.11}$$

Figure 2.2 shows two possible distributions of the outputs λ from the decision device when signal s_1 is present, on the left hand side, and when signal s_2 is present on the right hand side of the decision axis. The indices d_a and d' are measures of the overlap of the two distributions, or in other words of the inherent inability of the system to discriminate signal s_2 from signal s_1.

The concept of a decision threshold is basic to signal detection theory. The choice of a decision threshold allows for the classification of the images and also for the determination of the percent correct within each class. Historically, the performance or accuracy of diagnostic systems was measured by the total percent correct of diagnostic

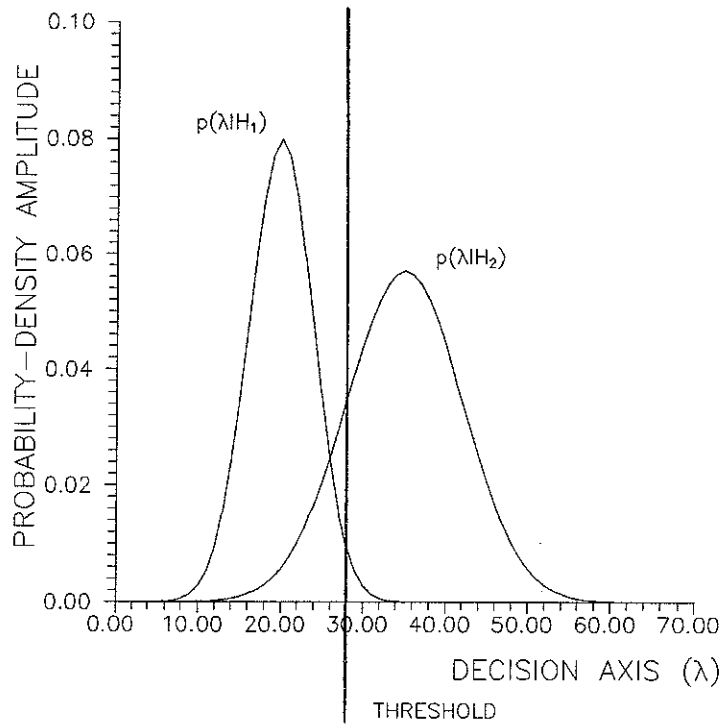


Fig. 2.2. Probability density of λ under hypotheses H_1 and H_2

decisions. This information is important to the evaluation but incomplete since the percent correct is a strong function of disease prevalence. This point is quite important for understanding the necessity of using ROC analysis, and the following case is often chosen as an illustrative example. If only 5% of the patients in a population have a particular disease, then a system may be 95% accurate simply by diagnosing those 5% to be normal patients. Stated more generally, the percent correct is a relative measure of true-positive and true-negative decisions in the tested population, but does not give any information on the proportion of those decisions within each subclass (normal and abnormal) separately. Better indices of performance are, therefore, the TPF and the FPF defined earlier in chapter 1, so that the error of diagnosis for diseased and healthy patients can be measured separately. Two more indices, known as the true-negative fraction (TNF) and the false-negative fraction (FNF), can also be used beside the TPF and FPF. Let us define the TNF as the ratio of the number of normal images for which the test statistic is lower than the threshold to the total number of normal images. Similarly, the FNF is defined as the ratio of the number of abnormal images for which the test statistic is lower than the threshold to the total number of abnormal images. Abnormal and normal images are also termed positive and negative images, respectively, positive (negative) to indicate that the disease is present (absent). With this terminology, hypothesis H_2 can also be termed the positive hypothesis while H_1 refers to the negative hypothesis. The four indices, TPF, FPF, TNF, and FNF are related as FNF=1-TPF and FPF=1-TNF, and all the relevant information with regard to accuracy can be captured by recording one member of each of the complementary pairs of ratios. The usual choice is the TPF and the FPF, also called the "hits" and "false alarms", defined in integral forms as

$$TPF = \int_{\lambda_c}^{\infty} p(\lambda|H_2)\, d\lambda \tag{2.12}$$

and

$$FPF = \int_{\lambda_c}^{\infty} p(\lambda|H_1)\, d\lambda \quad , \tag{2.13}$$

respectively. Due to the inherent inability of a system to completely discriminate between two signals because of the overlap of the probability densities of the test statistic under the two hypotheses, any system, unless perfect, will yield nonzero values of the FNF and FPF. Moreover, to discriminate between positive and negative cases, physicians usually choose different decision thresholds, and the performance of the diagnostic system becomes a strong function of the observers who read the images. Allowing this threshold to vary is allowing for the trading between the fraction of true positive and false positive decisions. By varying the decision threshold, the resulting TPF and FPF will trace the receiver operating characteristic (ROC) curve (Green and Swets 1966; Swets and Pickett 1982). The ensemble of pairs (TPF, FPF) resulting from varying the threshold can be chosen as a measure of the performance since they do not depend on either the prior probability of positive and negative events or the decision threshold. For example, we do not want the performance of a diagnostic system for detecting lesions to be specific to the relative frequency of lesions in the specimens chosen for the test sample.

The ROC curve came out of the developments of radar during world war II (Van Meter and Middleton 1954; Peterson et al. 1954; and Grettenberg 1963), but was intensively used in psychology and psychophysics starting in the 1960s before its integration in medical diagnosis by Lusted (1960). Lusted applied ROC analysis to previously published data on

the detection of pulmonary tuberculosis to show that apparent differences in performance were due to the use of different operating points on the ROC curve, each point corresponding to different values of TPF and FPF. A good review on the application of ROC analysis in radiology has been given in a paper by Metz (1986). The idea behind using ROC analysis is to extract a single number from the ROC curve to provide a measure of performance. The now preferred measure is the proportion of the area of the entire graph that lies beneath the curve (Hanley and Neil 1982, Swets 1988). One is then faced with the task of fitting a continous ROC curve to the data points lying on the curve. A nonparametric estimate of the area under the ROC curve can be obtained by connecting the successive ROC points by lines and using the trapezoidal rule to measure the area beneath the connected points (Green and Swets 1966). A parametric approach to solving this problem is to assume some functional form for the ROC. The most popular functional form in medical imaging is the binormal form where each ROC curve is assumed to have the same functional form as that implied by two normally distributed decision variables with generally different means and variances (Green and Swets 1966; Egan 1975; Swets 1979, 1986). A binormal functional form has the interesting property that all ROC curves are transformed into straight lines if they are plotted on "normal-deviate" axes. If we make the change of variable $z = (\lambda-\overline{\lambda}_1)/\sigma_1$, and if we denote as $\Psi(z)$ the cumulative standard normal distribution defined by

$$\Psi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} dx\, e^{-x^2/2} \quad , \tag{2.14}$$

then instead of graphing the TPF as a function of the FPF as in conventional ROC plots,

"normal-deviate" plots show z_{TPF} as a function of z_{FPF}, where z_{TPF} and z_{FPF} are defined by

$$TPF \equiv \Psi(z_{TPF}) \quad \text{and} \quad FPF \equiv \Psi(z_{FPF}) \quad . \tag{2.15}$$

A binormal ROC curve can be interpreted in terms of an effective decision variable λ arising from two Gaussian density distributions such that

$$p(\lambda|H_1) = \frac{1}{\sqrt{2\pi}} \exp(-\lambda^2/2) \quad , \tag{2.16}$$

and

$$p(\lambda|H_2) = \frac{1}{\sqrt{2\pi}} \exp(-(a-b\lambda)^2/2) \quad . \tag{2.17}$$

In this interpretation, $p(\lambda|H_1)$ has zero mean and unit standard deviation, while $p(\lambda|H_2)$ has a mean of a/b and standard deviation of 1/b . On normal deviates, any binormal ROC plots as a straight line of the form $z_{TPF} = a + b\, z_{FPF}$ with ordinal intercept a and slope b. A maximum-likelihood algorithm is then used to estimate the slope and intercept of the fitted line, as well as the area under the ROC curve, denoted as A_z, where the subscript z serves as a reminder that the measure was obtained from a binormal graph (Dorfman and Alf 1969; Swets and Pickett 1982; Metz 1986). The justification for this binormal assumption is that even when the probability densities are not Gaussians, there may be a monotonic transformation that can transform the distributions into Gaussians (Metz 1986). The ROC itself stays unchanged by any linear transformation even though the probability densities themselves may be changed.

A conventional ROC curve and its equivalent normal-deviate graph are plotted in Figure 2.3. For the conventional plot, the closer the curve to the upper left corner, the

higher the inherent discriminability, since for the same FPF, it yields a higher TPF. Points along the main diagonal will be obtained when the two probability densities perfectly overlap. In general, ROC curves must be specified by two parameters a and b if a full description of the shape of binormal ROC curve is required. In practice, however, we are not so much interested in the shape of the ROC curve, but rather in assigning a single index of performance to the system. This is the case, for example, when one is interested in ranking several diagnostic systems. If we assume that we are in this latter case, figures of merit such as A_z , d′, d_e or d_a will be used instead (Whalen 1971; Lusted 1971). For ROC curves with binormal forms, however, the indices of detectability may be calculated from the values estimated for a and b. The expression for d_a is given by

$$[d_a]^2 = \frac{(\overline{\lambda_2}-\overline{\lambda_1})^2}{(\sigma_1^2 + \sigma_2^2)/2} = \frac{2a^2}{1+b^2} \quad . \tag{2.18}$$

The index of detectability d_e is different from d_a by the fact that it uses the average standard deviation instead of the average variance. Thus, d_e is defined by

$$[d_e]^2 = \frac{(\overline{\lambda_2}-\overline{\lambda_1})^2}{((\sigma_1+ \sigma_2)/2)^2} = \frac{4a^2}{(1+b)^2} \quad . \tag{2.19}$$

If $\sigma_1= \sigma_2= \sigma_\lambda$ then b=1 and both d_a and d_e reduce to the figure of merit d′ related to a by

$$[d']^2 = \frac{(\overline{\lambda_2}-\overline{\lambda_1})^2}{(\sigma_\lambda)^2} = a^2 \tag{2.20}$$

The area under the ROC curve itself can be expressed as a function of the appropriate index of detectabilty d, where d denotes here d_a or d′, as

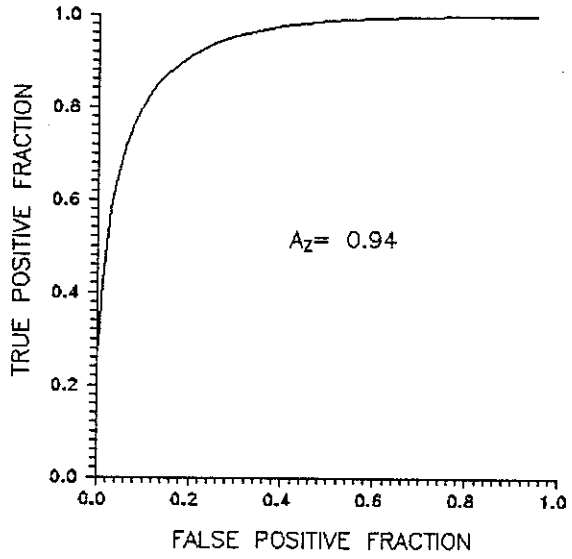


Fig. 2.3. (a) An example of an ROC curve plotted on conventional axes.

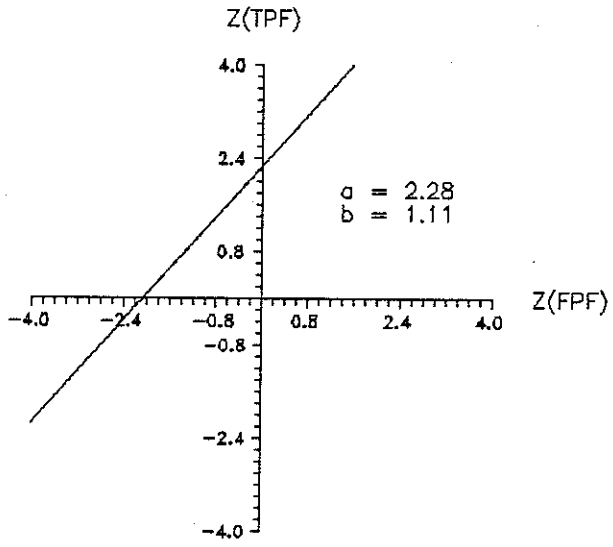


Fig. 2.3. (b) The same curve plotted on "normal-deviate" axes.

$$A_z = \Psi(d/\sqrt{2}) \quad , \tag{2.21}$$

which is most often expressed as

$$A_z = \frac{1}{2}\,(\,1 + \mathrm{erf}(d/2)) \quad , \tag{2.22}$$

where erf(z) is the error function defined as

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z dx\, \exp(-x^2) \quad , \tag{2.23}$$

and erf(z) and $\Psi(z)$ are related by

$$\Psi(z) = \frac{1}{2}\,(\,1 + \mathrm{erf}(z/\sqrt{2}) \quad . \tag{2.24}$$

We shall now define mathematical observers that are commonly used for image quality assessment, that are calculable for the type of task that we are looking at, and give expressions for the test statistics that they use. We choose to review the ideal, the nonprewhitening (npw) ideal, the channelized ideal, and the Hotelling observers, which we shall define as we encounter them below. We shall show that for each observer much insight may be gained by studying the expression of the test statistic itself. As we stressed earlier, a mathematical figure of merit is only useful if it is a good predictor of human performance, therefore, we shall discuss as well the motivations for considering those figures of merit in relation to human performance.

The Ideal Observer

The ideal Bayesian observer is a mathematical observer whose strategy is to base its decision on the likelihood ratio. The likelihood ratio is defined as the ratio of the probability of the data given hypothesis H_2 to the probability of the data given an alternative hypothesis H_1,

$$L = \frac{p(\mathbf{g}|H_2)}{p(\mathbf{g}|H_1)} \quad . \tag{2.25}$$

The output of an ideal observer device is thus simply a number. To be able to compute this number, however, the observer must know the probability densities of the data under each hypothesis. The knowledge of the probability densities requires the knowledge of the statistics of the object to be imaged, the imaging system, and the noise statistics. Under the assumption that the likelihood ratio can be computed, it is then compared to a decision threshold to decide whether hypothesis 1 or 2 is true.

The motivation for using the ideal observer as a standard of comparison to predict the performance of the human observer comes from the fact that the ideal observer has been found to be a good predictor of human performance for well defined and simple tasks such as the detection of a signal known-exactly (SKE) in white Gaussian noise (Barlow 1957, 1958, 1962; Burgess et al. 1979, 1981, 1982; Watson 1983; Judy et al. 1987; Legge et al. 1987; Wagner and Brown 1982). On an absolute scale, however, real observers are much less sensitive than ideal observers. A quantitative measurement of this loss in sensitivity is the efficiency factor, defined as the squared ratio of the detectability of the human d_{human} to the detectability of the ideal observer d_{ideal} (Tanner and Birdsall 1958; Barlow 1978). Efficiencies in the range of 25% to 75% for the human observer relative to the ideal

observer have been reported by Burgess in most of his experiments, under optimal display conditions (Burgess et al. 1981, 1982). We could argue that a 25% efficiency is already in itself a counterargument for using the ideal observer as a standard of comparison for human performance, but there are stronger counterarguments against using the ideal observer. The first of them is that cases of severe discrepancies between human and ideal performances have been demonstrated. Among others, Barlow (1958) showed that the predicted relations between threshold energy and background intensity (the so-called square-root or de Vries-Rose law), target area (Piper's law) and target duration (Pieron's law) hold over limited ranges of those variables. Some of these results are summarized in a review article by Geisler (1989) where many other psychophysical experiments with actual human data compared to the ideal observer predictions can also be found. Another more recent example of discrepancy between the human and the ideal observer is the detection of signals in correlated noise. It is now well established that humans fall far below ideal in correlated noise experiments, as suggested by Hanson (1979) and shown later on by Myers et al. (1985). Another counterargument against using the ideal observer is that the tasks that may be considered have to be quite simple, and thus their relevance to clinical efficacy is not always clear.

These observations on the ideal observer have led to the use of other constructs for optimal detection, such as the npw ideal observer, the channelized ideal observer, and the Hotelling observer. On the other hand, to consider the visual system as an optimum or quasi-optimum detector or equivalently to compare its performance to an optimum observer tells us how far from perfect detection the human stands.

Example 2.1. Detection of A Signal in White Gaussian Noise.

A classical example is the detection of a low-contrast, exactly known signal

superimposed on a constant background and embedded in white Gaussian noise. The likelihood ratio is now the probability that the data are due to the presence of the signal (in noise) divided by the probability that the data are due to the noise alone (Van Trees 1968). If N is the number of pixels, the probability density of $\mathbf{g}$ given hypothesis H_k, $p(\mathbf{g}|H_k)$, is a multivariate Gaussian distribution given by

$$p(\mathbf{g}|H_k) = ((2\pi)^N |\mathbf{K}_k|)^{-1/2} \exp\left[-\frac{1}{2} (\mathbf{g}-\overline{\mathbf{g}_k})^t \mathbf{K}_k^{-1} (\mathbf{g}-\overline{\mathbf{g}_k}) \right] , \qquad (2.26)$$

where $\overline{\mathbf{g}_k} = \langle \mathbf{g}|H_k \rangle = [\langle g(1)|H_k\rangle,...,\langle g(i)|H_k\rangle,...,\langle g(N)|H_k\rangle]^t$, $\mathbf{K}_k$ is the N x N covariance matrix under hypothesis k, with k=1,2, given by $\langle (\mathbf{g}-\overline{\mathbf{g}_k})\ (\mathbf{g}-\overline{\mathbf{g}_k})^t|H_k\rangle$, and $|\mathbf{K}_k|$ is the determinant of $\mathbf{K}_k$. Any monotonic transformation of the test statistic will yield the same performance; therefore, because of the exponential form of the probability density given by equation (2.26), we shall work with the log-likelihood ratio defined as

$$\ell = \ln \frac{p(\mathbf{g}|H_2)}{p(\mathbf{g}|H_1)} , \qquad (2.27)$$

where ln is the natural logarithm. By substituting $p(\mathbf{g}|H_k)$ given by equation (2.26) into equation (2.27), the expression for the log-likelihood ratio becomes

$$\ell = -\frac{1}{2} (\mathbf{g}-\overline{\mathbf{g}_2})^t \mathbf{K}_2^{-1} (\mathbf{g}-\overline{\mathbf{g}_2}) + \frac{1}{2} (\mathbf{g}-\overline{\mathbf{g}_1})^t \mathbf{K}_1^{-1} (\mathbf{g}-\overline{\mathbf{g}_1}) + \text{terms independent of } \mathbf{g}. \qquad (2.28)$$

Since we assumed signal-independent noise, $\mathbf{K}_1 = \mathbf{K}_2 = \mathbf{K}$, and ℓ can be written as

$$\ell = \frac{1}{2}\ \mathbf{g}^t\ \mathbf{K}^{-1}\ (\bar{\mathbf{g}}_2-\bar{\mathbf{g}}_1) + \frac{1}{2}\ (\bar{\mathbf{g}}_2-\bar{\mathbf{g}}_1)^t\ \mathbf{K}^{-1}\ \mathbf{g} + \text{terms independent of } \mathbf{g}. \qquad (2.29)$$

Since the covariance matrix is symmetric, its inverse is symmetric, and if we denote by $\Delta\bar{\mathbf{g}}$ the difference of the mean images $\bar{\mathbf{g}}_2 - \bar{\mathbf{g}}_1$, then

$$(\Delta\bar{\mathbf{g}})^t\ \mathbf{K}^{-1}\ \mathbf{g} = \mathbf{g}^t\ \mathbf{K}^{-1}\ \Delta\bar{\mathbf{g}}\ , \qquad (2.30)$$

and ℓ simplifies to

$$\ell = (\Delta\bar{\mathbf{g}})^t\ \mathbf{K}^{-1}\ \mathbf{g}\ , \qquad (2.31)$$

where we have dropped the terms independent of g. The test statistic is found to be linear with the data. Thus, the ideal-observer discriminant is not only the optimum discriminant but also the optimum linear discriminant for the detection of a signal in signal-independent white Gaussian noise.

The covariance matrix K being positive definite, if the inverse $\mathbf{K}^{-1}$ of the covariance matrix exists then $\mathbf{K}^{-1/2}$ exists, and we can gain more insight into the expression for ℓ by writing equation (2.31) as

$$\ell = (\mathbf{K}^{-1/2}\ \Delta\bar{\mathbf{g}})^t\ (\mathbf{K}^{-1/2}\ \mathbf{g})\ . \qquad (2.32)$$

This expression shows that the data are first filtered by the inverse of the square root of the covariance matrix, an operation known as prewhitening. The ideal Baysesian observer performs, then, a matched-filtering operation, correlating the prewhitened received data $\mathbf{K}^{-1/2}\mathbf{g}$ with the expected signal $\Delta\bar{\mathbf{g}}$ seen through the prewhitening filter. For this reason this ideal observer is often referred to as the prewhitening matched-filter. A scheme of the ideal observer, as just defined, is given in Figure 2.4. By identification of equation (2.9)

and (2.31), the feature operator for the ideal observer is given by

$$\mathbf{a} = \mathbf{K}^{-1}\,\Delta\bar{\mathbf{g}}\ , \tag{2.33}$$

and for the two-hypotheses detection task, the detectability index d_a given by equation (2.10) becomes

$$[d_a]^2 = \frac{[\bar{\ell}_2 - \bar{\ell}_1]^2}{P_1\,[\sigma_1]^2 + P_2\,[\sigma_2]^2}\ , \tag{2.34}$$

where $[\sigma_k]^2 = [\langle(\ell-\bar{\ell}_k)^2|H_k\rangle$ with k=1,2, and $\bar{\ell}_k = \langle\ell|H_k\rangle$. For signal-independent noise, $\sigma_1 = \sigma_2 = \sigma_\ell$, and the variance reduces to a single term. If we denote $(\langle\ell|H_2\rangle-\langle\ell|H_1\rangle)$ as $\Delta\bar{\ell}$, then $\Delta\bar{\ell}$ and σ_ℓ may be calculated using the expression for the test statistic given by equation (2.31). The results are

$$\begin{aligned}
\Delta\bar{\ell} &= \langle\,\ell|H_2\,\rangle - \langle\,\ell|H_1\,\rangle \\
&= \langle\,(\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\mathbf{g}\,|H_2\,\rangle - \langle\,(\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\mathbf{g}\,|H_1\,\rangle \\
&= (\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\langle\,\mathbf{g}|H_2\,\rangle - (\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\langle\,\mathbf{g}|H_1\,\rangle \\
&= (\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\Delta\bar{\mathbf{g}}
\end{aligned} \tag{2.35}$$

$$\begin{aligned}
[\sigma_\ell]^2 &= \langle\,(\ell-\bar{\ell}_2)^2|H_2\,\rangle \\
&= \langle\,(\,(\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\mathbf{g} - (\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\langle\mathbf{g}|H_2\rangle)^2\,|H_2\,\rangle \\
&= \langle\,(\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,(\mathbf{g}-\bar{\mathbf{g}}_2)\,(\mathbf{g}-\bar{\mathbf{g}}_2)^t\,\mathbf{K}^{-1}\,\Delta\bar{\mathbf{g}}\,|H_2\,\rangle \\
&= (\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\langle\,(\mathbf{g}-\bar{\mathbf{g}}_2)\,(\mathbf{g}-\bar{\mathbf{g}}_2)^t|H_2\,\rangle\,\mathbf{K}^{-1}\,\Delta\bar{\mathbf{g}} \\
&= (\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\mathbf{K}\,\mathbf{K}^{-1}\,\Delta\bar{\mathbf{g}} \\
&= (\Delta\bar{\mathbf{g}})^t\,\mathbf{K}^{-1}\,\Delta\bar{\mathbf{g}}\ .
\end{aligned} \tag{2.36}$$

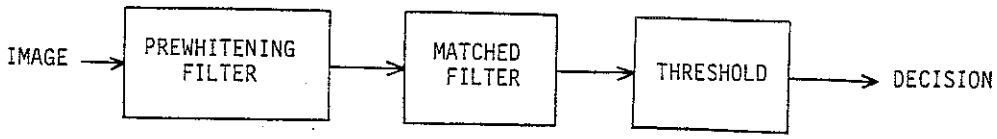


Fig. 2.4. Model of the ideal observer

Equations (2.35) and (2.36) show that $[\sigma_\ell]^2$ equal $\Delta\bar{\ell}$. This equality results from the assumption that the noise is signal independent. For the ideal observer, d_a will be denoted as d_{ideal}, and the substitution of $\Delta\bar{\ell}$ and σ_ℓ into equation (2.34) leads to

$$[d_{ideal}]^2 = (\Delta\bar{\mathbf{g}})^t \, \mathbf{K}^{-1} \, \Delta\bar{\mathbf{g}} \quad . \tag{2.37}$$

Example 2.2 Detection of a Signal Embedded in Uncorrelated Poisson Noise.

The probability density of the data **g** given hypothesis H_k (k=1,2) is now given by a multivariate Poisson distribution. Because of the uncorrelated nature of Poisson noise, the multivariate probability density is simply the product over the pixel values of Poisson probability densities. It can be expressed as

$$p(\mathbf{g}|H_k) = \prod_{i=1}^{N} \frac{\bar{g}_k(i)^{g(i)}}{g(i)!} \exp(- \bar{g}_k(i)) \tag{2.38}$$

The likelihood ratio is then given by

$$\ell = \ln \frac{p(\mathbf{g}|H_2)}{p(\mathbf{g}|H_1)}$$

$$= \sum_{i=1}^{N} g(i) \left[\ln \frac{\bar{g}_2(i)}{\bar{g}_1(i)} \right] + \text{terms independent of } g(i) \quad . \tag{2.39}$$

In the low-contrast approximation we can write $\bar{g}_2(i) = \bar{g}_1(i) + \Delta\bar{g}$, with $\Delta\bar{g} << 1$, and the expression for ℓ becomes

$$\ell = \sum_{i=1}^{N} g(i) \left[\ln \left(1 + \frac{\Delta\bar{g}(i)}{\bar{g}_1(i)} \right) \right]$$

$$\simeq \sum_{i=1}^{N} g(i) \frac{\Delta\bar{g}(i)}{\bar{g}_1(i)}$$

$$= (\Delta\bar{\mathbf{g}})^t \mathbf{K}_1^{-1} \mathbf{g} \quad , \tag{2.40}$$

since for uncorrelated Poisson noise the covariance matrices $\mathbf{K}_k$ (k=1,2) are diagonal and their elements may be written as

$$\mathbf{K}_k(i,j) = \bar{g}_k(i) \delta_{ij} \quad . \tag{2.41}$$

Moreover, in the low-contrast approximation, $\mathbf{K}_1 = \mathbf{K}_2 = \mathbf{K}$, and equation (2.40) becomes

$$\ell = (\Delta\bar{\mathbf{g}})^t \mathbf{K}^{-1} \mathbf{g} \quad , \tag{2.42}$$

which is also the expression for the test statistic given by equation (2.31) that we found for the case of signal-independent white Gaussian noise. The detectability index associated with ℓ is then again given by

$$d_{ideal} = (\Delta\bar{\mathbf{g}})^t \mathbf{K}^{-1} \Delta\bar{\mathbf{g}}$$

$$= \sum_{i=1}^{N} \frac{[\Delta\bar{g}(i)]^2}{\bar{g}_1(i)} \quad . \tag{2.43}$$

The Nonprewhitening Ideal Observer

The npw ideal observer and the ideal observer differ in their ability to take into account noise correlations. The npw ideal observer, also referred to as the quasi-ideal observer (Wagner et al. 1985), is a suboptimal observer which assumes white uncorrelated noise. The assumption of uncorrelated noise is not sufficient since the noise may be uncorrelated but not white. The strategy of the npw ideal observer, while performing a discrimination task, is to apply a mask or template on the image in order to perform a cross-correlation of the expected signal with the image data. The test statistic is then given by

$$\lambda = (\Delta\bar{\mathbf{g}})^t \ \mathbf{g} \quad , \tag{2.44}$$

where $\Delta\bar{\mathbf{g}}$ is the expected signal or template laid on the image data $\mathbf{g}$. This template-matching strategy, described in the space domain by equation (2.44), may also be performed in the Fourier domain. In the Fourier domain, the data transform is multiplied by a matched filter, that is the complex conjugate of the transform of the expected signal. We shall represent the discrete Fourier operator by $\mathbf{F}$. We use $\mathbf{F}$ to map the expected signal into the Fourier domain as follows

$$\mathbf{F}\,\Delta\bar{\mathbf{g}} = \Delta\bar{\tilde{\mathbf{g}}} \quad , \tag{2.45}$$

where $\Delta\bar{\tilde{\mathbf{g}}}(m)$ is the Fourier amplitude of $\Delta\bar{\tilde{\mathbf{g}}}$ at frequency ρ_m. Note that the Fourier operator, as we have defined it, maps a lexicographically ordered vector from the space domain to a lexicographically ordered vector in the Fourier domain. The test statistic, expressed in the Fourier domain and denoted as λ_F, is then given by

$$\lambda_F = (\Delta\tilde{\bar{g}})^\dagger \tilde{g}$$
$$= (\Delta\tilde{\bar{g}})^t \tilde{g} \quad , \tag{2.46}$$

and since $\Delta\bar{g}$ is real, the transpose conjugate denoted by the dagger symbol $\dagger$ becomes simply the transpose. Moreover, we note that, according to Parseval's theorem (Gaskill 1978), λ_F given by equation (2.46) equals λ given by equation (2.44). This equality follows since $\mathbf{F}$ is a unitary operator satisfying the equation

$$\mathbf{F}^\dagger \mathbf{F} = \mathbf{I} \quad , \tag{2.47}$$

where $\mathbf{I}$ is the identity matrix.

The npw ideal observer is said to be suboptimal because it treats the noise as if it were white. For white noise, $\mathbf{K} = \sigma^2 \mathbf{I}$, and the npw ideal observer becomes equivalent to the ideal observer.

The npw ideal observer is of great importance since human performance has been interpreted using its strategy in two cases. The first case was for the detection of signals known exactly embedded in signal-independent, white, Gaussian noise (Barnard 1972; Burgess 1979; Wagner 1985), while the second case was presented by Myers et al. (1985) who showed that for signals embedded in correlated noise, the performance of the human observer could be interpreted in terms of the npw ideal observer rather than the ideal observer. They also showed that the npw ideal observer was equivalent to the channelized ideal observer.

The Channelized Ideal Observer

The concept of the channelized ideal observer comes from findings in the

psychophysics and the physiology of human vision. We shall review some of the experimental results that show that retinal images are decomposed into several frequency channels at the visual cortex level.

The existence of frequency channels in human vision has been demonstrated by the pioneer works of Campbell and Robson (1968). Their experiments to prove the existence of channels were done in parallel with other experiments to measure the contrast sensitivity function of the human observer (DePalma and Lowry 1962; Campbell and Green 1965; Campbell et al. 1969; Kulikowski and King-Smith 1973). Both sets of experiments used vertical sinusoidal gratings of different spatial frequencies ω as the stimuli. The contrast sensitivity function is defined as the inverse of the contrast threshold, which is the minimum contrast of the grating necessary to detect it on a uniform background. The outcome of those experiments was that the contrast sensitivity drops at high frequencies of the grating due to the finite resolution set by the eye, as well as at very low frequencies where the drop could be explained by some inhibition process taking place at the neural level. This inhibition process, also referred to as the brightness-contrast effect (Ratliff 1965), will prove to be of great interest to us as we study the impact of long-tailed PSFs on human detectability.

These observed variations in performance of the human observer due to different frequencies of the stimuli led to some further experiments based on adaptation techniques. If a stimulus of frequency ω_0 is shown to an observer for a prolonged time, the contrast sensitivity decreases. This is the adaptation process. Campbell and Robson (1968) showed that this adaptation process holds for any stimulus whose frequency is in a frequency band around ω_0, but that outside this band the sensitivity is not affected. Further experiments were then presented by Sachs, Nachmias and Robson (1971) using compound gratings. They found that, for a two-frequency grating, the two frequencies were detected separately unless

they differ by less than a certain amount. The results from these different experiments demonstrate that at some stage in the visual process the visual information is processed separately in different frequency bands. To model this process, Campbell and Robson assumed that the retinal image is decomposed through independent band-pass linear filters. These band-pass filters are now referred to as frequency channels.

Some other groups of researchers (Georgeson 1975; Nachmias and Weber 1975) have further characterized these channels by showing that the bandwidth of these filters is around an octave, which yields filters of constant bandwith on a logarithmic scale. On the other hand, earlier research by Hubel and Wiesel (1962) had been carried out on cells belonging to the visual cortex of cats and monkeys who have a visual cortex similar to the human. They discovered a class of cells, called simple cells, whose response is a function of the frequency of the visual stimuli. Their responses have been shown by Maffei and Fiorentini (1976) to be linear. The observed linearity allows them to model these cells as linear filters. Finally, Andrew and Pollen (1979), Maffei et al. (1979), and De Valois et al. (1979) showed that the bandwidth of simple cells ranges from 0.6 to 2.0 octaves with an average value of 1.3 octaves, findings that agree with those found by Georgeson (1975) and Nachmias and Weber (1975). The findings of channels in the human vision process seem to indicate that the low-level processing of the visual cortex can be modeled by multiple frequency channels.

The idea behind the channelized ideal observer is to add frequency-selective channels to the ideal-observer construct. The modification of the ideal-observer calculations to implement the concept of channels has been investigated by Myers et al. (1987). As for the ideal observer, one can compute the likelihood ratio of the data given two hypotheses, but now the likelihood ratio of the data is taken after the data have been processed by the frequency channels of the observer. A model of the channelized ideal-observer for the case

of white, Gaussian noise has been derived by Myers et al. (1987), where details on the implementation can be found.

The Hotelling Observer

The Hotelling observer is a linear classifier, first presented in a classic paper by Harold Hotelling (1931). This classifier is sometimes better known from the work of Fisher (1936). We shall review the two-class problem relevant to the studies presented in this work; the generalization to M classes may be found in texts on statistical pattern recognition such as that by Fukunaga (1972). The use of the Hotelling observer as a means to assess image quality has been demonstrated by Gu and Lee (1984), Barrett et al. (1985, 1986) and Smith and Barrett (1986). To be a good figure of merit, the Hotelling observer must correlate with human performance. Some examples of this correlation have been presented by Fiete et al. (1987) and White et al. (1989), who have shown a high correlation between the human and Hotelling observer for detection of tumors, variable in size, contrast, and position, in simulated liver scans imaged through different collimators.

We shall show that the Hotelling observer is very intuitive and a simpler construct than the ideal observer since it does not require the knowledge of the probability densities, but is based, instead, on the so-called scatter matrices, $\mathbf{S}_1$ and $\mathbf{S}_2$, that describe the first-order and second-order statistics of the two classes. The interclass scatter matrix $\mathbf{S}_1$, a measure of the separation of the class means, is defined as

$$\mathbf{S}_1 = \sum_{k=1}^{2} P_k \, [\bar{\mathbf{g}}_k - \bar{\mathbf{g}}_0 \,] \, [\bar{\mathbf{g}}_k - \bar{\mathbf{g}}_0 \,]^t \quad , \tag{2.48}$$

where the subscript k and angle brackets indicate the ensemble average over the elements of

the k^{th} class, P_k is the a priori probability of occurrence of class k, and $\overline{g_0}$ is the grand mean given by

$$\overline{g_0} = P_1 \overline{g_1} + P_2 \overline{g_2} \quad . \tag{2.49}$$

By substituting equation (2.49) into equation (2.48), S_1 may also be written as

$$S_1 = P_1 P_2 (\overline{g_2} - \overline{g_1}) (\overline{g_2} - \overline{g_1})^t \quad . \tag{2.50}$$

The intraclass scatter matrix S_2, a measure of the scatter of samples around their class means, is the average covariance matrix given by

$$S_2 = \sum_{k=1}^{2} P_k K_k \quad , \tag{2.51}$$

where K_k is the covariance matrix for class k, defined as

$$K_k = \langle [g - \overline{g_k}] [g - \overline{g_k}]^t \rangle \quad . \tag{2.52}$$

In order to formulate an index of classification, we define the Hotelling trace criterion (HTC) as

$$J = tr (S_2^{-1} S_1) \quad , \tag{2.53}$$

where tr (A) denotes the trace (sum of the diagonal elements) of the matrix A. The beauty of the HTC is that it requires the knowledge of only S_1 and S_2 to calculate how separable the classes are, while the Bayesian ideal-observer approach requires the knowledge of the

probability densities. Moreover, the value of J increases as the class means get farther apart, which corresponds to larger values of $\mathbf{S}_1$, and (or) the individual classes cluster tighter to their mean values, which corresponds to smaller values of $\mathbf{S}_2$.

Given the Hotelling classifier, we must find the feature or test statistic λ_{Hot} that maximizes J. We assume that we select λ_{Hot} by applying a 1xN transformation matrix $\mathbf{a}^t$ to the data vector $\mathbf{g}$,

$$\lambda_{Hot} = \mathbf{a}^t \mathbf{g} \quad . \tag{2.54}$$

To simplify the notation we shall refer to λ instead of λ_{Hot}. The scatter matrices in the feature space, denoted as $\mathbf{S}_{1\lambda}$ and $\mathbf{S}_{2\lambda}$ are in fact 1x1 matrices and are given by

$$\begin{aligned} \mathbf{S}_{1\lambda} &= \sum_{k=1}^{2} P_k [\, \overline{\lambda_k} - \lambda_0 \,][\, \overline{\lambda_k} - \lambda_0 \,]^t \\ &= \sum_{k=1}^{2} P_k \, [\, \overline{\lambda_k} - \lambda_0 \,]^2 \end{aligned} \tag{2.55}$$

$$\begin{aligned} \mathbf{S}_{2\lambda} &= \sum_{k=1}^{2} P_k \langle [\, \lambda - \overline{\lambda_k} \,][\, \lambda - \overline{\lambda_k} \,]^t \rangle \\ &= \sum_{k=1}^{2} P_k \, \langle [\, \lambda - \overline{\lambda_k} \,]^2 \rangle \quad , \end{aligned} \tag{2.56}$$

where $\lambda_0 = P_1 \, \overline{\lambda_1} + P_2 \, \overline{\lambda_2}$. By using equations (2.56), (2.55), (2.54), (2.51), (2.52), and (2.48), it can be shown that $\mathbf{S}_{1\lambda}$ and $\mathbf{S}_{2\lambda}$ are related to $\mathbf{S}_1$ and $\mathbf{S}_2$ by

$$\mathbf{S}_{1\lambda} = \mathbf{a}^t \, \mathbf{S}_1 \, \mathbf{a} \quad \text{and} \quad \mathbf{S}_{2\lambda} = \mathbf{a}^t \, \mathbf{S}_2 \, \mathbf{a} \quad . \tag{2.57}$$

Let α_p, ϕ_p (p=1,..,P) be the eigenvalues and eigenvectors of $\mathbf{S}_2^{-1}\mathbf{S}_1$, respectively. The trace may then be expressed as a function of the eigenvalues α_p as

$$J = \mathrm{tr}\,(\mathbf{S}_2^{-1}\,\mathbf{S}_1) = \sum_{p=1}^{P} \alpha_p \quad . \tag{2.58}$$

Since we showed that $\mathbf{S}_1$ can be written as $\mathbf{y}\,\mathbf{y}^t$ (see equation (2.50)), the rank of $\mathbf{S}_1$ is one and a single eigenvalue is nonzero. The sum over the eigenvalues given by equation (2.58) reduces then to α_1. On the other hand in feature space, since $\mathbf{S}_{1\lambda}$ and $\mathbf{S}_{2\lambda}$ are 1x1 matrices, the Hotelling trace denoted as J_λ is a scalar, and we shall write

$$J_\lambda \;=\; \mathbf{S}_{2\lambda}^{-1}\,\mathbf{S}_{1\lambda} \;=\; \beta_1 \quad . \tag{2.59}$$

Since $\mathbf{S}_1$ and $\mathbf{S}_2$ are Hermitian matrices, $\mathbf{S}_2^{-1}\mathbf{S}_1$ is also an Hermitian matrix, and the eigenvectors in feature space form a complete orthogonal basis set. Therefore, the feature operator can be expressed as a linear combination of the eigenvectors of $\mathbf{S}_2^{-1}\mathbf{S}_1$ as

$$\mathbf{a} \;=\; \sum_{p=1}^{P} \gamma_p\;\phi_p \quad , \tag{2.60}$$

and we shall now show that $\mathbf{a} = \phi_1$ leads to a maximum value for J_λ. By substituting equation (2.60) into the expressions for $\mathbf{S}_{1\lambda}$ and $\mathbf{S}_{2\lambda}$ given by equation (2.57), we can write J_λ as

$$J_\lambda = \frac{J \sum_{\ell} \gamma_\ell \gamma_1 (\phi_\ell)^t S_2 \phi_1}{\sum_{\ell,p} \gamma_\ell \gamma_p (\phi_\ell)^t S_2 \phi_p}$$

$$= J \frac{(x,\omega)}{(\omega,\omega)} \quad , \tag{2.61}$$

where the parentheses denote the scalar product of two vectors, and x and ω are defined as

$$x = \gamma_1 S_2^{1/2} \phi_1 \tag{2.62}$$

$$\omega = \sum_{\ell} \gamma_\ell S_2^{1/2} \phi_\ell \quad . \tag{2.63}$$

If for $\ell = 1$, $\gamma_\ell = \gamma_1 = 1$ and for $\ell \neq 1$, $\gamma_\ell = 0$, then $\omega = x$ and $J_\lambda = J$. In any other cases, $x \neq \omega$ but (x,ω) is always equal to (x,x) since the eigenvectors of an orthogonal eigenvector basis set are orthogonal if their associated eigenvalues are different, as is the case for ϕ_1 and the other eigenvectors of the basis set. Equation (2.61) then becomes

$$J_\lambda = J \frac{||x||^2}{||\omega||^2} \leq J \quad . \tag{2.64}$$

We have then shown that $a = \phi_1$ maximizes J_λ. The eigenvalue α_1 (or J) may then be computed by using equation (2.58) as

$$\begin{aligned}\alpha_1 &= \mathrm{tr}\,(\mathbf{S}_2^{-1}\mathbf{S}_1) \\ &= \mathrm{tr}\,(\,\mathbf{S}_2^{-1}\,\{\,P_1P_2\,(\bar{g}_2 - \bar{g}_1)\,(\bar{g}_2 - \bar{g}_1)^t\,\}\,) \\ &= P_1P_2\,\mathrm{tr}\,(\,\mathbf{S}_2^{-1}\,\,(\bar{g}_2 - \bar{g}_1)\,(\bar{g}_2 - \bar{g}_1)^t\,\,) \\ &= P_1P_2\,\,(\bar{g}_2 - \bar{g}_1)^t\,\,\mathbf{S}_2^{-1}\,\,(\bar{g}_2 - \bar{g}_1)\quad ,\end{aligned} \tag{2.65}$$

and the feature operator is then given by

$$\mathbf{a} = \mathbf{S}_2^{-1}\,(\bar{g}_2 - \bar{g}_1)\quad . \tag{2.66}$$

The validity of the expression for **a** can be shown by substituting equations (2.50) and (2.66) into the eigenvalue equation for $\mathbf{S}_2^{-1}\mathbf{S}_1$,

$$\begin{aligned}\mathbf{S}_2^{-1}\,\mathbf{S}_1\,\mathbf{a} &= \alpha_1\,\mathbf{a} \\ \mathbf{S}_2^{-1}\,P_1P_2\,(\bar{g}_2-\bar{g}_1)\,(\bar{g}_2-\bar{g}_1)^t\,\,\mathbf{S}_2^{-1}\,(\bar{g}_2-\bar{g}_1) &= P_1P_2\,(\bar{g}_2-\bar{g}_1)^t\,\mathbf{S}_2^{-1}\,(\bar{g}_2-\bar{g}_1)\,\mathbf{a} \\ \mathbf{S}_2^{-1}\,(\bar{g}_2-\bar{g}_1) &= \mathbf{a}\quad ,\end{aligned} \tag{2.67}$$

and, by combining equations (2.54) and (2.67), the test statistic then becomes

$$\lambda_{\mathrm{Hot}} = [\,\bar{g}_2 - \bar{g}_1\,]^t\,\,\mathbf{S}_2^{-1}\,\,\mathbf{g}\quad . \tag{2.68}$$

The test statistic given by equation (2.68) is, in fact, a generalization of the prewhitening matched filter. The presence of $\mathbf{S}_2^{-1}$ in the expression for λ means that the Hotelling observer takes into account not only the randomness due to the presence of noise in the data but also the randomness due to object variability. For no randomness in the

object itself and in the low-contrast approximation, $\mathbf{S}_2$ reduces to $\mathbf{K}$, the noise covariance matrix introduced in equation (2.52). The test statistic λ_{Hot}, despite its analogy to the log-likelihood ratio, is suboptimal because λ_{Hot} is linear with the data, while the log-likelihood ratio is usually non-linear.

We shall now show that the HTC is also closely related to the detectability index d_a defined by equation (2.34). Since we saw that $J = J_\lambda$, we may write J as

$$J = \mathbf{S}_{2\lambda}^{-1} \mathbf{S}_{1\lambda} \quad , \tag{2.69}$$

and in the case where the probability densities of λ for each class of objects are Gaussians, we get

$$J = \frac{P_1(\overline{\lambda_1} - \lambda_0)^2 + P_2(\overline{\lambda_2} - \lambda_0)^2}{P_1 [\sigma_1]^2 + P_2 [\sigma_2]^2} \quad . \tag{2.70}$$

If we substitute λ_0 by $P_1\overline{\lambda_1} + P_2\overline{\lambda_2}$ and use $P_1 + P_2 = 1$ then,

$$J = \frac{P_1P_2 (\overline{\lambda_1} - \overline{\lambda_2})^2}{P_1 [\sigma_1]^2 + P_2 [\sigma_2]^2} \quad , \tag{2.71}$$

or

$$J = P_1P_2 [d_a]^2 \quad . \tag{2.72}$$

CHAPTER3

LONG-TAILED POINT SPREAD FUNCTIONS

A perfect imaging system would give, as an image of a point source, a point image identical to the point source. This is, however, idealistic, and in reality many factors contribute to the degradation of a point image, in other words to the mislocation of the detected photons. A measure of this degradation is given by the point spread function (PSF), which is defined as the average spatial distribution of a very large number of detected photons emitted by a single point-like object. Ideally, the PSF is symmetric and centered around its expected location defined geometrically by the point object and the center of the aperture. Moreover, it is sharply peaked for high-resolution systems. For real systems, however, the PSF can be altered in several ways; its center of symmetry can be shifted from its expected value, while its peak can become broader and assymmetric.

In conventional radionuclide imaging, the imaging device is usually an Anger-type camera, which consists of a collimator, a scintillation crystal, a set of photomultiplier tubes sealed to a glass window, and some analog circuitry used to record the position of the gamma-ray events that are being detected. A collimator placed between the object and the detector serves as the imaging device. The bores of the collimator, usually made of lead alloy, should be small in diameter and extend over an appreciable length of material to be able to tranform the nondirectional incoming radiation into multiple, highly directional photon beams which permit high resolution. A gain in resolution is always, however, accompanied by a loss in sensitivity of the detection device, that is, a longer exposure time is now required to reach equivalent number of detected photons. This loss in sensitivity can be, however, compensated for by increasing the strengh of the emitted source, but only up

to the maximum dose of radiation allowed to be injected into the patient's body. Thus, in practice, the size and the length of the bores are chosen to satisfy the best trade-off between resolution and sensitivity, and the resulting loss in resolution often leads to some broadening of the central peak of the PSF. Moreover, physical processes, such as Compton scattering in the body, can cause even further broadening of the PSF, adding long spatial tails to its compact central peak. We refer to such PSFs as long-tailed. Long-tailed PSFs can also arise due to veiling glare in image intensifiers, or due to septal penetration in nuclear imaging, and are commonly encountered in image science where they frequently constitute an important source of contrast reduction and resolution loss.

We are interested in investigating the effect that this type of PSF has on human detection and whether any improvement in performance can be gained by deconvolving the tails of the PSF. We shall find out what figures of merit are useful for characterizing human performance by comparing the performance of the human observer evaluated through psychophysical studies with the performance of statistical observers derived from statistical decision theory. Thus, we shall also refer to these statistical observers as standards of comparison. For the ideal observer, we shall show that the performance is independent of any linear, invertible deconvolution filter. The question then arises whether the ideal observer is a good predictor of human performance, and we shall investigate its performance, as well as the performances of the npw ideal and the Hotelling observers, before and after deconvolution. We shall also investigate the performance of new statistical observers; we shall define the no-low-frequency (nlf) ideal observer, which is a pseudoideal observer that filters out some of the low-frequency components of the image before performing the detection task. We shall also introduce a combination of the npw ideal observer with the nlf ideal observer, referred to as the nlp ideal observer.

Image Simulations

The nuclear imaging system shown schematically in Figure 3.1 (a) consists of a radiating object, an image-forming element and a detector. Let us describe the object $f(\mathbf{r})$ as the superposition of a mean background $b(\mathbf{r})$ and a low-contrast signal $s(\mathbf{r})$, where $\mathbf{r}$ denotes a 2D position vector. The overall imaging system may be characterized by a PSF $p_1(\mathbf{r})$, which we take to be the sum of a sharply peaked function, also referred to as the core $c(\mathbf{r})$ of the PSF, and an extended tail $t(\mathbf{r})$. The subcript "1" on $p_1(\mathbf{r})$ refers to the primary or imaging PSF, since we shall introduce later another PSF $p_2(\mathbf{r})$ related to the filtering process. If $f(\mathbf{r})$ is the mean number of counts emitted per unit area and per unit time from the object plane, the mean number of photons detected per unit area during the imaging time T would be, for a perfect imaging system, κ T $f(\mathbf{r})$, where κ is an efficiency factor taking into account the geometry of the system. An expression for κ was derived in chapter 2 (equations 2.1 - 2.6) for the case of a pinhole aperture. If the imaging system is not perfect, the PSF departs from a delta function, and the number of photons detected per unit area during the imaging time T can be evaluated, using equation (2.6), by the following expression

$$
\begin{aligned}
g(\mathbf{r}) &= (\kappa T)\,(f(\mathbf{r}) * p_1(\mathbf{r})) + n(\mathbf{r}) \\
&= (\kappa T)\,([s(\mathbf{r}) + b(\mathbf{r})] * [c(\mathbf{r}) + t(\mathbf{r})]) + n(\mathbf{r}) \\
&= (\kappa T)([s(\mathbf{r}) * c(\mathbf{r})] + [b(\mathbf{r}) * c(\mathbf{r})] + [s(\mathbf{r}) * t(\mathbf{r})] + [b(\mathbf{r}) * t(\mathbf{r})]) + n(\mathbf{r}) . \quad (3.1)
\end{aligned}
$$

To investigate the impact that long-tailed PSFs have on the detectabilities achievable by different observers, we vary the extent of the tail of the PSF by choosing ten different widths of the tail of the PSF for a fixed value of the contrast of the core of the PSF relative

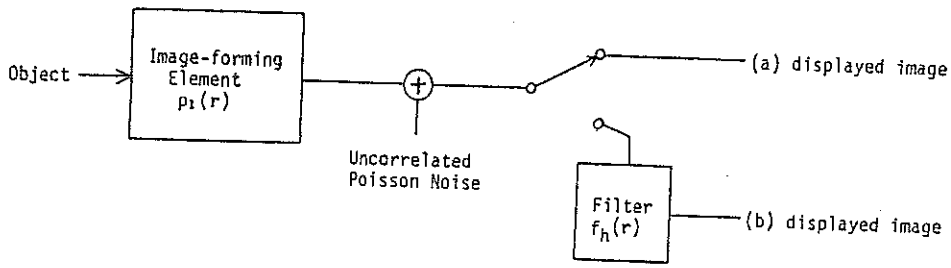


Fig. 3.1. Model of the imaging system used to generate computer-simulated images. (a) The object is convolved with the psf $p_1(\mathbf{r})$ and Poisson noise is added to the data. (b) The second function $f_h(\mathbf{r})$ serves as a high-pass filter to deconvolve the images.

to the tail. Thus, we generate ten different sets of images, each corresponding to different parameters of the long-tailed PSF. Each PSF is expressed as the sum of two Gaussians, as suggested, for example, by Tsui et al. (1978), to model the core and the tail of the PSF, respectively. Their expression is given by

$$p_1(\mathbf{r}) = a_c \exp(-r^2/r_c^2) + a_t \exp(-r^2/r_t^2) \quad , \tag{3.2}$$

where a_c and a_t are the amplitudes of the core and the tail of the PSF, respectively, and r_c and r_t are the half widths at 1/e of the peak intensity of their respective distributions. The width of the core of the PSF is held constant across the ten sets of images, so that the degradation from one set to the next is essentially due to the broadening of the PSF resulting from adding tails to the core. For the same reason, the contrast of the PSF, defined as

$$C_{psf} = \frac{a_c}{a_t} \quad , \tag{3.3}$$

is held constant across the ten sets of images. The area under the PSF, defined as the integral over the space domain of $p_1(\mathbf{r})$, is given by $\pi\, r_c^2\, a_c + \pi\, r_t^2\, a_t$ and is set to 1 mm² for normalization purposes. Therefore, given r_c and r_t , a_c and a_t can be determined by solving the system of two equations with two unknowns, given by

$$\pi\,(r_c^2\, a_c + r_t^2\, a_t) = 1 \text{ mm}^2 \tag{3.4}$$

and

$$\frac{a_c}{a_t} = C_{psf} \quad . \tag{3.5}$$

For mathematical convenience, we also model the signal $s(\mathbf{r})$ and the background

b(**r**) as 2D Gaussian distributions, so that each term in equation (3.1) involving a convolution operation takes on a very simple form. The object f(**r**) is then described as

$$f(\mathbf{r}) = \frac{a_s}{\pi r_s^2} \exp(-r^2/r_s^2) + \frac{a_b}{\pi r_b^2} \exp(-r^2/r_b^2) \quad , \tag{3.6}$$

where a_s and a_b are now the integrals under the signal and the background distributions respectively, and r_s and r_b are the half widths at 1/e of the peak intensity of their respective distributions. With such conventions, we can show that the analytical expression of each of the first four terms of equation (3.1) is a 2D Gaussian distribution itself. For example, the convolution of the background b(**r**) with the core of the PSF c(**r**) yields a 2D Gaussian distribution given by

$$b(\mathbf{r}) * c(\mathbf{r}) = a_b\, a_c \frac{r_c^2}{(r_b^2 + r_c^2)} \exp\left(-r^2/(r_b^2 + r_c^2)\right) \quad . \tag{3.7}$$

Finally, we define the contrast of the signal on the background to be

$$C_s = \frac{a_s/r_s^2}{a_b/r_b^2} \quad , \tag{3.8}$$

and given r_s , r_b , and C_s , we can solve for a_s and a_b using

$$(\kappa T)(a_s + a_b) = N_{tot} \tag{3.9}$$

and

$$\frac{a_s}{a_b} = C_s \frac{r_s^2}{r_b^2} \quad , \tag{3.10}$$

where N_{tot} is the total number of counts detected. Note that we used the normalization condition given by equation (3.4) to derive equation (3.9) and that κ must be in mm^{-2} for the units on each side of equation (3.9) to be consistent.

To determine a set of parameters that characterizes the object and the PSFs, we assume that the images are sampled on a 128x128-pixel grid. This grid size yields a maximum value for the modulus of the position vector, denoted as r_{max}, of 90.5 pixels. This modulus constitutes the radius of reference that is used to adjust the widths of the four Gaussian distributions. The width of the background is taken to be the widest, with $r_{max}/r_b = 1$, while the signal is chosen to be 8 pixels in diameter at 1/e of the peak intensity. The core of the PSF is quite sharp, with $r_c = 1.5$ pixels, and the values of r_t, which are varying from one set of images to the next, starting at set number 1 and ending at set number 10, are such that $r_{max}^2/r_t^2 = N + 2$ where N is the set number. The progression of the extent of the tail of the PSFs from set number 1 (PSF_1) to 10 (PSF_{10}) is represented in Figure 3.2. The radius of the outer circle corresponds to the value of r_t for set number 1 while the radius of the inner circle corresponds to the value of r_t for set number 10. The intermediary circles correspond to set number 3, 5 and 8 from the outer to the inner circle, respectively. Given the r_t values and a value for C_{psf} , we can solve equations (3.4) and (3.5) for a_t and a_c. The parameters describing the ten PSFs are given in Table 3.1 for $C_{psf} = 11$, where the values of r_t are given in pixels and a_c and a_t are dimensionless. The ratio $(r_t^2\, a_t)/(r_c^2\, a_c)$ is a measure of the integral under the tail of the PSF relative to the integral under the core, or the relative strength of the tail with respect to the core of the PSF. To illustrate the relative strengths of the tails of the PSFs across the ten experiments, four different PSFs corresponding to image-set numbers 1, 4, 7, and 10 are plotted on Figure 3.3.(a), (b), (c), and (d), respectively. By looking at Table 3.1, we note that

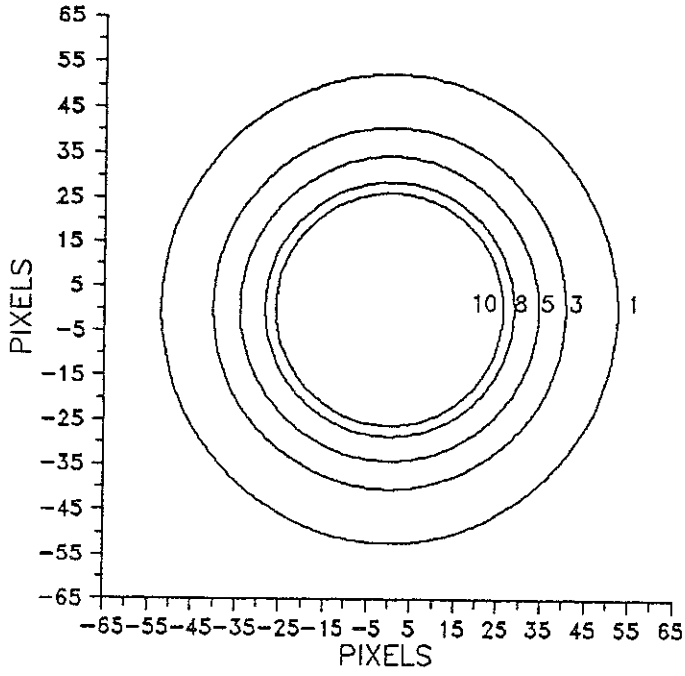


Fig. 3.2. Scheme of the extent of the tail of the PSFs from set number 1 to set number 10. Each circle is centered at the origin of the 128x128-pixel array and has a radius r_t that satisfies $[r_{max}]^2/[r_t]^2 = N + 2$, where N is the set number and $r_{max} = 90.5$ pixels.

POINT SPREAD FUNCTION PARAMETERS				
Set Number	r_t	a_t	a_c	α
1	52.3	1.1E-03	1.3E-02	110
2	45.3	1.5E-03	1.7E-02	83
3	40.5	1.9E-03	2.1E-02	66
4	37.0	2.3E-03	2.5E-02	55
5	34.2	2.7E-03	2.9E-02	47
6	32.0	3.0E-03	3.3E-02	41
7	30.2	3.4E-03	3.7E-02	37
8	28.6	3.8E-03	4.2E-02	33
9	27.3	4.1E-03	4.6E-02	30
10	26.1	4.5E-03	5.0E-02	28

Table 3.1. Point-spread-function parameters: the extent, r_t, of the tail of the PSF and the amplitudes of the two Gaussians representing the tail and the core of the PSF, a_t and a_c, respectively. The parameter α is a measure of the ratio $(a_t\ r_t^2)/(a_c\ r_c^2)$.

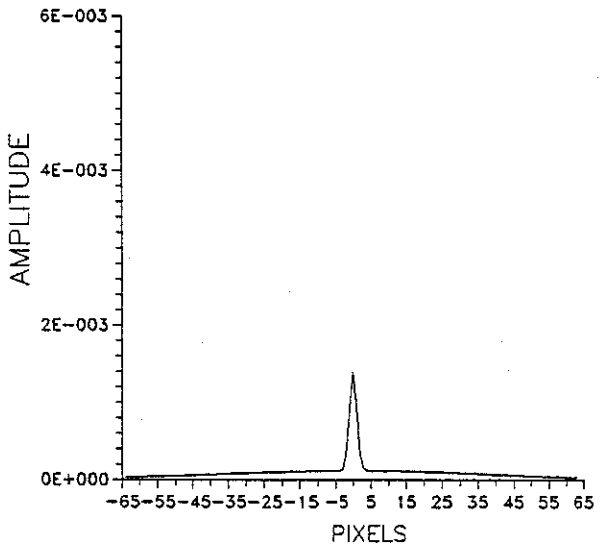


Fig. 3.3. (a) Profile of $p_1(r)$ for set number 1 with r_t = 52.3 pixel.

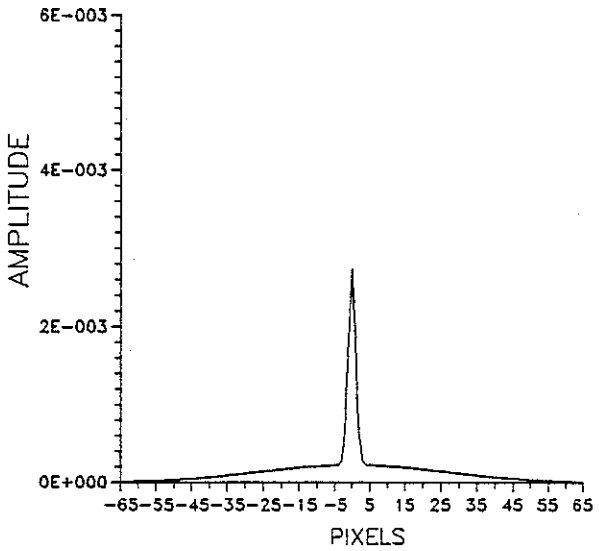


Fig. 3.3. (b) Profile of $p_1(r)$ for set number 4 with r_t = 37.0 pixel.

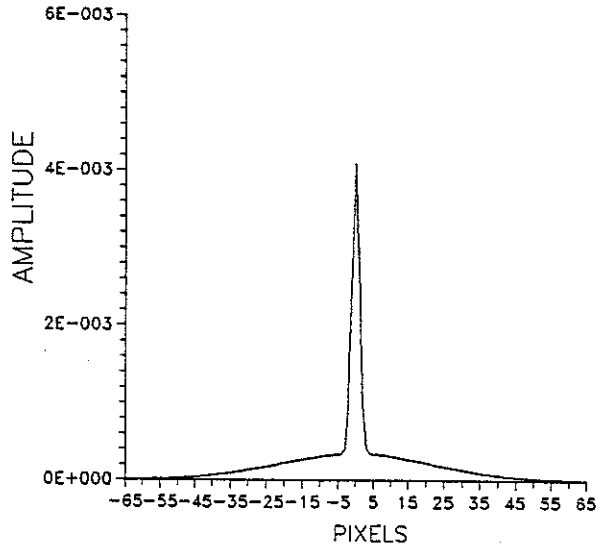


Fig. 3.3. (c) Profile of $p_1(r)$ for set number 7 with r_t = 30.2 pixel.

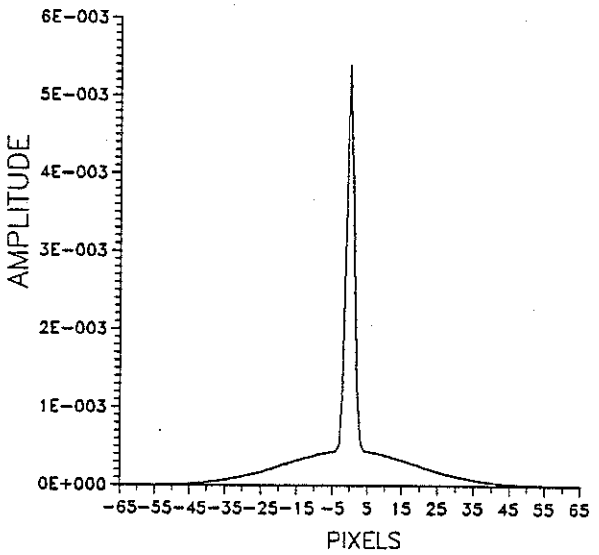


Fig. 3.3. (d) Profile of $p_1(r)$ for set number 10 with r_t = 26.1 pixel.

the ratios $(r_t^2\, a_t)/(r_c^2\, a_c)$ are greater than 1, which is due to the small value of 11 given to C_{psf} . Though it is, in general, quite unlikely that we shall encounter ratios greater than 10 in medical practice (Wagner 1977), the study of these extreme cases will lead to a deeper understanding of several statistical observers, and more especially the ideal observer, as we shall see. The object parameters are adjusted by setting $C_s = 3$ and by adjusting the exposure time such that the noise variance at the peak value of the mean image, denoted as σ_{peak}, is 7% of the peak value, m_{peak}. Moreover, the noise variance is related to the mean pixel value through Poisson statistics, at each point in the image. Thus the noise variance at the peak value of the mean image has to satisfy two conditions

$$\sigma_{peak} = 0.07\ m_{peak} \tag{3.11}$$

and

$$\sigma_{peak} = [\, m_{peak} \,]^{1/2} \quad , \tag{3.12}$$

which yields a value of 204.82 counts/pixel for m_{peak}. For the peak value of the mean image to be 204.82 counts/pixel as we let the PSF vary, the exposure time, or equivalently the total number of counts N_{tot}, has to be adjusted for each set of images; the resulting values for N_{tot} are given in Table 3.2 along with the corresponding values T. Note that, while the total number of counts varies from one set of images to the next as shown in Table 3.2, the contrast of the signal before imaging is held constant across the ten sets of images. A plot of the object profile is given in Figure 3.4, before and after imaging using a long-tailed PSF. The contrast of the object after imaging, denoted as C_{is}, ranges from 5.4% corresponding to set number 1 to 17% corresponding to set number 10 as shown in Table 3.2.

We shall now derive the test statistics for several statistical observers and compute

EXPOSURE TIME VALUES			
Set Number	T	N_{tot}	C_{is} %
1	666	6.7E06	5.4
2	617	6.2E06	6.8
3	584	5.8E06	8.1
4	562	5.6E06	9.4
5	544	5.4E06	10.7
6	529	5.3E06	12.0
7	517	5.2E06	13.3
8	506	5.1E06	14.6
9	496	5.0E06	15.8
10	487	4.9E06	17.1

Table 3.2. Values of the exposure time T across the ten sets of images as well as the total number of counts, N_{tot}. The contrast of the signal after imaging through the ten PSFs is given by C_{is}.

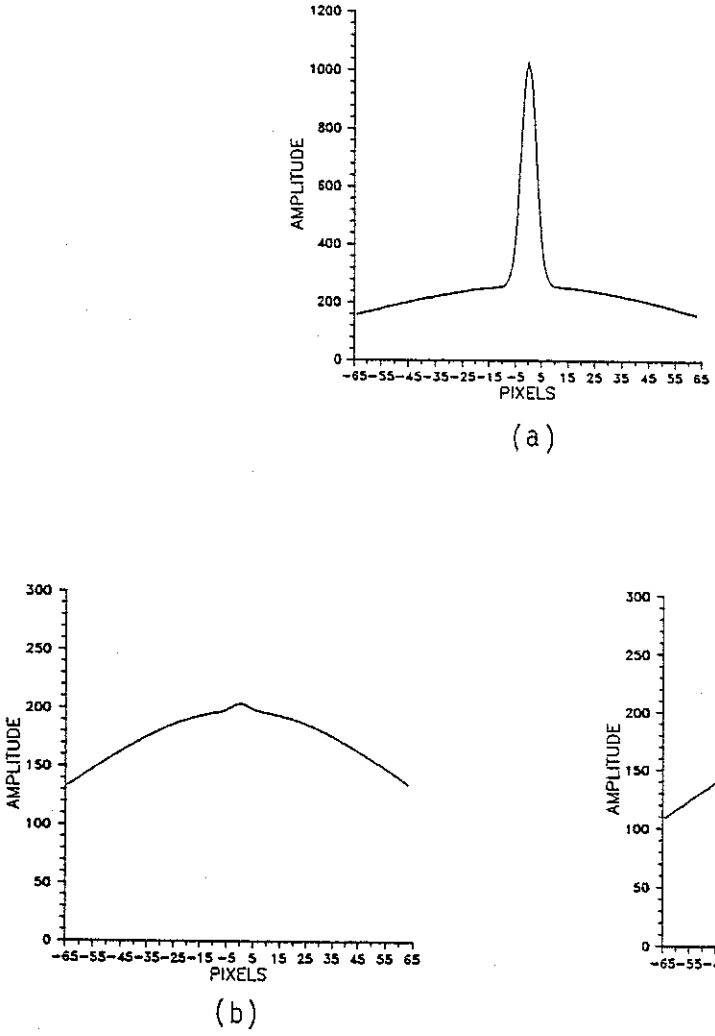


Fig. 3.4. Plots of the profile of the object : (a) before imaging (b) after imaging using PSF_1 and (c) after imaging using PSF_{10}.

their performances by using the values of the object and PSF parameters given in Tables 3.1 and 3.2.

Performance of Standard Observers Before Deconvolution

Given the specific task of detecting the low-contrast signal s(**r**) (when present) on the background b(**r**), one can determine the optimum detectability by making best possible use of the information in the image. Bayesian theory tells us that this is accomplished by applying the strategy of the ideal observer, which was described in detail in chapter 2.

The Ideal Observer

If we call H_1 and H_2 the hypotheses that the signal is absent or present, respectively, we can define the mean images $\overline{g_1}(\mathbf{r})$ and $\overline{g_2}(\mathbf{r})$ under H_1 and H_2, respectively, as

$$\overline{g_1}(\mathbf{r}) = (\kappa\, T)\, [\, b(\mathbf{r}) * (c + t)(\mathbf{r})\,] \tag{3.13}$$

$$\overline{g_2}(\mathbf{r}) = (\kappa\, T)\, [\, (b + s)(\mathbf{r}) * (c + t)(\mathbf{r})\,] \quad , \tag{3.14}$$

where the operator symbol $*$ represents a 2D convolution, and c(**r**) and t(**r**) the core and the tail of the PSF, respectively. The only random variable in this problem is the noise of the system, the signal and the background being known exactly (SKE/BKE). Working with digitized images, we substitute the actual image g(**r**) by a set of numbers g(i) that defines a vector **g**.

The noise is strictly uncorrelated Poisson noise, so the probability densities $p(\mathbf{g}|H_k)$ are multivariate Poissons given by equation (2.38). We showed in chapter 2 that the substitution of equation (2.38) into the expression for the log-likelihood ratio given by

equation (2.27) led to equation (2.40), given by

$$\ell = \sum_{i=1}^{N} \frac{\Delta\bar{g}(i)}{\bar{g}_1(i)}\, g(i) \quad , \tag{3.15}$$

where $\Delta\bar{g}(i)$ is the difference of the mean images under the two hypotheses, $(\bar{g}_2(i) - \bar{g}_1(i))$. If we now compare equation (3.15) with equation (2.9), the ratio $\Delta\bar{g}(i)/\bar{g}_1(i)$ is the i^{th} component of the feature operator **a**. To increase our insight into the ideal-observer construct, we show in Figure 3.5 the profile of the feature operator in the space domain and in the Fourier domain. By comparing Figure 3.4 (a) and 3.5 (a), we note that the ideal observer tries to match the whole region in the image where the signal is located. If we now look at the feature operator in the Fourier domain, we see that the ideal observer behaves mainly as a low-pass filter for the long-tailed PSF problem. In other words, the ideal observer looks where the maximum energy is.

The detectability index d_{ideal}, which is simply the signal-to-noise ratio associated with the log-likelihood ratio, was shown to be given by equation (2.43) as

$$[d_{ideal}]^2 = \sum_{i=1}^{N} \frac{[\Delta\bar{g}(i)]^2}{\bar{g}_1(i)} \quad . \tag{3.16}$$

This expression for d_{ideal} can be explicitly written as a function of $s(\mathbf{r})$, $b(\mathbf{r})$, $c(\mathbf{r})$, and $t(\mathbf{r})$ by using equations (3.13) and (3.14) to yield

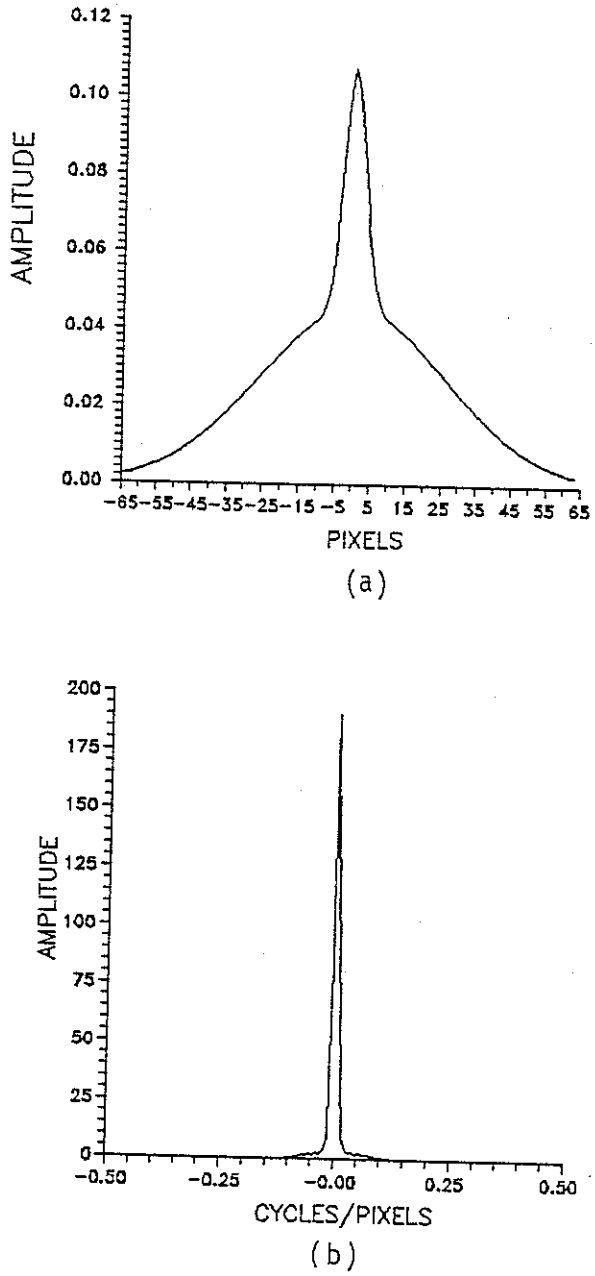


Fig. 3.5. The ideal-observer feature operator. (a) in the space domain (b) in the Fourier domain.

$$[d_{ideal}]^2 = (\kappa\, T) \sum_{i=1}^{N} \frac{[\,(s*c)(i) + (s*t)(i)\,]^2}{[\,(b*c)(i) + (b*t)(i)\,]} \quad . \tag{3.17}$$

The performance of the ideal observer can now be evaluated for each set of images, and the results are presented in Table 3.3.

The Nonprewhitening Ideal Observer

The npw ideal observer performs a simple matched-template operation, and the test statistic is given by

$$\lambda_{npw} = (\Delta\bar{\mathbf{g}})^t \; \mathbf{g} \quad , \tag{3.18}$$

where $\Delta\bar{\mathbf{g}}$ is the difference in the expected means under hypothesis H_1 and H_2. The general expression for the index of detectability index d_{npw} is given by equation (2.11) as

$$[d_{npw}]^2 = \frac{(\Delta\bar{\lambda})^2}{(\sigma_\lambda)^2} \quad . \tag{3.19}$$

By substitution of equation (3.18) into (3.19), equation (3.19) becomes

$$[d_{npw}]^2 = \frac{[\,(\Delta\bar{\mathbf{g}})^t \, \Delta\bar{\mathbf{g}}\,]^2}{(\Delta\bar{\mathbf{g}})^t \; \mathbf{K}_1 \, \Delta\bar{\mathbf{g}}} \quad , \tag{3.20}$$

where $\mathbf{K}$ is the covariance matrix given by equation (2.41). The expression for d_{npw} expressed in terms of $s(\mathbf{r})$, $b(\mathbf{r})$, $c(\mathbf{r})$, and $t(\mathbf{r})$, is then given by

THEORETICAL DETECTABILITIES BEFORE DECONVOLUTION				
Set Number	Ideal	Npw	Nlf	Nlp
1	22.38	22.37	2.32	2.34
2	24.15	24.18	2.86	2.90
3	25.70	25.79	3.39	3.45
4	27.14	27.28	3.91	3.99
5	28.48	28.69	4.42	4.52
6	29.75	30.01	4.92	5.04
7	30.96	31.28	5.40	5.55
8	32.10	32.49	5.88	6.05
9	33.20	33.64	6.35	6.54
10	34.24	34.75	6.80	7.03

Table 3.3. Theoretical values of detectability computed before deconvolution for 4 statistical observers and 10 sets of images.

$$[d_{npw}]^2 = \frac{(\kappa\, T)\left[\sum_{i=1}^{N} [\,(s*c)(i) + (s*t)(i)\,]^2\right]^2}{\sum_{i=1}^{N} [(s*c)(i) + (s*t)(i)]^2\,[(b*c)(i) + (b*t)(i)]} \quad . \tag{3.21}$$

The values of d_{npw} across the ten experiments are given in Table 3.3. We note, as expected, that the performance of the npw ideal and the ideal observers are about the same, although the expressions for d_{ideal} and d_{npw} given by equations (3.17 and 3.21), respectively, are somewhat different. The similarity in performance is a result of the fact that the noise, although not white, is uncorrelated.

The Hotelling Observer

We shall show that the Hotelling observer is equivalent to the ideal observer for a simple two-hypotheses SKE/BKE detection task. The performance of the Hotelling observer may be evaluated using the HTC given by

$$J = \mathrm{tr}\,(\,\mathbf{S}_2^{-1}\,\mathbf{S}_1\,) \quad , \tag{3.22}$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ are the two scatter matrices given by equations (2.48) and (2.51), respectively. We also saw that J may be expressed as a function of the test statistic λ_{Hot} by

$$J = \frac{P_1\,P_2\,[\,\Delta\overline{\lambda}_{Hot}\,]^2}{[\,\sigma_{\lambda_{Hot}}\,]^2} \quad , \tag{3.23}$$

with

$$\lambda_{Hot} = [\,\langle \mathbf{g} \rangle_2 - \langle \mathbf{g} \rangle_1\,]^t\ \mathbf{S}_2^{-1}\ \mathbf{g} \quad . \tag{3.24}$$

Equations (3.23) and (3.24) lead to an expression for J given by

$$J = P_1 P_2 \, (\Delta\bar{g})^t \, \mathbf{S}_2^{-1} \, \Delta\bar{g} \quad . \tag{3.25}$$

If we now use the result from appendix A, where we show that $\mathbf{S}_2$ can be generally expressed as the sum of two terms, one being the contribution of the quantum noise to the covariance matrix while the second term shows the contribution of the background noise to the covariance matrix, we find that, if the background is known exactly, the matrix $\mathbf{S}_2$ reduces to the noise covariance matrix $\mathbf{K}$. We have therefore demonstrated that J and d_{ideal} are related, for the background-known-exactly case, as

$$J = P_1 P_2 \, [d_{ideal}]^2 \quad . \tag{3.26}$$

Under the assumption that the probabilities of occurrences of the two classes are equal with $P_1 = P_2 = 1/2$, then J becomes

$$J = \frac{1}{4} \, [d_{ideal}]^2 = \frac{1}{4} \, (\kappa \, T) \sum_{i=1}^{N} \frac{[\,(s * c)(i) + (s * t)(i)\,]^2}{[\,(b * c)(i) + (b * t)(i)\,]} \quad . \tag{3.27}$$

No-Low-Frequency Ideal Observer

Let us now explore the predictions of some new statistical observers that use specific information about long-tailed PSFs. Referring back to the brightness-contrast effect introduced in chapter 2, we know that the human observer cannot use an absolute brightness level as a useful piece of information when performing a detection task; instead the human observer is essentially sensitive to contrasts in brightness. This property was also

part of the channelized ideal-observer construct that suppressed the lowest-frequency channel in its model. For the long-tailed PSF problem, the term $(s * t)(\mathbf{r})$ itself, appearing in the numerator of equation (3.17), is a broad, slowly varying term that contains mainly low spatial-frequencies. The idea, then, is to test the hypothesis that the human observer does not use this low-frequency information when performing the detection task. To test this hypothesis, we shall collect human-observer data and compare human performance with a pseudo ideal-observer performance that is derived from a test statistic λ_{nlf} defined as

$$\lambda_{nlf} = (\Delta\bar{\mathbf{g}})^t \ \mathbf{K}^{-1} \ \mathbf{g} \quad , \tag{3.28}$$

where the subscript nlf refers to no low-frequency and $\Delta\bar{g}(\mathbf{r}) = \kappa\, T\, (s * c)\, (\mathbf{r})$ instead of $\kappa\, T\, [(s*c) + (s*t)](\mathbf{r})$ used by the ideal observer. The mean values of the data under the two hypotheses H_1 and H_2 are then given by

$$\bar{g}_1(\mathbf{r}) = (\kappa\, T)\ b(\mathbf{r}) * (c(\mathbf{r}) + t(\mathbf{r})) \tag{3.29}$$

$$\bar{g}_2(\mathbf{r}) = (\kappa\, T\,)\ b(\mathbf{r}) * (c(\mathbf{r}) + t(\mathbf{r})) \ + s(\mathbf{r}) * c(\mathbf{r}) \quad , \tag{3.30}$$

and the detectability index associated with the test statistic is expressed as

$$\begin{aligned} [d_{nlf}]^2 &= (\Delta\bar{\mathbf{g}})^t \ \mathbf{K}^{-1} \ \Delta\bar{\mathbf{g}} \\ &= \sum_{i=1}^{N} \frac{[\,\Delta\bar{g}(i)\,]^2}{\bar{g}_1(i)} \\ &= \sum_{i=1}^{N} \frac{(\kappa\, T)\,[\,(s*c)(i)\,]^2}{[\,(b*c)(i) + (b*t)(i)\,]} \quad . \end{aligned} \tag{3.31}$$

The denominator in equation (3.31) is the same as in equation (3.17) since uncorrelated noise is rich in high spatial frequencies, and the deletion of low frequencies does not significantly affect the noise variance. The exact function of the nlf ideal observer is to suppress only the term $(s_* t)(\mathbf{r})$, but since the support of $(s_* c)(\mathbf{r})$ is much smaller than $(s_* t)(\mathbf{r})$, we can also approximate the feature operator of the nlf ideal-observer by the feature operator of the ideal observer multiplied in the Fourier domain by a high-pass filter that takes the value 0 over the support of $(\tilde{s}.\tilde{t})(\boldsymbol{\rho})$ and 1 elsewhere, as shown in Figure 3.6, where $\boldsymbol{\rho}$ is the 2D spatial frequency vector conjugate to r.

The Nlp Ideal Observer

The same strategy used with the ideal observer to derive the nlf ideal observer can be applied to the npw ideal observer. A new observer, referred to as the nlp ideal observer, where nlp stands for no-low-frequency and non-prewhitening, can be derived. The expression for the test statistic λ_{nlp} is given by

$$\lambda_{nlp} = (\Delta\bar{\mathbf{g}})^t \; \mathbf{g} \quad , \tag{3.32}$$

where $\Delta\bar{g}(\mathbf{r})$ can be computed from equations (3.29) and (3.30) and is given by $\kappa\, T\, (s_* c)(\mathbf{r})$. The detectability index associated with the test statistic is thus given by

$$[d_{nlp}]^2 = \frac{(\kappa\, T)\left[\sum_{i=1}^{N} [\,(s_* c)(i)\,]^2\right]^2}{\sum_{i=1}^{N} [(s_* c)(i)]^2\,[(b_* c)(i) + (b_* t)(i)]} \quad . \tag{3.33}$$

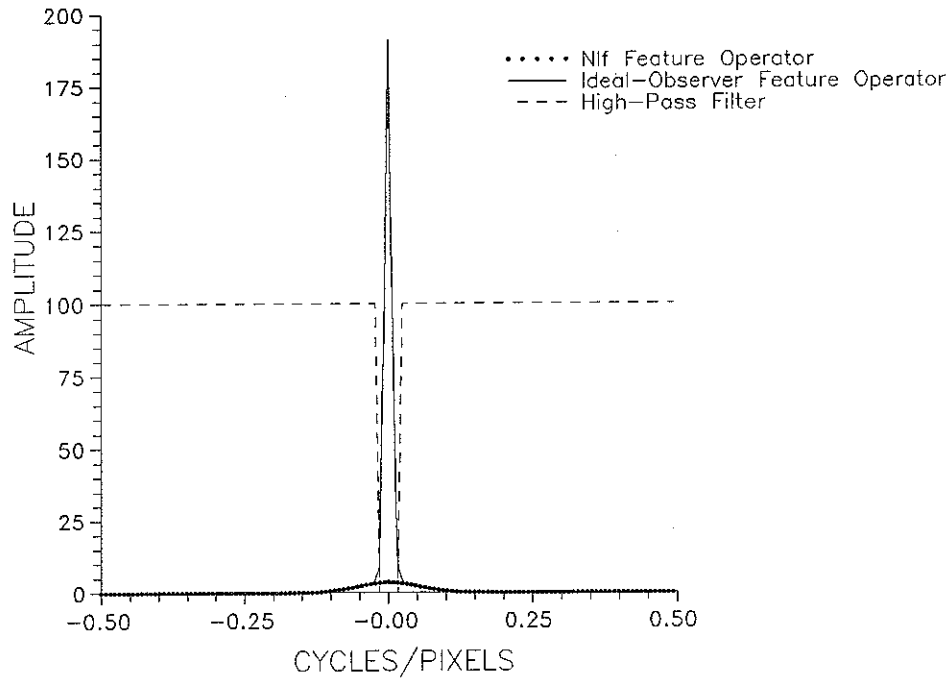


Fig. 3.6. Profile, in the Fourier domain, of the nlf ideal-observer feature operator superimposed on the ideal-observer feature operator. The high-pass filter represented by a dashed line shows how the nlf ideal-observer feature operator can be approximated by the ideal-observer feature operator multiplied by a high-pass filter that takes the value 0 over the support of $\tilde{s}(\rho)\,\tilde{t}(\rho)$ and the value of 1 otherwise. Note that the high-pass filter has been scaled on the figure by a factor of 100 for clarity.

The predicted values of detectabilities of the nlp ideal observer are also given in Table 3.3.

Performance of Standard Observers after Linear Deconvolution

One concrete prediction of the ideal-observer construct is that linear filtering of an image cannot improve performance; if filtering is desirable, the ideal observer can do it internally as part of the decision strategy. Thus the question arises as to whether or not the human observer performance will improve with linear filtering. High-pass filtering is commonly applied to suppress low-frequency components in an image, generally enhancing at the same time high frequencies. Among others, Jaszczak et al. (1984), Clough (1986), Msaki et al. (1987), Seibert and Boone (1988), Yanch et al. (1988), Heuscher and Brunnett (1989), and Fahimi and Macovski (1989) have expended considerable effort in filtering radiographic images to remove scatter. They showed that the removal of scattered radiation led to the enhancement of the contrast in the image. For the types of images that we are looking at, which have a highly grainy structure due to the small number of counts, high-pass filtering will not only remove some low-frequency components in the image but also boost the noise, possibly nullifying the advantage of deconvolution (Kennett and Prestwich 1979). Furthermore, the filtering will introduce noise correlations, which are known to interfere with the performance of human observers as shown by Myers et al. (1985). Although it is unclear if any improvements can be achieved, the intuitive feeling persists that sharper images should somehow be better images. In particular, the author's intuition suggests that filtering to remove scattering from radiographs should improve their quality, at least for the human observer.

A schematic of the imaging process is given in Figure 3.1. (b), where the images are convolved by a filter function before display. We want to filter the images to correct for the degradation resulting when the PSF departs from a sharply peaked function. One

can invert the convolution process by multiplying the Fourier transform of the image by the reciprocal of the Fourier transform of the PSF, provided this Fourier transform is nonzero over the range of frequencies of interest (Bracewell 1986; Goodman 1988). Such an operation is called deconvolution. If we denote by $\tilde{f}_h(\rho)$ the filter used to deconvolve the images, where the subscript h specifies the high-pass nature of the filter, we design the appropriate filter for each set of images as depicted in Figure 3.7. The Fourier transform of the long-tailed PSF, $p_1(\mathbf{r})$, represented in the top-left quadrant, is first performed to yield the transfer function $\tilde{p}_1(\rho)$ plotted in the top-right quadrant. We then superimpose on the same graph the desired transfer function $\tilde{p}_2(\rho)$ that would have yielded high-resolution images, if it had been used in the first place instead of $\tilde{p}_1(\rho)$. A plot of the Fourier transform of $\tilde{p}_2(\rho)$ is shown in the bottom right quadrant of Figure 3.7 to show its sharpness. We can then design the appropriate filter in the Fourier domain by dividing $\tilde{p}_2(\rho)$ by $\tilde{p}_1(\rho)$, and the filter expression is then given by

$$\tilde{f}_h(\rho) = \frac{\tilde{p}_2(\rho)}{\tilde{p}_1(\rho)} = \frac{(r_c^2 a_c + r_t^2 a_t) \exp(-\pi^2 |\rho|^2 r_c^2)}{r_c^2 a_c \exp(-\pi^2 |\rho|^2 r_c^2) + r_t^2 a_t \exp(-\pi^2 |\rho|^2 r_t^2)} , \qquad (3.34)$$

where $|\rho|$ is the modulus of the 2D frequency vector ρ. An inverse Fourier tranform is then performed on the filtered images to take us back in the space domain. The nature of the filtering operation is high-pass as shown in Figure 3.8, where an example of the four filters designed to deconvolve the images from set numbers 1, 4, 7, and 10 is given.

The Ideal Observer

Since we stated earlier that linear filtering does not affect the performance of the ideal observer, we propose to show now that it is the case. Given a data set **g** satisfying the

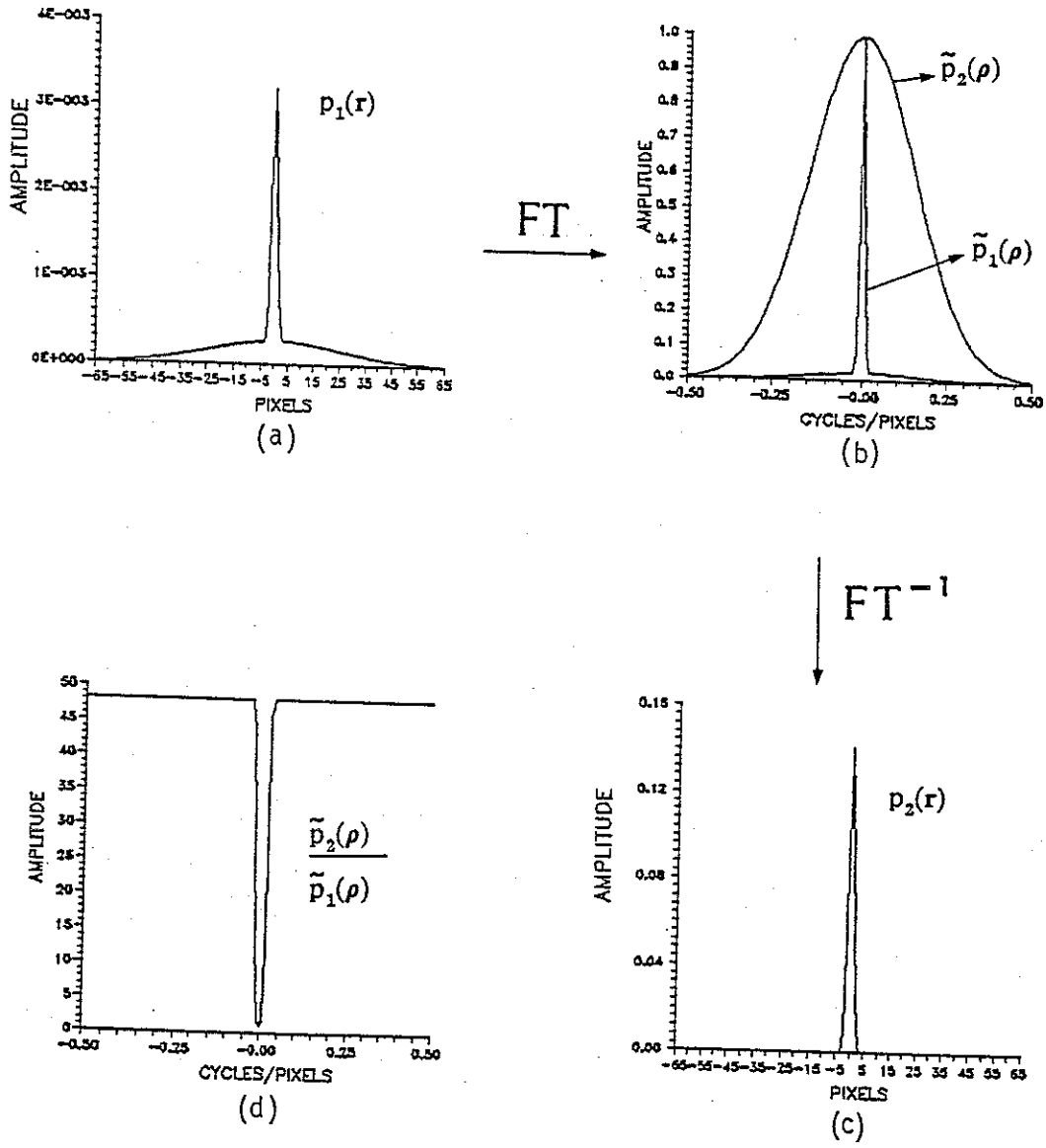


Fig. 3.7. Scheme of the filtering process: Profiles of (a) $p_1(r)$ for set number 5 (b) Fourier transform of $p_1(r)$, $\tilde{p}_1(\rho)$, superimposed on $\tilde{p}_2(\rho)$ (c) Inverse Fourier transform of $\tilde{p}_2(\rho)$, and (d) High-pass filter $F_h(\rho)$ formed by taking the ratio $\tilde{p}_2(\rho)/\tilde{p}_1(\rho)$.

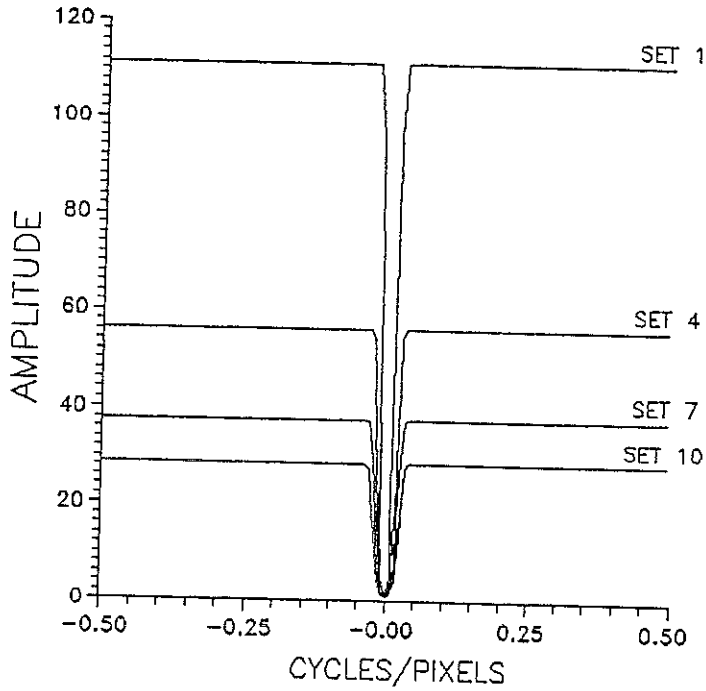


Fig. 3.8. Four examples of the filter used in the Fourier domain to deconvolve the images. Here we show a 1D plot of the shapes of the filters used with the image set number 1, 4, 7, and 10. The filters are rotationally symmetric around the zero frequency, where they take the value of 1 .

imaging equation (2.7), if we denote by $\mathbf{F}_h$ the matrix operator associated with the filtering operation, the filtered data set $\mathbf{g}^f$ is given by

$$\mathbf{g}^f = \mathbf{F}_h \; \mathbf{g} = \mathbf{F}_h \; \mathbf{H} \; \mathbf{f} + \mathbf{F}_h \; \mathbf{n} \quad . \tag{3.35}$$

The covariance matrix of the filtered data $\mathbf{g}^f$ under hypothesis H_k is then given by

$$\begin{aligned} \mathbf{K}^f{}_k &= \langle (\mathbf{g}^f - \overline{\mathbf{g}^f})\,(\mathbf{g}^f - \overline{\mathbf{g}^f})^t \,|H_k \rangle \\ &= \langle (\mathbf{F}_h \; \mathbf{n})\,(\mathbf{F}_h \; \mathbf{n})^t \,|H_k \rangle \\ &= \mathbf{F}_h \; \langle \mathbf{n} \; \mathbf{n}^t |H_k \rangle \; \mathbf{F}_h{}^t \\ &= \mathbf{F}_h \; \mathbf{K}_k \; \mathbf{F}_h{}^t \quad . \end{aligned} \tag{3.36}$$

To avoid cumbersome notation, we shall consider the low-contrast case where $\mathbf{K}_1 = \mathbf{K}_2 = \mathbf{K}$. We then have $\mathbf{K}^f{}_1 = \mathbf{K}^f{}_2 = \mathbf{K}^f$, and the detectability index associated with the filtered data-set, denoted as $d^f{}_{ideal}$, where the superscript f is used to specify "filtered", is given by

$$\begin{aligned} [d^f{}_{ideal}]^2 &= (\Delta\overline{\mathbf{g}^f})^t \;\; \mathbf{K}^f{}_1{}^{-1} \; \Delta\overline{\mathbf{g}} \\ &= (\mathbf{F}_h \; \Delta\overline{\mathbf{g}})^t \; (\mathbf{F}_h \; \mathbf{K}_1 \; \mathbf{F}_h{}^t)^{-1} \; (\mathbf{F}_h \; \Delta\overline{\mathbf{g}}) \\ &= (\Delta\overline{\mathbf{g}})^t \; \mathbf{F}_h{}^t \; (\mathbf{F}_h{}^t)^{-1} \; \mathbf{K}_1{}^{-1} \; (\mathbf{F}_h)^{-1} \; \mathbf{F}_h \;\; \Delta\overline{\mathbf{g}} \\ &= (\Delta\overline{\mathbf{g}})^t \; \mathbf{K}_1{}^{-1} \; \Delta\overline{\mathbf{g}} \\ &= [d_{ideal}]^2 \quad , \end{aligned} \tag{3.37}$$

which demonstrates the invariance of the ideal-observer performance under linear filtering for the low-contrast case. Since we are now familiar with the mathematics involved in calculating the ideal-observer performance, we see that this result can be easily generalized

to any contrast. Instead of calculating the variance contribution to the detectability as a single term, we would need to consider a term with $\mathbf{K}_1$ and a term with $\mathbf{K}_2$.

The Nonprewhitening Ideal Observer

Let us now look at the effect of linear filtering on the performance of the npw ideal observer. The test statistic is now given by

$$\lambda_{npw} = \mathbf{b}^t \ \mathbf{g}^f \tag{3.38}$$

where $\mathbf{g}^f$ is the filtered data set and $\mathbf{b}$ the feature operator for the filtered data. If we define as $f_h(r)$ the inverse Fourier transform of $\tilde{f}_h(\rho)$ given by equation (3.34), the feature operator for the npw ideal observer is simply the expected signal seen through the equivalent imaging device of response function $p_2(r)$, where $p_2(r)$ equal $p_1(r) * f_h(r)$. By using equation (3.35), we can rewrite the test statistic as

$$\begin{aligned} \lambda_{npw} &= \mathbf{b}^t \ \mathbf{F}_h \ \mathbf{g} \\ &= \mathbf{a}^t \ \mathbf{g} \quad , \end{aligned} \tag{3.39}$$

where $\mathbf{a}$ is now the feature operator associated with the unfiltered data. Moreover, with $\Delta\mathbf{f}$ being the difference object under the two hypotheses and $\mathbf{H}_2$ the system operator associated with $p_2(r)$, $\mathbf{b} = \mathbf{H}_2 \ \Delta\mathbf{f}$, and the feature operator $\mathbf{a}$ becomes

$$\mathbf{a}^t = (\mathbf{H}_2 \ \Delta\mathbf{f})^t \ \mathbf{F}_h \quad . \tag{3.40}$$

In the case of the two-hypotheses SKE/BKE detection task, $\Delta\mathbf{f}$ is simply the signal $\mathbf{s}$ to be

detected, and equation (3.39) yields

$$\lambda_{npw} = s^t \; H_2^t \; F_h \; g \quad . \tag{3.41}$$

The detectability index d_{npw} is given by the ratio of $\Delta\overline{\lambda}_{npw}$ to $\sigma_{\lambda_{npw}}$ with

$$\begin{aligned} \Delta\overline{\lambda}_{npw} &= \langle \lambda_{npw} | H_2 \rangle - \langle \lambda_{npw} | H_1 \rangle \\ &= s^t \; H_2^t \; F_h \; (\langle g | H_2 \rangle - \langle g | H_1 \rangle) \\ &= s^t \; H_2^t \; F_h \; \Delta\overline{g} \end{aligned} \tag{3.42}$$

$$\begin{aligned} [\sigma_{\lambda_{npw}}]^2 &= \frac{1}{2} [\langle (\lambda_{npw} - \langle \lambda_{npw} | H_2 \rangle)^2 | H_2 \rangle + \langle (\lambda_{npw} - \langle \lambda_{npw} | H_1 \rangle)^2 | H_1 \rangle] \\ &= \frac{1}{2} [\langle [\, s^t \; H_2^t \; F_h \; (g - \overline{g_2}) \,]^2 \rangle + \langle [\, s^t \; H_2^t \; F_h \; (g - \overline{g_1}) \,]^2 \rangle] \\ &= \frac{1}{2} \sum_{k=1}^{2} [s^t \; H_2^t \; F_h \; \langle (g - \overline{g_k})(g - \overline{g_k})^t \rangle \; (F_h^t \; H_2 \; s)] \\ &= \frac{1}{2} \sum_{k=1}^{2} s^t \; H_2^t \; F_h \; K_k \; F_h^t \; H_2 \; s \; , \end{aligned} \tag{3.43}$$

where K_k is the unfiltered noise covariance matrix for class k. For the case of uncorrelated Poisson noise, and in the low-contrast approximation, equation (3.43) yields

$$[\sigma_{\lambda_{npw}}]^2 = s^t \; H_2^t \; F_h \; K \; F_h^t \; H_2 \; s \quad , \tag{3.44}$$

and the detectability index d_{npw} becomes

$$[d_{npw}]^2 = \frac{[\, s^t \;\; H_2{}^t \;\; F_h \;\; \Delta\bar{g}\,]^2}{s^t \;\; H_2{}^t \;\; F_h \;\; K \;\; F_h{}^t \;\; H_2 \;\; s}$$

$$= \frac{\left[\sum_i \sum_j \sum_k s(i)\, H_2(i,j)\, F_h(j,k)\, \Delta\bar{g}(k)\right]^2}{\sum_i \sum_j \sum_k \sum_l \sum_m \sum_n s(i)\, H_2(i,j)\, F_h(j,k)\, K(k,l)\, F_h(l,m)\, H_2(m,n)\, s(n)}$$

$$= \frac{\left[\sum_i \sum_j \sum_k s(i)\, H_2(i,j)\, F_h(j,k)\, \Delta\bar{g}(k)\right]^2}{\sum_k A(k)\, \bar{g}_1(k)\, A(k)}$$

$$= \frac{\left[\sum_k A(k)\, \Delta\bar{g}(k)\right]^2}{\sum_k [A(k)]^2\, \bar{g}_1(k)} \;, \qquad (3.45)$$

where A(k) is given by

$$A(k) = \sum_i \sum_j s(i)\, H_2(i,j)\, F_h(j,k) \quad . \qquad (3.46)$$

We can write d_{npw} as a continuous integral in the space domain by taking the limit of the sum when the area of the sampling element tends to zero and by undoing the lexicographic ordering to transform the 1D sums into 2D integrals as

$$[d_{npw}]^2 = \frac{\left[\int_{\infty} d^2\mathbf{r}\ A(\mathbf{r})\ \Delta\bar{g}(\mathbf{r})\right]^2}{\int_{\infty} d^2\mathbf{r}\ [A(\mathbf{r})]^2\ \bar{g}_1(\mathbf{r})} = \frac{[\,(\Delta\bar{g}(\mathbf{r}) * A(\mathbf{r}))\,(0)\,]^2}{(\bar{g}_1(\mathbf{r}) * [A(\mathbf{r})]^2)\,(0)}\ , \qquad (3.47)$$

where (0) means that the convolution is evaluated at zero shift. Once d_{npw} is expressed as a ratio of convolution operations, it is straightforward to evaluate both the numerator and the denominator in the Fourier domain, a convolution in space domain being transformed into a simple product in Fourier domain. We shall, as an example, translate the numerator of equation (3.47) into its Fourier expression. It can be written as

$$[\Delta\bar{g}(\mathbf{r}) * A(\mathbf{r})]\,(0) = \sum_{i=1}^{N} \Delta\bar{g}(i)\ A(i) = \sum_{i=1}^{N} \Delta\tilde{\bar{g}}(i)\ \tilde{A}(i)$$

$$= \sum_{i=1}^{N} \kappa\, T\,(\,\tilde{s}(i)\,)\,(\tilde{c}(i) + \tilde{t}(i))\ \frac{\tilde{s}(i)\ |\,\tilde{p}_2(i)\,|^2}{(\tilde{c}(i) + \tilde{t}(i))}$$

$$= \sum_{i=1}^{N} \kappa\, T\,[\,\tilde{s}(i)\,]^2\ |\,\tilde{p}_2(i)\,|^2\ . \qquad (3.48)$$

A similar exercise can be performed on the denominator. The results of the computation of

d_{npw} are given in Table 3.4.

It is quite interesting to look at the shape of the feature operator **a** for the npw ideal observer after deconvolution. The profile of the feature operator given by equation (3.39) is plotted in Figures 3.9 and 3.10. In the space domain, the feature operator first calculates the contribution of the local background surrounding the expected signal-location and substract it from the signal-plus-background contribution. This is very different from the feature operator associated with the ideal observer that was shown to be equivalent to a low-pass filter.

The Hotelling Observer

We have already shown that, before deconvolution, the Hotelling observer is equivalent to the ideal observer, and we have also shown that the ideal observer is invariant under filtering. Therefore, the Hotelling observer is also invariant under linear filtering.

THEORETICAL DETECTABILITIES AFTER DECONVOLUTION			
Set Number	Ideal	Nlf	Npw
1	22.38	2.32	1.67
2	24.15	2.86	2.16
3	25.70	3.39	2.65
4	27.14	3.91	3.13
5	28.48	4.42	3.61
6	29.75	4.92	4.09
7	30.96	5.40	4.57
8	32.10	5.88	5.04
9	33.20	6.35	5.52
10	34.24	6.80	5.99

Table 3.4. Theoretical values of detectability computed after deconvolution across the ten sets of images for 3 statistical observers, the ideal, the nlf ideal, and the npw ideal observers.

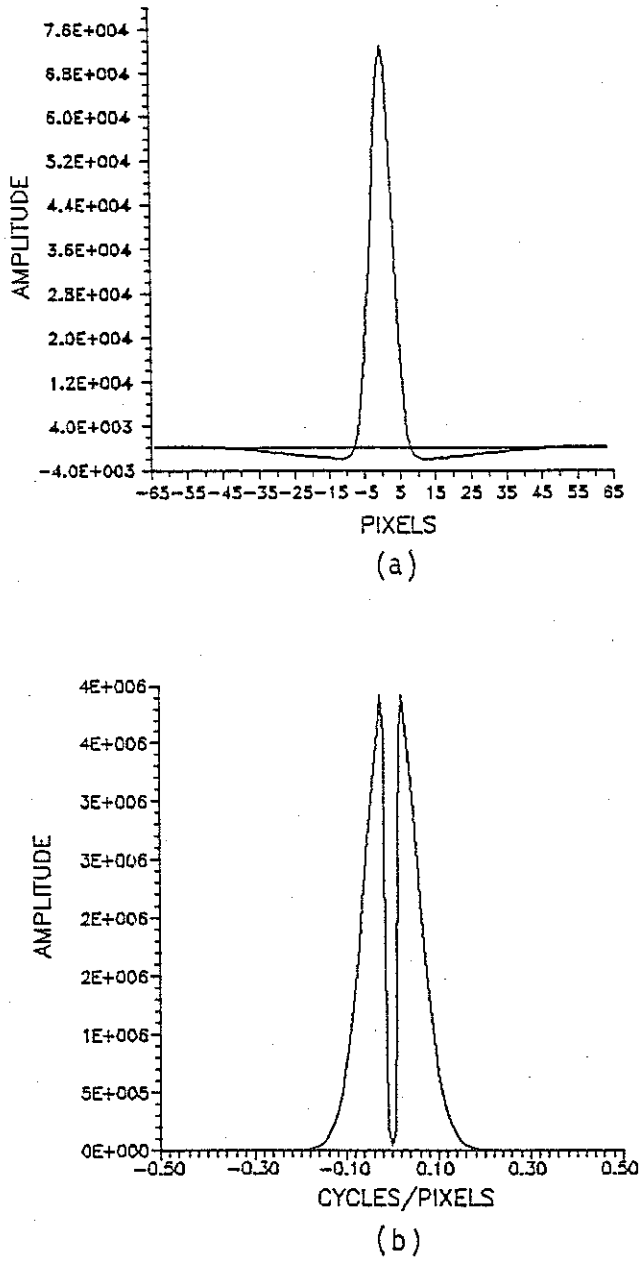


Fig. 3.9. The npw ideal-observer feature operator after deconvolution using $p_1(r)$ defined for set number 1 (a) in the space domain (b) in the Fourier domain.

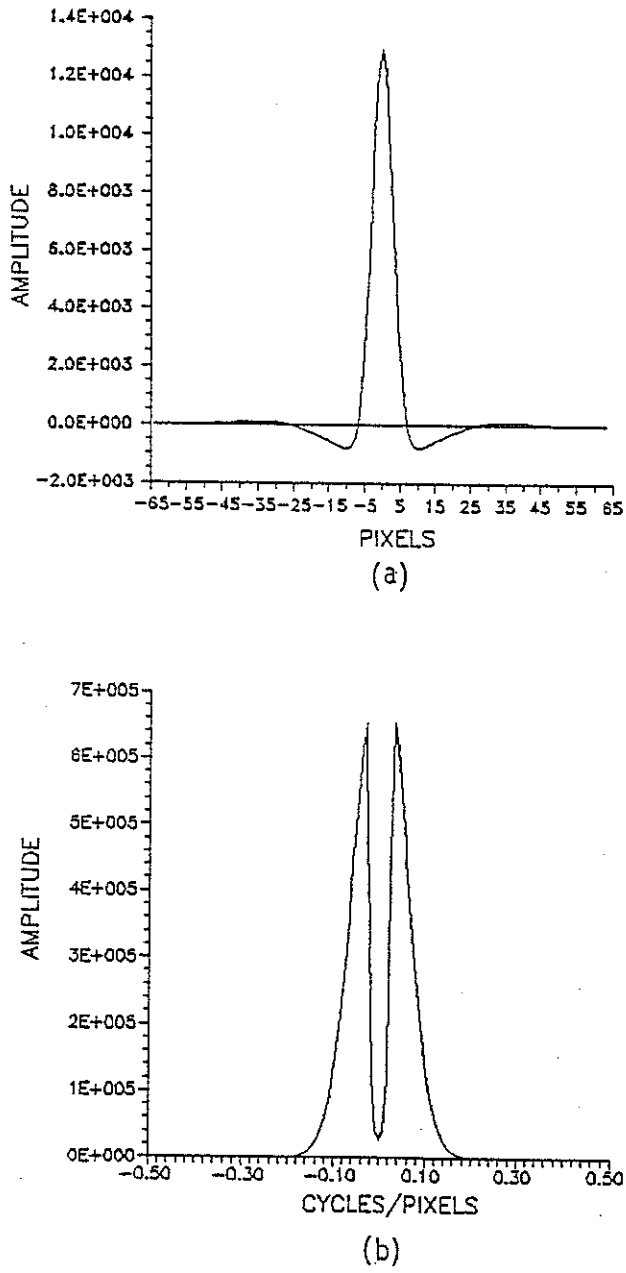


Fig. 3.10. The npw ideal-observer feature operator after deconvolution using $p_1(r)$ defined for set number 10 (a) in the space domain (b) in the Fourier domain.

CHAPTER4

EXPERIMENTAL METHODS AND RESULTS FOR LONG-TAILED PSFS

For the detectability predicted by any statistical observer to be a good figure of merit, it must not only correlate with the human observer performance, but also yield sufficient efficiency. Since efficiency depends strongly on the amount of training that the human observers undertake, it is in some ways difficult to define what sufficient efficiency means. However, if we assume that, prior to the experiments, the human observers were familiarized with the types of images to be viewed and trained to perform the task according to the experimental procedure, we can define as an acceptable range of efficiencies any value between 10% and 50%. Below 10% efficiency, however, we shall conclude that the human observer is somehow either not using all the information available or using it improperly when compared to the statistical observer. In this research, because we are interested in comparing the performance of different statistical observers with the performance of the human observer, we are not so much concerned with working at 50% efficiency; rather we set our expectations around 20% and focus on the relative differences between the different figures of merit when compared to the human performance.

We shall now investigate the effect that long-tailed PSFs have on human visual detection and whether any improvement in performance can be gained by deconvolving the tails of the PSF.

Methods and Results before Deconvolution

Experimental Methods

To investigate the human observer performance, 10 psychophysical experiments using the ten sets of images described in chapter 3 were designed. Within each experiment 50 images were generated; 25 images simulated a bright tumor, located in the center of the picture and superimposed on a smooth background, while the tumor was absent in the other 25 images. The numerical values assigned to the different parameters and given in chapter 3 represent one possible set of parameters that gives, once Poisson noise is introduced, a visually measured detectability, d_{human}, in the range 0.5 to 3.0, where small changes in detectability can be detected using ROC analysis. Contrast-detail diagrams were used to select an appropriate set of parameters to design the experiments, but only as a working tool, and were not used in any ways to evaluate performance.

The computer-simulated images, generated as described in chapter 3, were displayed using the eight bits of available greylevels. They were then displayed as 128x128-pixel arrays using the greylevel-to-brightness transfer curve shown in Figure 4.1. The procedure followed to measure the output brightness vs. the displayed greylevels will be described in detail later, as we study the impact of the shape of the transfer curve on the performance. One image covered 5.5 cm x 4.5 cm on the display, and each pixel measured 0.43 x 0.35 mm^2. The light in the room was dimmed down to the same value for all the observers. The contrast and brightness levels of the display monitor were fixed during the study, and the observers were not allowed to vary them to optimize their performance. Each subject viewed the images binocularly at a comfortable distance chosen by the observer, and the observers were required to wear their usual correction lenses.

Six observers ran the study, each performing all 10 experiments. The observers first viewed sample images to ascertain the shape, size and location of the signal. They did

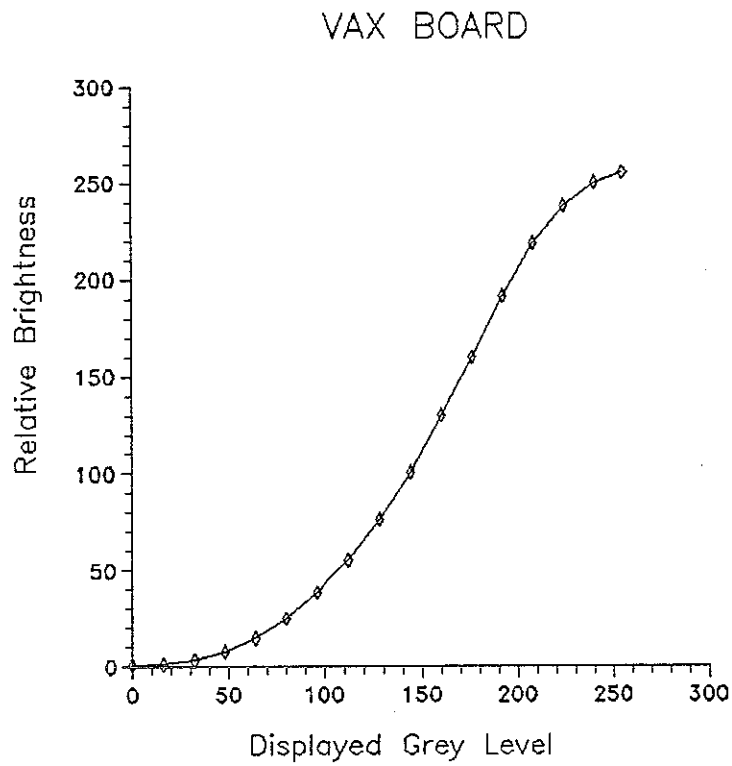


Fig. 4.1. Plot of the measured values of the grey-level-to-brightness transfer curve using an imaging board controlled by the VAX.

not have any information on the number of occurrences of the signal. The images were displayed in a random order, and the 10 experiments were also performed in a randomized order for each observer. The observer was presented a single stimulus on each trial, either the signal on a noisy background or a noisy background alone. An example of the simuli presented to the human observer, with the signal present, is shown in Figure 4.2. The observers were instructed to rate their certainty on a six-point scale (Green and Swets 1966) given in Table 4.1, and their responses were recorded and analysed using a computer program written by Seeley et al. (1982).

From the rating-scale data, an ROC curve was generated for each observer and each experiment (Swets 1979; Swets and Pickett 1982). Five points on a curve were obtained for each observer. All statistics describing the data were derived from non-parametric estimates of the ROC curves generated for each observer and then averaged over all observers (Hanley and Mc Neil 1982; Metz 1989). We used the estimated value of the area under the ROC curve to compute the detectability index d_a using Equation (2.22), where standard Tables of the erf(x) function were used to invert the Equation.

Results

The results of the experiments are given in column 2 of Table 4.2. We note first that the detectability values achieved by the human observer, which we shall denote as d_{human}, range from 0.4 to 3 as we expected since the parameters used to design the experiments were chosen such that the performance would fall within this range. Such a range of values is favorable since 0.4 corresponds to an area under the ROC curve of 0.610 while 3 corresponds to an area under the ROC curve of 0.983, and even small variations in performance can usually be observed when working within this range. To picture the performance of the human observer relative to any statistical observer, we plot the

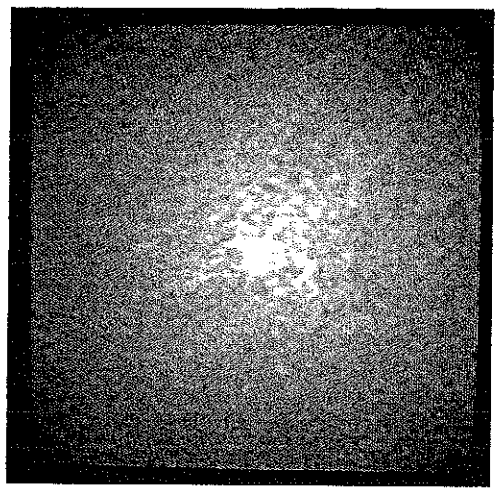


Fig. 4.2. (a) An example of the stimuli presented to the human observer during the psychophysical studies: here, we show a simulus where the signal, located in the center of the image, is present.

0	Relatively certain that no signal is present
1	Relatively certain that no signal is present
2	Just guessing that no signal is present
3	Just guessing that signal is present
4	Relatively certain that signal is present
5	Absolutely certain that signal is present

Table 4.1. Rating scale

EXPERIMENTAL RESULTS BEFORE DECONVOLUTION			
Experiment Number	d_{human}	η(ideal)%	η(nlf)%
1	0.41	0.03	3.12
2	0.72	0.09	6.34
3	0.74	0.08	4.78
4	1.16	0.18	8.80
5	1.41	0.25	10.18
6	1.46	0.24	8.81
7	1.93	0.39	12.77
8	2.03	0.40	11.92
9	2.79	0.71	19.30
10	3.04	0.79	19.99
AVERAGE EFFICIENCY		0.3%	10.8%
SLOPE		0.21	0.57
INTERCEPT		-4.50	-1.06
CORRELATION COEFFICIENT		0.96	0.98

Table 4.2. Results of the psychophysical study carried out to investigate the impact of long-tailed PSFs on human performance using 10 experiments numbered in column 1, where the tail of the PSF is decreasing from experiment 1 to 10. For each experiment, the performance is given in column 2 in terms of the detectability value associated with the human observer, and in column 3 to 5 in terms of the efficiency of the human observer relative to 3 statistical observers and calculated using the theoretical values of detectability given in table 3.3. The slopes and intercepts of the regression lines that fit the pair of points given by the experimental data points and the theoretical values are also given, as well as their correlation coefficients.

detectability of the human observer versus the predicted theoretical values. The performance of the human observer versus the ideal observer and the nlf ideal observer is plotted Figure 4.3. A linear regression is performed to fit the experimental data values plotted versus the theoretical values. Results show a correlation of the human observer relative to the ideal observer of 0.96 and relative to the nlf ideal observer of 0.98 (see Table 4.2). Despite of the high correlation of the data with both figures of merit, Figure 4.3 shows a strong difference in performance between the detectability values that the human observer and the ideal observer are operating on. We shall quantify this discrepancy by computing the efficiency of the human observer relative to the statistical observers.

By analogy with Barlow's definition of statistical efficiency (Barlow 1978), which is defined as the squared ratio of d_{human} to d_{ideal} and which we shall denote η(ideal), we define a pseudo-efficiency η(nlf) as the squared ratio of d_{human} to d_{nlf}. The values of η(ideal) and η(nlf), computed by using the experimental detectability values given in column 2 of Table 4.2 and the theoretical detectability values given in columns 2 and 3 of Table 3.3, are given in Table 4.2, columns 3 and 4. From these values, the average efficiency over the ten experiments is calculated and is given in Table 4.2. Results show an average efficiency η(ideal) of only 0.3% for the human observer with respect to the ideal observer. This low efficiency value shows that the ideal observer greatly overestimates the performance of the human observer. If we now look at the performance of the human observer relative to the nlf ideal observer, results show an efficiency of the human observer relative to the nlf ideal observer of 10.8% which is a factor 36 greater than relative to the ideal observer. These efficiency values are the quantification with a single number of the fact that the values of d_{human} are in better agreement with the values of d_{nlf} than with the values of d_{ideal}.

Another way of looking at the performance is to compare the performance of the ideal, the nlf ideal, and the human observers with respect to the contrast of the signal after

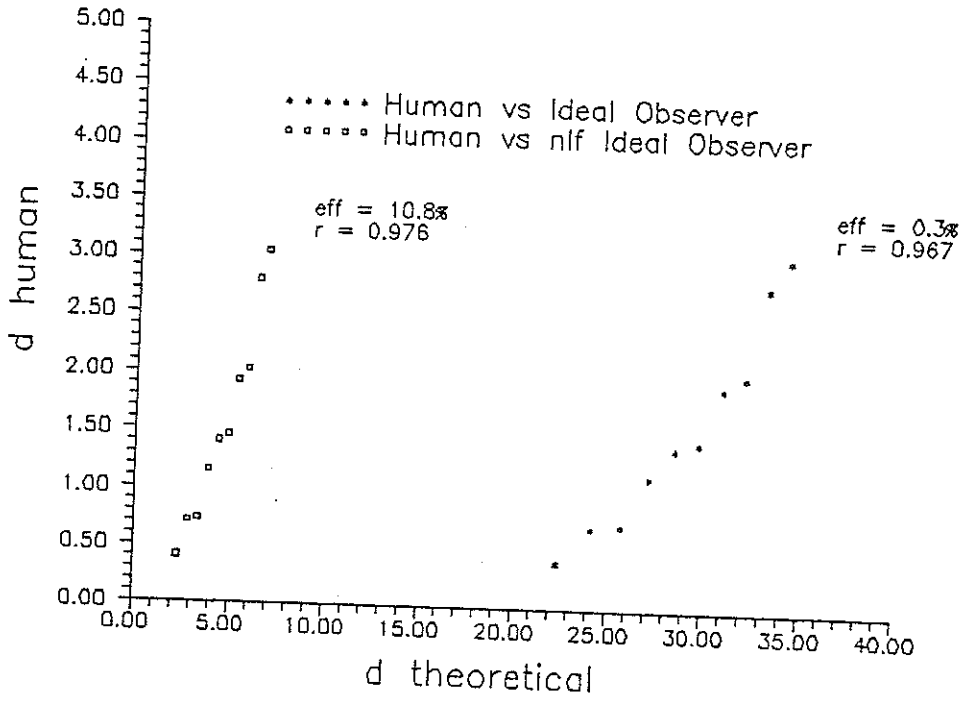


Fig. 4.3. Plot of the detectabilities achieved by the human observer across the ten experiments, d_{human} , where d_{human} increases from experiment 1 to 10 as the extent of the tail of the PSF decreases, versus the performances of the ideal observer and the nlf ideal observer. The respective values of the correlations (r) and the average (across the ten experiments) efficiencies (eff) of the human observer versus the two statistical observers are also given.

EXPERIMENTAL RESULTS
AFTER DECONVOLUTION

Experiment Number	d_{human}	η(ideal)%	η(nlf)%	η(npw)%
1	0.65	0.08	7.85	15.15
2	1.05	0.19	13.48	23.85
3	1.23	0.23	13.16	21.87
4	1.70	0.39	18.90	30.07
5	2.00	0.49	20.48	31.39
6	2.15	0.52	19.10	28.32
7	2.48	0.64	21.09	30.37
8	2.82	0.77	23.00	32.33
9	3.59	1.17	31.96	43.87
10	3.43	1.00	25.44	34.03
AVERAGE EFFICIENCY		0.6 %	19.5 %	29.1 %
SLOPE		0.24	0.65	0.69
INTERCEPT		-4.82	-0.88	-0.5
CORRELATION COEFFICIENT		0.98	0.99	0.99

Table 4.3. Results of the psychophysical study carried out to investigate the impact of deconvolution on human performance using 10 experiments. For each experiment the performance is given in column 2 in terms of the detectability values associated with the human observer, and in column 3 to 5 in terms of the efficiency of the human observer relative to 3 statistical observers and calculated using the theoretical values of detectability given in table 3.4. The slopes and intercepts of the regression lines that fit the pair of points given by the experimental data points and the theoretical values are also given, as well as their correlation coefficients.

imaging. The contrast of the signal after imaging, denoted as C_{is}, is calculated as follows

$$C_{is} = \frac{(s*c)(0) + (s*t)(0)}{(b*c)(0) + (b*t)(0)} \tag{4.1}$$

and the values of C_{is} are given in Table 3.2 across the ten sets of images. A plot of the detectabilities predicted by the ideal and the nlf ideal observers is given in Figure 4.4 where the values corresponding to the ten experiments (C_{is} between 5.4% and 17%) are plotted as well as some lower values of the contrast to better demonstrate the behavior of the ideal observer performance. To reach lower contrast values, the extent of the tail of the PSF is simply increased. This plot first shows that the ideal observer performance is highly non-linear as a function of C_{is} for C_{is} below 5.4%, but can be approximated by a linear function of C_{is} for C_{is} greater than 5.4% that includes the range over which the images for our psychophysical experiments were generated. This non-linear behavior of the ideal observer performance can be explained by the fact that, as we increase the extent of the tail of the PSF, the term $(s*t)(\mathbf{r})$ extends beyong the finite dimensions of the image that results in a truncation of the computation of d_{ideal} given by Equation (3.17). The detectability predicted by the nlf ideal observer, on the other hand, increases linearly as the contrast of the signal increases since the term $(s*t)(\mathbf{r})$ is being now ignored in Equation (3.32). Finally we can extract a minimum contrast perceived by the human observer of 4.5%, which is given by the intercept of the regression line d_{human} vs. C_{is} with the x-axis.

We conclude, then, that the results of the experiments indicate that the nlf ideal observer better predicts the performance of the human observer than the ideal observer does. However, the average efficiency of the human observer relative to the nlf ideal observer is only 10%, so we shall investigate if contrast enhancement, using high-pass filtering, would be of any help to the human observer. The main question is whether or not

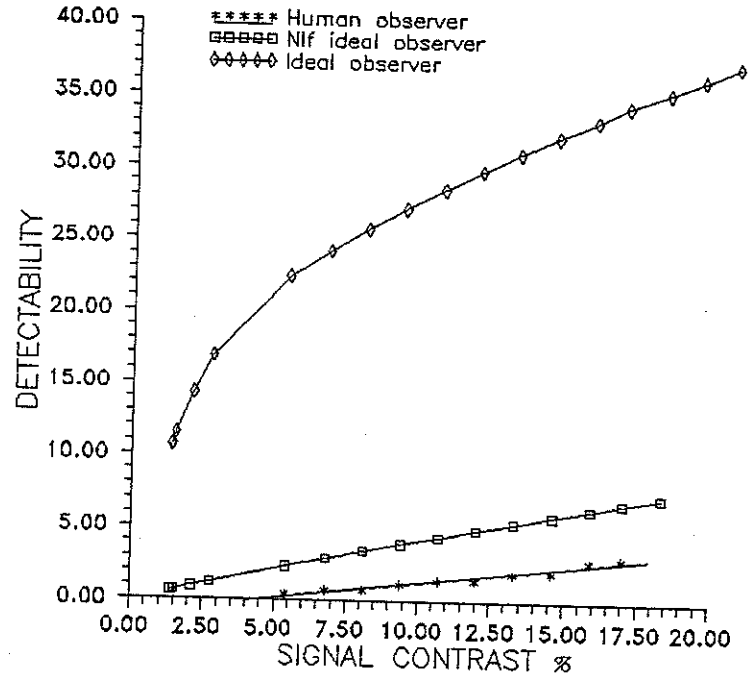


Fig. 4.4. Plot of the detectabilities associated with the ideal, the nlf ideal, and the human observers vs the contrast of the signal after imaging by a long-tailed PSF. The contrast of the signal decreases as the extent of the tail of the PSF increases.

the boost and the correlation in the noise induced by high-pass filtering will nullify the advantage of deconvolution. We shall now investigate the impact of deconvolution on the human observer performance.

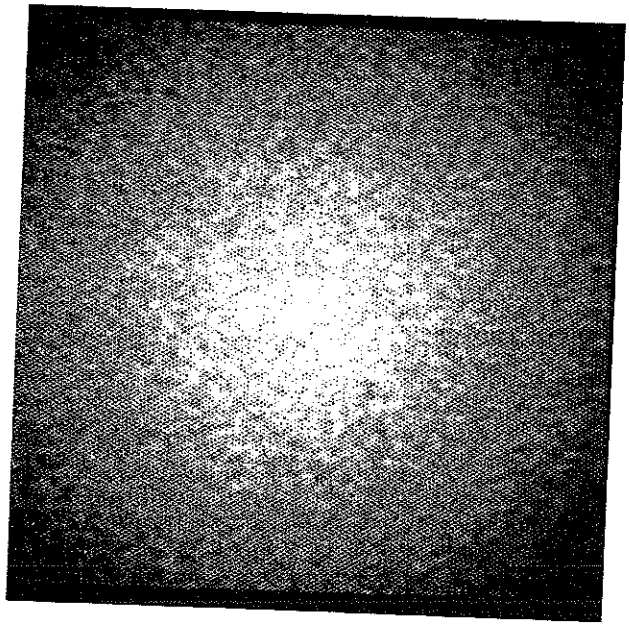
Methods and Results after Linear Deconvolution

Experimental Methods

We know that the impact of the tail of the PSF is to yield a broad signal given by $(s * t)(\mathbf{r})$ superimposed on a sharp signal formed by the core of the PSF and given by $(s * c)(\mathbf{r})$. In planar imaging, image restoration is a general term to describe a technique that allows the restitution of the object in its planar original form. We described in chapter 3 the deconvolution operation used to deblur the images. If no noise were present in the image, the deconvolution would yield a perfect restoration of the object (s+b)(r). For each width of the tail of the PSF, as given in Table 3.1, a filter $\tilde{f}_h(\rho)$ was designed using Equation (3.35) and applied to the computer-simulated images generated for the previous study. An example of four filters used to perform deconvolution was given in Figure 3.4. In the Fourier domain, the edges of a 128x128 image correspond to a maximum spatial frequency of 0.5 cycles/pixel. An example of an image which was presented to the human observers before deconvolution is shown in Figure 4.5 on the left-hand side, while the same image after deconvolution is shown on the right-hand side.

The experimental protocol was similar to the study before deconvolution. Each image was displayed using eight bits of greylevels. The images were viewed by the same six observers using the same certainty scale. Sample images were again displayed to the observers to clarify the object parameters and location before the study was run.

(a)



(b)

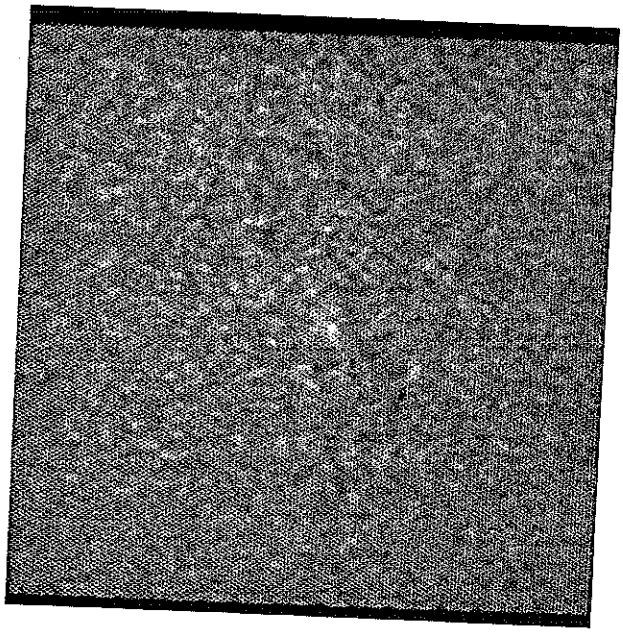


Fig. 4.5. An example of (a) a computer-simulated image before deconvolution and (b) the same image after deconvolution.

Results

The analysis carried out on the data is equivalent to the analysis performed on the data of the previous study, and results are presented in Table 4.3. By comparing the detectability values given in column 1 of Table 4.3 with the values given in column 1 of Table 4.2, we find that the performance of the human observer has improved due to linear filtering, but is still far below the prediction of the ideal observer, which is invariant under linear filtering as shown in chapter 3. In fact, the efficiency of the human observer with respect to the ideal observer is still as low as 0.6% after deconvolution. Such an efficiency again shows that the ideal observer is a rather poor indicator of human observer performance for the problem at hand, although the correlation is again high.

Let us now look at the performance of the human observer relative to other statistical observers. If we look at column 3 of Table 4.3, we observe, after deconvolution, a significant increase in efficiency of the human observer relative to the nlf ideal observer across the ten experiments. The average efficiency over the ten experiments of the human observer with respect to the nlf ideal observer has increased from about 11% before deconvolution to about 20% after deconvolution. A plot of the performance of the nlf ideal observer versus the human observer shown in Figure 4.6 shows that deconvolution does help the human observer, while the nlf ideal-observer performance remains constant as shown by the vertical shift of the points after deconvolution. We can now compare this result with the performance that would result from applying a simple matched filter on the deconvolved images. The construct of the npw ideal observer applied to the long-tailed PSF problem was given in chapter 3, and an illustration of the feature operator was shown in Figures 3.9 and 3.10. The results, given in Table 4.3, show a high correlation of the human observer with the npw ideal observer with a value of 0.99 and also an efficiency of 29%. The performance of the human observer relative to the nlf and the npw ideal observers is

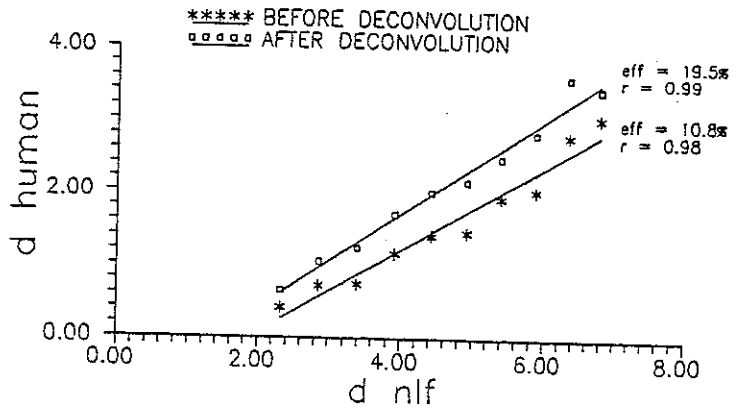


Fig. 4.6. Plot of the detectability values obtained from the psychophysical studies using images before and after deconvolution vs the detectability values predicted by the nlf ideal observer.

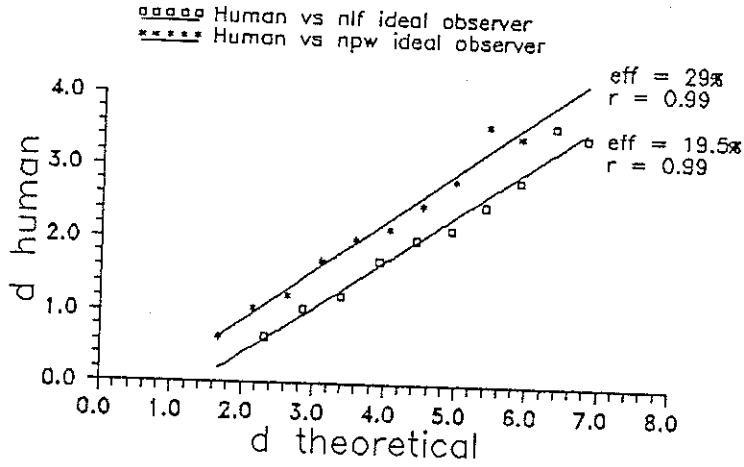


Fig. 4.7. Plot of the detectability values obtained from the psychophysical studies using deconvolved images vs the detectability values predicted by the nlf and npw ideal observers.

graphed in Figure 4.7. We note that the regression lines are almost parallel, and that the increase in efficiency from the nlf ideal observer to the npw ideal observer is shown by an upward shift of the regression line.

Discussion

The study of the impact of long-tailed PSFs on detection before deconvolution led to the important result that both the ideal observer and the npw ideal observer are poor figures of merit for the problem at hand, the efficiencies taking values of less than 1%. We found that the nlf ideal observer seems to better predict human performance, leading to an average efficiency of 11% over the set of ten experiments. Such a high discrepancy between the ideal or the npw ideal observers and the nlf ideal observer seems to indicate that the human observer cannot use the signal information encoded in the low-frequency component $(s*t)(\mathbf{r})$.

We then showed that linear deconvolution does help the human observer in the detection task, but that the improvement is such that the efficiency of the human relative to the ideal observer increases only from 0.3 to 0.6%. This small increase indicates that the ideal observer is certainly not the appropriate figure of merit to describe the problem of long-tailed PSFs, not only before deconvolution but after deconvolution as well. The fact that the ideal observer does not predict human performance after deconvolution is not so surprising if we recall the work of Myers et al. (1985). They found very low efficiencies for the human observer with respect to the ideal observer when images with negative noise correlations were considered and showed that the human observer is unable to prewhiten the noise. The negative correlations in the noise resulted from filtering the images using high-pass filters, that is similar to our study where we also use high-pass filters to deconvolve our images. They accepted the npw ideal observer as a better figure of merit to predict human-

observer performance. We also find that the npw ideal observer yields good efficiencies when compared to the performance of the human observer using the deconvolved images. We thus ask whether this increase in performance of the human observer is due to deconvolution or is simply a result of displaying the images in a more efficient way (Cormack 1980), or both. Thus, to show unambiguously that deconvolution does help the human observer in the detection task, we shall investigate the impact of the display process on the performance achieved.

Display Mappings and Human Performance

We already stressed the point that an imaging process may be conceived of as a two-stage process, a detection stage encompassing the acquisition and the recording of the data, and a processing and display stage where the data are displayed in the form of an image to a human observer. The former stage of imaging has been described in detail using statistical decision theory, while one aspect of image processing, high-pass filtering, was also investigated. We want to address now an important question, which is to find what is the impact of the CRT-display factors on human-observer performance. We shall be looking more specifically at two main factors: the utilization of the available grey-scale levels and the photometric linearization of the greylevel-to-brightness transfer curve.

We shall show an example where the utilization of the full scale of available greylevels is beneficial to the human observer, at least for a nonlinear grey-scale transfer curve. We shall also show that linearization of this curve, despite its importance in comparing displays, leads to a significant loss in human observer efficiency for the problem at hand.

Utilization of the Available Grey-Scale Levels

To prove that deconvolution did help the human observer in the detection task, we investigated the effect that linear grey-scale mappings have on human observer performance for the processed and unprocessed images.

To study this effect we first carried out a psychophysical study with the computer-generated images, before and after deconvolution, and displayed the images without greylevel manipulation. In fact, the values of the unprocessed images, as generated, ranged within the greylevel values of 0 to 255 and could therefore be displayed without scaling using the default CRT greylevel-to-brightness transfer curve that was given in Figure 4.1. For the processed images, on the other hand, their range of greylevel values fell outside of the 0 to 255 greylevel range, and the images had to be scaled--without any stretch--before display. To perform this operation of scaling without stretching, we looked, among all the images, for the image with the highest range of values and derived the appropriate scaling factor, common to all the images, to multiply each image of the set by, before display.

The results of these studies are presented in Figure 4.8, where we plot the detectabilities of the human observer, calculated using ROC analysis, vs. the detectabilities predicted by the nlf ideal observer. The two lower regression lines correspond to the performance, before and after deconvolution, for the images displayed without stretch. The observed increase in slope from 0.49 to 0.57 shows an improvement in performance due to deconvolution apart from the improvement that would follow contrast stretching. The images were then stretched over the full range of available greylevels. The results are given by the two upper regression lines plotted in Figure 4.8. Results show an increase in slope for both processed images (from 0.49 to 0.57) and unprocessed images (from 0.57 to 0.65). An additional upward shift in intercept results for the processed images after stretch. The

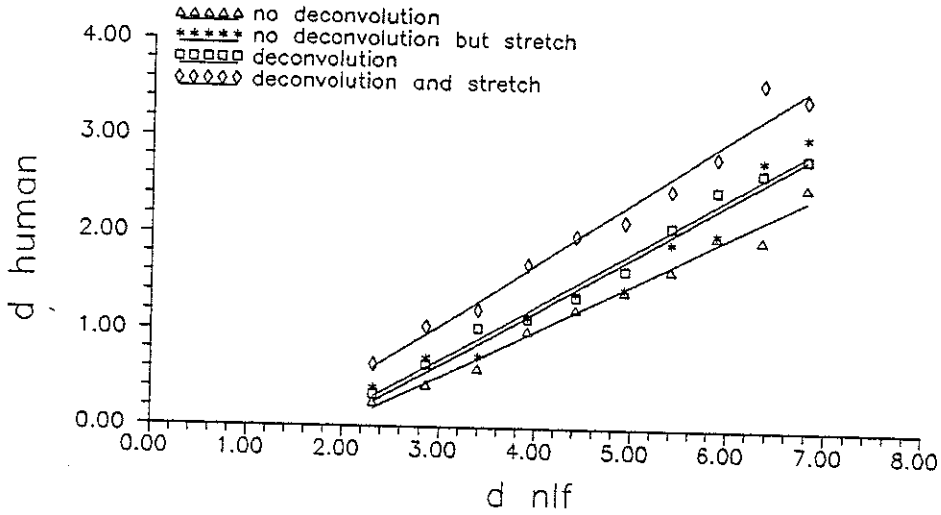


Fig. 4.8. Plot showing the effect of stretch of the images on human performance when graphed vs the performance of the nlf ideal observer. The two lower regression lines correspond to the performance before deconvolution, while the two upper curves correspond to the performance after deconvolution.

overlap of the two central lines tells us that by performing either deconvolution or linear stretch on the images, the performance is increased by about the same amount. However, if we now apply deconvolution and stretch sequentially, the performance is greatly improved. From these results, we see that deconvolution plus stretch leads to the highest performance of those just described for the observer/display system, but also that deconvolution alone does help the human observer in the detection task.

Questions still remain however. Will further stretch on the images improve performance, and what is the impact of the shape of the greylevel-to-brightness transfer curve on the performance? We shall answer the first question by looking at the effect of further greylevel windowing. We performed a more extensive stretch by applying a greylevel window to the image before deconvolution. The greylevel window used in the study consisted of keeping half of the full range of the actual greylevel values in the images, without altering the signal. The results are presented in Figure 4.9. They show that the performance achieved with this windowing technique is essentially equivalent to the performance achieved when deconvolution and stretch were applied sequentially. To get an idea of the different types of images, that is before deconvolution, after deconvolution, stretched images, and windowed images, Figure 4.10 shows on the top left an example of an image before deconvolution and stretch, and on the top right, the same image after deconvolution. The bottom right image represents the effect of deconvolution and stretch using 256 greylevels, while the bottom left image illustrates the greylevel windowing technique on the unprocessed image.

Linearization of Grey-Scale Transfer Function

There is still another concern with the display of the images. We need to understand the effect of the shape of the greylevel-to-brightness transfer curve on the

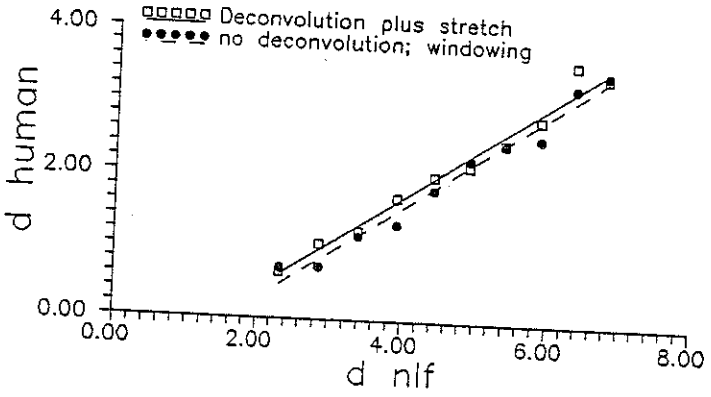


Fig. 4.9. Comparison of the performance of the human observer vs the nlf ideal observer after deconvolution or grey-level windowing. An example of the images seen by the human observers after deconvolution and stretch is given in the lower right corner of figure 4.10, while an example of the images resulting from grey-level windowing is shown in the lower left corner of figure 4.10.

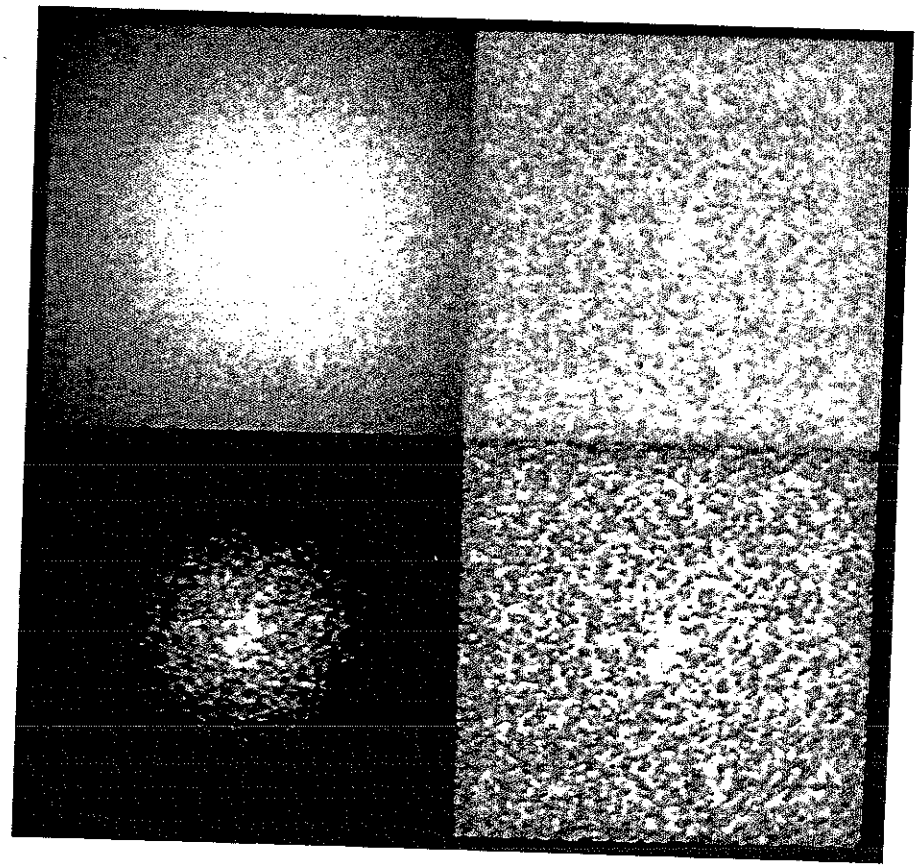


Fig. 4.10. An example of four computer-simulated images with : (a) upper left, before deconvolution and before stretch. (b) upper right, after deconvolution but before stretch. (c) lower right, after deconvolution and stretch. (d) lower left, before deconvolution but with grey-level windowing using 256 grey-levels on only a central part of the image.

performance, before and after deconvolution. It was suggested by Burgess (1988) that we verify that the increase in performance resulting from deconvolution is not simply due to a different location of the signal on the greylevel scale. If we look, in fact, at the transfer curve given in Figure 4.1, we observe different regions of increase in brightness along the curve due to different local slopes of the transfer curve, with a maximum magnification in the linear part of the curve. To clarify the effect of deconvolution on the images, apart from the role of the display characteristic curve, Burgess suggested linearizing the greylevel-to-brightness transfer curve and repeating the study of deconvolution. We shall now describe the procedure used to linearize the display photometrically and present the results of the experiments. This procedure of linearization is not to be confused with perceptual linearization as described by Pizer 1981.

Because of the need to increase the control over the display parameters and the experimental conditions associated with the psychophysical studies, we decided that for any following studies we would drive our display monitor with a PCvision board. This decision allowed us to have a system especially for the purpose of psychophysical studies, such that the control over the experimental conditions, viz. the contrast and brightness of the CRT display, and the environmental setup in general, could be facilitated. Since the previous studies were run using an imaging board controlled by the VAX, we measured the transfer curves of our CRT (Sierra Scientific Corporation) for the two boards, the board controlled by the VAX and the PCvision board.

To photometrically linearize a CRT display, we first needed to measure the transfer charateristic-curve of the CRT, that is the relation between the input greylevel and the corresponding displayed brightness value at the CRT screen. CRT displays have been characterized accurately, for example, by Hans Roehrig et al. (1989), who considered, among other factors, the nonuniformity in brightness across the CRT screen and the

existence of scatter within the phosphor of the screen. Because we wanted to make measurements as accurate as possible, we decided to control for both of these effects even though they were small. We chose to measure brightness within only a 16x16 array at the center of the screen; we called this array our measurement area.

For precise control of these measurements, we needed to account for the ambient light and two types of scatter: light scattered in the imaging lens and light scattered from activated phosphor surrounding the measurement area. We first created a dark-room environment and then imaged the 16x16 array onto the sensitive area of a photodiode such that the photodiode only saw the light coming from the measurement area. While the first type of scatter always needs to be eliminated, the second type of scatter must, on the other hand, be included as part of the measurements because this type of scatter is part of the interactive process that takes place in all CRTs. Any one pixel's brightness is affected by the scatter within the CRT screen from surrounding bright pixels. Thus if we want an accurate measure of a displayed greylevel, we must allow for the scatter from the surrounding area. To account for the scatter from surrounding areas and its effect on the measured greylevel brightness, we set a 128x128-pixel area around the 16x16 array to a greylevel of 128 during the whole measurement sequence. A greylevel value of 128 was used because this was the average mean value of the images used in the phychophysical studies.

The measurements of brightness were made using a 55 mm camera lens as the imaging device and a photodiode of 95 mm^2 sensitive area as the detection device. The 16x16 array was displayed in the center of the CRT screen and imaged by the lens on the photodiode. The greylevel values assigned to the suround were 0, 128, and 255. A value of 128 was used to simulate the mean value of the images that were used for psychophysical studies, while the values 0 and 255 were used during calibration of the imaging system. The

calibration sequence of the imaging system was as follows. We first proceeded to an alignment of the system to form a focused image of the 16x16 square array on the photodiode. To bring the object into sharp focus, the easiest way was to display a computer-simulated fine grid on the CRT screen and to form its image on the photodiode. The grid was brought into focus by opening the lens to F/2.8 and translating the imaging system (lens and detector) on its xyz translation table. We then displayed a 16x16 area centered in a 128x128 surround of greylevel 255 while the value of the 16x16 array was set to zero. We verified that the photodiode saw only the central area, and not the surround, by scanning transversely the image formed on the photodiode with a 5 mm field stop positioned at the photodiode plane. The results of the scan, presented in Figure 4.11, show a significant flat plateau in the center of the scan that corresponds to the detector seing only the greylevel value of zero. The rising edges, on the other hand, correspond to a field of view that includes more and more suround, up to the second plateau at the edges of the scan, which occurs when the photodiode saw only the surround of greylevel value 255. This calibration step was to ascertain that the scanning aperture was of the apropriate size to capture most of the 16x16 array but none of the surround. The last step in the calibration was to prevent any light emitted by the surround from reaching the detection area due to its scattering in the imaging lens. To avoid this scatter, we first stopped down the aperture of the camera to F/16. We then measured the residual scatter by superimposing a black piece of material of density greater than 4 on the 16x16 array already set to zero and measured the amount of light reaching the detector. Since we still measured some residual scatter occuring in the lens after closing it down to its smallest aperture size, we used a 10 mm circular aperture positioned at the entrance plane of the camera lens to serve as a baffle. The size of this aperture was found to be sufficient to make the system immune to scatter.

We then set the surround to a greylevel value of 128 to simulate the display of a

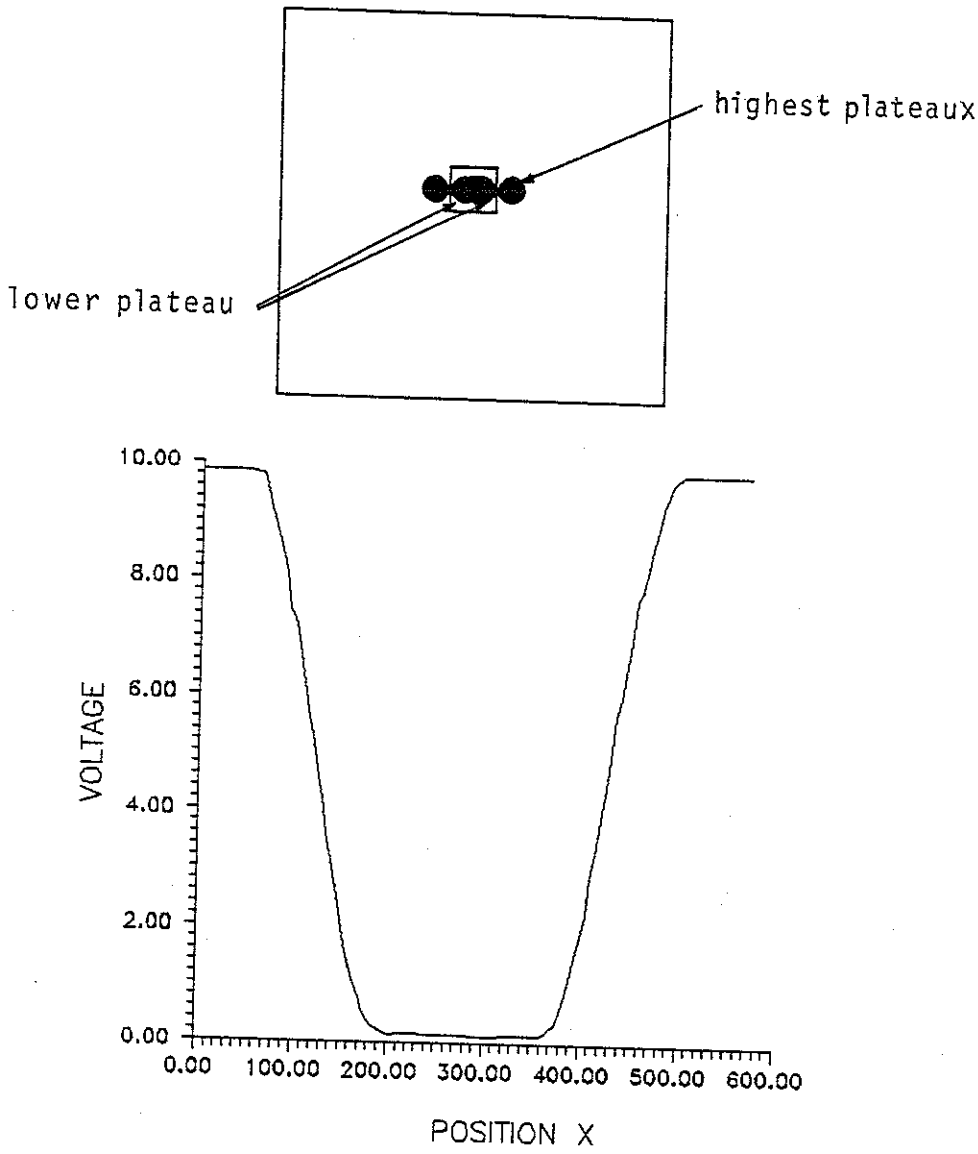


Fig. 4.11. Plot of the profile of the relative brightness accross an image of a 16x16-pixel array of grey-level value zero centered in a 128x128-pixel array of grey-level value 255. The image was formed on a photodiode by a 55 mm camera lens, and the relative brightness was measured as a voltage as indicated on the figure. The image was scanned using a 5 mm field stop positioned as close as possible to the photodiode. In the central flat plateau the scanning aperture overlapped only with the 16x16-pixel array. It overlapped part of the 128x128-pixel array and part of the 16x16-pixel array for the rising portion of the curve. Finally it overlapped only the 128x128-pixel array for the two highest plateaux at the edges of the profile.

real image. If the mean level of an image is 128, as it is generally the case in our psychophysical studies, we assume that the scatter reaching one pixel from its surrounding pixels would be in average the same as the scatter caused by a constant surrounding assigned to this mean greylevel value. The greylevel value of the measurement area varied from 0 to 255 in steps of 16. Figure 4.12 shows the results of the measurements when the displayed greylevels were driven either by the VAX board or the PCvision board, respectively. These measured data points were then fit to a polynomial of degree 7, and the fitted curve was used to derive a lookup table (LUT) needed to linearize the greylevel-to-brightness transfer curve.

We found a change in the shape of the transfer curve of the same CRT monitor when driven by a PCvision board instead of the display board on our VAX 8600 as shown in Figure 4.12. We postulate that this difference in shape of the two transfer curves is due to the different voltage levels of the video signal at the output of the two boards as shown Figure 4.13. No significant difference was found, however, due to the setting of the surround to a greylevel value of 0 instead of 128 as shown in Figure 4.12 (b) by the superposition of the two curves for those two values of the surround. To verify that the LUT provided a good photometric linearization of the transfer curve, we repeated the measurements using the LUT. The linear measured curves are shown in Figure 4.12 as straight lines.

Stretch and Linearity

We carried out two psychophysical studies to investigate the effect of using different boards to drive the CRT display and the effect of linearization of the CRT on human performance. For each study, we ran 5 of the 10 experiments designed to study the effect of deconvolution for the long-tailed PSFs images. We ran experiments 1, 3, 5, 7, and

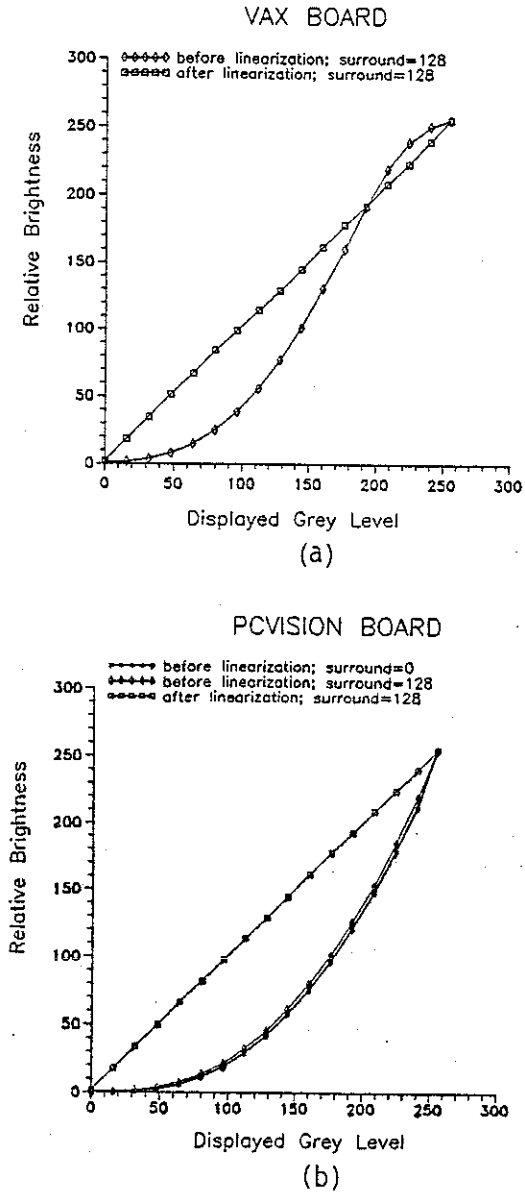


Fig. 4.12. Plots of the measured relative brightness vs the displayed grey-levels at the CRT screen. The measurements were done with a photodiode/55 mm camera assembly looking at a 16x16-pixel array of grey-level values ranging from 0 to 255. The 16x16-pixel array was centered in a 128x128-pixel array of grey-level value 0 or 128 and referred to as the surround. The display monitor was driven by either (a) a VAX board or (b) a PCvision board. The measurements were repeated after photometric linearization of the display and the results are shown on the same plots as straight lines.

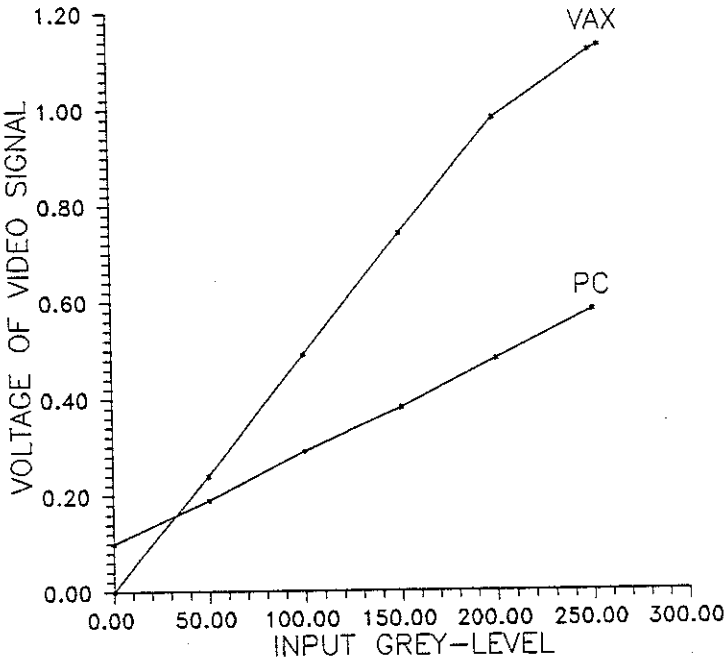


Fig. 4.13. Plot of the voltages of the video signals generated from the numerical values of the displayed grey-levels for two driving boards, a VAX board and a PCvision board. This difference in values explains the difference in the shape of the transfer curves for the two boards as shown in figure 4.12. The display monitor was the same for the two driving boards.

9 with the images displayed using eight bits of greylevels. The images covered 5.4 x 5.4 cm on the screen and each pixel measured 0.42 x 0.42 mm^2. The light in the room was dimmed down to the same value for all the observers. The contrast and brightness levels of the display monitor were fixed during the study and the observers were not allowed to vary them during the study. Each subject viewed the images binocularly at a comfortable distance chosen by the observer, and the observers were required to wear their usual correction lenses.

Six observers ran the study, each performing all 5 experiments. The observers first viewed sample images to ascertain the shape, size and location of the signal to detect. They did not have any information on the number of occurrences of the signal. The images were displayed in a random order either with the non-linearized transfer curve of the display at fixed contrast and brightness levels, or using the LUT derived to photometrically linearize the display. The 5 experiments were also performed in a randomized order for each observer. The observer was then presented a single stimulus on each trial, either the signal on a noisy background or the noisy background alone. The observer was instructed to rate his certainty of seing the signal on a six-point scale, and the responses were recorded and then analysed using ROC analysis.

The first psychophysical study compared detection using two different boards to drive the same CRT monitor. We found that the shape of the transfer curve, found when the monitor was driven with the PCvision board, led to a loss in efficiency compared to the performance resulting from driving the monitor with the VAX board as shown in Figure 4.14 (a) and (b). The loss in efficiency when going from 10% to 6.5% and 19% to 13% was 3.5/10 x100 $\simeq$ 35% and 6/19 x100 $\simeq$ 32% before and after deconvolution, respectively. This difference in performance shows how a difference in shape of the greylevel-to-brightness transfer curve can strongly affect human observer performance.

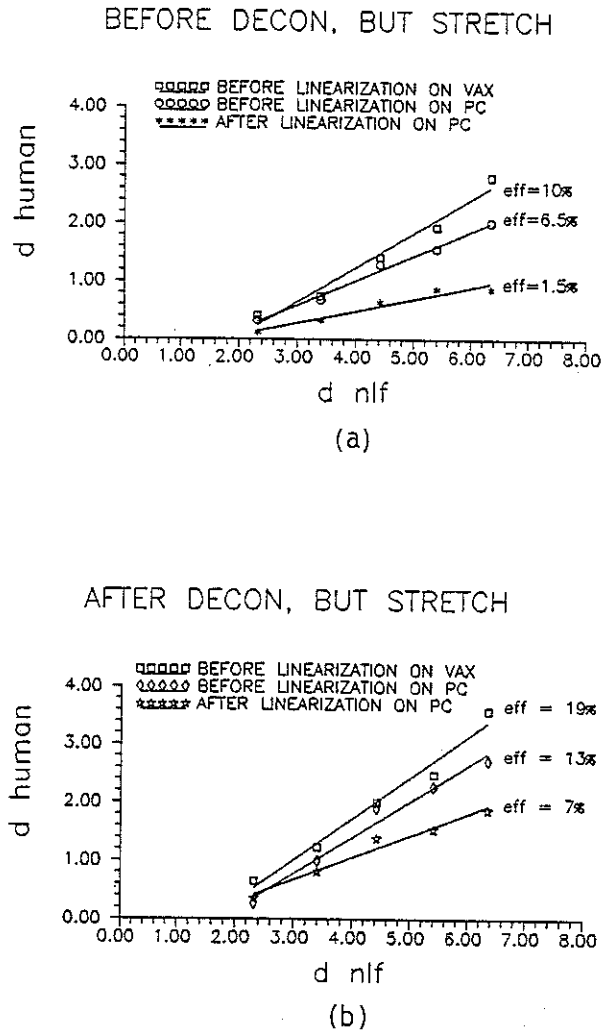


Fig. 4.14. Plots of the detectabilities obtained from the psychophysical studies using linearized and non-linearized displays vs the detectabilities predicted by the nlf ideal observer: (a) before deconvolution, but with stretch (b) after deconvolution and stretch.

The second psychophysical study investigated the effect of linearization of the transfer characteristic curve on perception. Because of ease of control we used the PCvision board to run this study. The results, also given in Figure 4.14, show a loss in efficiency of about 77% and 46% before and after deconvolution, respectively. The results of the linearization of the display are summarized in Figure 4.15 for the PCvision board before and after deconvolution. It is interesting to see that the measured performance of the human observer is about the same for the images before deconvolution and linearization than after deconvolution and linearization as shown by the two overlapping central regression lines.

Discussion

First, we have shown that setting the window and the level of the displayed images does affect the performance of the human observer when performing the detection task. Second, after linearization of the greylevel-to-brightness transfer curve, the performance of the human observer has decreased significantly.

To discuss the first finding, let us again consider Figure 4.10 which displays some examples of images containing a signal. First, note that both the signal and the noise become more visible after either deconvolution followed by stretch or simple greylevel windowing before deconvolution. We expected the performance of the human observer to remain constant after processing, however, unless the contrast of the signal was below the contrast threshold of the eye before processing (Blackwell 1946; Barlow 1957; Nagaraja 1964; Goodenough et al. 1975; Judy and Swensson 1985; Seltzer et al. 1988).

Several factors can be the source of a reduction in human efficiency: first, the inefficient filtering that occurs, for example, if the human observer systematically cross-correlates the images with a larger or a smaller template, or if the diameter of the disc signal to detect is larger than about 1 degree of visual angle (Burgess 1986). In our experiments,

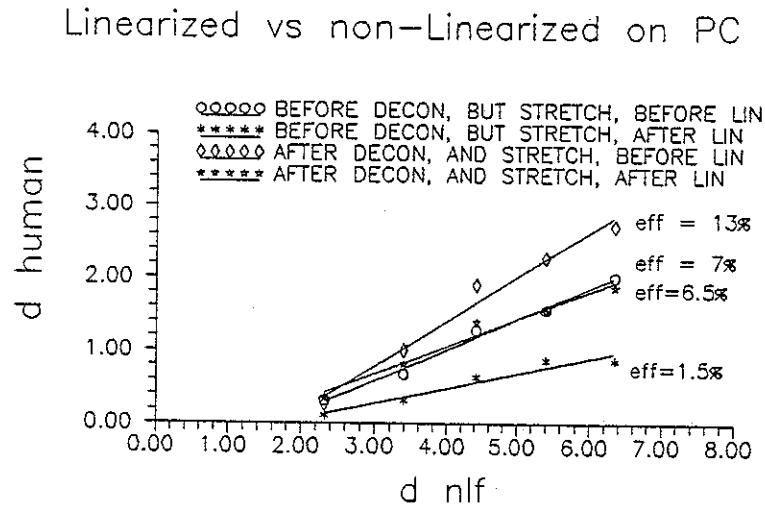


Fig. 4.15. Plot to summarize the performance obtained using the PCvision driving board for linearized and non-linearized displays, before and after deconvolution.

the fact that the signal was seen under about 0.6 degree of visual angle (the signal being about 5 mm wide and the viewing distance about 50cm), implies that this type of error was not the limiting factor. Another factor is the uncertainty within the observer about the given signal information, which has been discussed in great detail in an excellent paper by Pelli (1985), but which has also been discussed by Swensson and Judy (1981), and Hanson (1983); and more recently by Wagner and Barrett (1987), Brown and Insana (1988), Myers and Wagner (1988), Wagner et al. (1989), and Wagner et al. (1990). We should expect this type of error to remain the same before and after processing of the images, however. The third factor is the intrinsic source of variability within the human observer's eye-brain system, commonly referred to as internal noise, which can explain the observed contrast threshold of the eye. We shall distinguish between two sources of variability: one component, referred to as constant internal noise, does not depend on the image noise level, but is a function of the mean luminance in the image (Cattell 1893; Griffiths and Nagaraja 1963; Wickelgren 1968). The other component, referred to as induced internal noise, is a function of the amount of noise in the image and dominates when the noise is easily visible in the image (Burgess 1986).

Ishida et al. (1984) showed in a quite convincing paper how the improvement in detectability due to contrast enhancement using a windowing procedure can be explained by taking into account the constant internal noise level in the theoretical performance. They showed that, although the windowing technique increases both the signal contrast and the image noise by the same factor, the relative magnitude of internal noise with respect to the perceived noise in the image is reduced by image contrast enhancement, and detectability is therefore improved. These results seem to agree with our findings since the noise clearly became more visible after processing in our experiments. Moreover, as the perceived noise in the image increases after processing, the contribution of the constant internal noise to the

total noise decreases, and Ishida et al. showed that the performance reaches a plateau where no further improvement can be achieved. This result may also explain the result that we found with the greylevel windowing or clipping operation, where we found that the performance was as good as for the deconvolution plus stretch, but not better in spite of the the maximum contrast stretching applied. We shall note, however, that, although it was interesting to explore the effect of clipping part of the image that was outside of the signal region and to stretch the remaining portion of the image, this technique cannot usually be applied unless the exact location of the signal is known, while deconvolution of the images can be applied in any case as long as we know the parameters of the PSF of the imaging system.

The second finding, that is the decrease in detection following the linearization procedure, can be explained by the reduction in slope of the transfer curve in the greylevel region in which the signal resided. It should be noted that, although the transfer curve depicted in Figures 4.12 (a) and (b) are commonly encountered in a large range of displays, other transfer curves are possible, and some of them would certainly lead to different effects. For example, if linearization has the effect of increasing the slope of the transfer curve in the range of greylevels in which the signal resides, then detection would be most likely increased. However, for the range of displays that our transfer curves represent, photometric linearization is obviously not a good candidate to be used to evaluate human performance, since the displayed information can be increased when using a different transfer curve. Moreover, for the task at hand, the efficiencies we were working with before linearization were well below 50% as often reported, for example, by Burgess in his experiments (Burgess 1984 and 1985). We expect that linearization would have a stronger effect on the performance for those low efficiencies. This reduction in performance might have been more subtle if we were to work between 25% and 50% efficiencies, but our study

demonstrates how important it is to control for the display parameters, to report the shape of the display transfer curve in parallel with the psychophysical results, so that more unity among different groups of research can be achieved comparative studies carried out.

CHAPTER5

NONUNIFORM BACKGROUNDS

We have encountered cases where the performance of statistical observers could be easily calculated when both the signal and the background were known exactly. The performances of these observers were then compared to human performance to determine if they could be used as a basis for system assessment and optimization. It was implicitly assumed in this approach that a system optimized for a model observer and an exactly specified task would also be optimal not only for the human observer, but also for more realistic tasks. To verify the validy of this assumption, we propose to increase the complexity of the task by relaxing the property of an exactly specified background. The main objective of this study is to provide some starting points towards assessing image quality for more realistic images. The realisation of this objective should directly lead to a better understanding of the trade-off between resolution and sensitivity, as specified by the imaging aperture size. We shall start to motivate the interest in studying background uncertainty by showing, by the means of a simple example, a paradox between statistical decision theory and clinical experience. We shall then review some of the pioneer works that strongly triggered our interest to study the problem of background uncertainty as well.

If we consider the problem of aperture design when the task is the detection of a completely specified signal $s(\mathbf{r})$ in an infinite, uniform background of known strength B_0, we shall show that the ideal-observer strategy leads to an optimum aperture that is infinite in spatial extent. Let us describe the signal by a 2D Gaussian function of the form

$$s(\mathbf{r}) = \frac{a_s}{\pi r_s^2} \exp\left(- |\mathbf{r}|^2 / r_s^2 \right) \quad , \tag{5.1}$$

where $\mathbf{r}$ is a 2D position vector, r_s is the half width at 1/e of the peak intensity, and a_s represents the total energy under the signal. In a two-hypotheses detection task, the mean object $\overline{f(\mathbf{r})}$ can be described by either a constant background of strength B_0 under hypothesis H_1, or as the signal $s(\mathbf{r})$ superimposed on the constant background of strength B_0 under hypothesis H_2. For a task where the object is specified completely, $\overline{f(\mathbf{r})}$ and $f(\mathbf{r})$ are equivalent, and we shall then refer to the mean object as $f(\mathbf{r})$. If we follow the formalism of image formation described in chapter 2 (see equation (2.6)), a mean image $\overline{g(\mathbf{r})}$ is then given by

$$\overline{g(\mathbf{r})} = \kappa \, T \, \left(f(\mathbf{r}) * p(\mathbf{r}) \right) \quad , \tag{5.2}$$

where κ is an efficiency factor describing the geometry of the imaging system, $p(\mathbf{r})$ is the transmittance of the aperture when projected onto the image plane, and T is the exposure time. Let us describe the imaging aperture as a Gaussian pinhole of the form

$$p(\mathbf{r}) = \exp\left(- |\mathbf{r}|^2 / r_p^2 \right) \quad , \tag{5.3}$$

where r_p is the half width at 1/e of the peak transmittance measured in the image plane. We derived in chapter 2 an expression for the detectability index d_{ideal} in the form of a discrete sum of the ratio of the squared difference of the expected signal to the mean background as shown by equation (2.43). As the area of the sampling element approaches zero, equation (2.43) can be rewritten in integral form as,

$$[d_{ideal}]^2 = \int d^2\mathbf{r} \; \frac{[\Delta\bar{g}(\mathbf{r})]^2}{g_1(\mathbf{r})} \quad , \tag{5.4}$$

with

$$\Delta\bar{g}(\mathbf{r}) = \kappa T \; (s(\mathbf{r}) * p(\mathbf{r}))$$

$$= \kappa T \; \frac{a_s \, r_p^{\,2}}{r_s^{\,2} + r_p^{\,2}} \; \exp\left(- \frac{|\mathbf{r}|^2}{r_s^{\,2} + r_p^{\,2}} \right) \quad , \tag{5.5}$$

and

$$\bar{g_1}(\mathbf{r}) = B_0 * p(\mathbf{r}) = \kappa T \, A_{ap} \, B_0 \quad , \tag{5.6}$$

with A_{ap} equal to $\pi r_p^{\,2}$, the aperture area. By substituting equations (5.5) and (5.6) into equation (5.4), the expression for d_{ideal} becomes

$$[d_{ideal}]^2 = \int d^2\mathbf{r} \left[\kappa T \; \frac{a_s \, r_p^{\,2}}{r_p^{\,2} + r_s^{\,2}} \; \exp\left(- \frac{|\mathbf{r}|^2}{r_s^{\,2} + r_p^{\,2}} \right) \right]^2 \frac{1}{\kappa T \, \pi \, r_p^{\,2} \, B_0}$$

$$= \kappa T \; \frac{a_s^{\,2} \, r_p^{\,4}}{(r_p^{\,2} + r_s^{\,2})^2 \, (\pi \, r_p^{\,2} \, B_0)} \int d^2\mathbf{r} \; \exp\left(- \frac{2 \, |\mathbf{r}|^2}{r_s^{\,2} + r_p^{\,2}} \right)$$

$$= \kappa T \; \frac{a_s^{\,2} \, r_p^{\,2}}{2 \, B_0 \, (r_p^{\,2} + r_s^{\,2})} \left\{ \frac{2}{\pi \, (r_p^{\,2} + r_s^{\,2})} \int d^2\mathbf{r} \; \exp\left(- \frac{2 \, |\mathbf{r}|^2}{r_p^{\,2} + r_s^{\,2}} \right) \right\}$$

$$= \frac{\kappa T \; a_s^{\,2} \, r_p^{\,2}}{2 \, B_0 \, (r_p^{\,2} + r_s^{\,2})} \quad . \tag{5.7}$$

Equation (5.7) shows that, for a given imaging time T, the detectability is optimal for r_p infinite and tends to the asymptotic value $\kappa\ T\ a_s^2\ /\ 2\ B_0$. In other words, for the task at hand, resolution seems to be useless from the ideal-observer point of view. Equation (5.7) also shows that, for an aperture of a given size r_p , the detectability increases linearly with exposure time and tends to infinity as T tends to infinity.

In a generalization of the above task, Tsui et al. (1978) calculated the performance of three model observers as a function of aperture size for the problem of detection of a spherical tumor in a finite uniform background of unknown strength. The realism of the problem was increased further by assuming a finite detector area. One observer model in the Tsui paper closely approximated the Bayesian strategy for the task. Because the flat background was known only in a statistical sense, the observer had to make an estimate of the background in an annular region about the signal and compare this estimate to an estimate of the activity level at the location of the expected signal. The authors optimized the integration areas of the two regions of estimation for each aperture size they considered. The uncertainty in the flat background level together with the finite image field resulted in a finite optimum aperture size for maximum observer detectability. In other words, spatial resolution was an important system design parameter for this task.

More recently, an interesting problem was presented by Wagner et al. (1981). They considered a high-contrast Rayleigh task in which the ideal observer was asked to determine whether a noisy scene was the image of a single or double Gaussian signal. The Rayleigh discrimination task was first suggested as a useful test of imaging system performance by Harris (1964). Wagner et al., considered a SKE/BKE task in that the width of the sources and the separation of the pair were known, as was the strength of the uniform background on which they were superimposed. Both the background and the detector were assumed to have infinite spatial extent. Since the only randomness in the task

was due to Poisson statistics, the formulation of the Bayesian decision strategy was straightforward. Using a signal-detection framework, the authors determined how ideal observer performance varied as a function of source width and separation for three apertures: a simple pinhole, a large open aperture (twenty times the pinhole width), and a coded aperture known as the uniformly redundant array (URA). The authors found that the ideal observer performance was much better for the URA than for the small pinhole when the signal had high contrast. The ranking of the code and the large aperture depended on the source width and separation, with the large aperture outperforming the code for coarse tasks. The surprising conclusion of Wagner et al. was that the large aperture outperformed the simple pinhole for all source widths and separations.

We have just described studies of certain tasks that imply that a large open aperture is much better than a small pinhole, according to the calculated performance of the ideal observer. Yet, this conclusion does not seem to agree with current experimental practice in radionuclide imaging. For example, in the field of nuclear medicine, spatial resolution in the displayed image in the neighborhood of 1 cm seems to give the best subjective image quality. It is the purpose of our study to further understand the source of this paradox.

The detection and Rayleigh studies that found superior performance for large apertures have two common features. One is the complete specification of the background and signal, and the other is the infinite size of the detector area. To resolve the apparent contradiction between the results of these studies and current experimental practice, we shall investigate tasks in which the signal is still known exactly but the background is allowed to vary randomly from image to image and have a spatially varying ("lumpy") character. We shall show that even in the circumstance of infinite background and detector extent, a spatially varying background drives an imaging system to finite optimum resolution. Before

elaborating on performance calculations, we shall proceed to the description of two types of nonuniform or lumpy backgrounds which will be used in our computer-simulated images and in the calculation of statistical observers performance.

Lumpy Background Simulations

The increase in complexity due to the relaxation of the certainty in the background can be achieved in many different ways, but only certain models will lead to a description that is mathematically tractable. The mathematics involved in the formulation of a statistical observer performance require either the full knowledge of the probability densities of the images under the two hypotheses, as shown for the ideal observer in chapter 2, or the first- and second-order statistics as for the npw ideal and the Hotelling observers. In the case of the lumpy backgrounds, we may want to stay away from computing the ideal-observer performance since the full knowledge of the probability densities under the two hypotheses is generally difficult to derive or even to estimate as the complexity of the task increases. Rather, we shall consider less demanding statistical observers such as the Hotelling and the npw ideal observers. Some further assumptions can also be made to allow for the simplification of the mathematics involved. To compute the classification performance of these two statistical observers, we make the following assumptions: the first assumption is wide-sense stationarity, where the autocorrelation function is only a function of the shift variable $\mathbf{r}$. The second assumption is that the background autocorrelation function is a Gaussian function of the form

$$R_f(\mathbf{r}) = \frac{W_f(0)}{\pi r_b^{\,2}} \exp\left(-|\mathbf{r}|^2/2\, r_b^{\,2}\right) , \tag{5.8}$$

where r_b is the correlation length of the autocorrelation function and $W_f(0)$ the value of the power spectrum at zero frequency. The subscript f here emphasis the fact that the autocorrelation function refers to the object **f** rather than the image **g**. After imaging, we shall refer instead to R_g and W_g . The notation $W_f(0)$ comes from the notations used for the power spectrum since, for stationary statistics, the power spectrum $W_f(\rho)$ can be defined as the Fourier transform of the autocorrelation function, so that

$$W_f(\rho) = W_f(0) \exp(-2\pi^2 r_b^2 |\rho|^2) , \tag{5.9}$$

where ρ is the 2D frequency variable in the Fourier domain conjugate to **r**. We shall now proceed to a description of two simulation schemes that yield stationary statistics and Gaussian autocorrelation functions.

Lumpy Backgrounds of Type I

The first approach to lumpy backgrounds is to simulate uncertainty in the background by randomly superimposing Gaussian functions on a constant background of strength B_0 over the object space. We shall often refer to these Gaussian functions as Gaussian blobs or simply blobs. To keep the mathematics as simple as possible, we assume here that the Gausian blobs are of constant amplitude $b_0/\pi r_b^2$ and constant width r_b. One image formed with these Gaussian blobs can be described mathematically as the convolution of a set of delta functions of equal amplitudes and randomly located in the object space with a Gaussian function of constant strength and width. If we denote by $b(\mathbf{r})$ the lumpy component of the background, the expression for $b(\mathbf{r})$ is then given by

$$b(r) = \left[\sum_{j=1}^{K} \delta(r-r_j) \right] * \left[\frac{b_0}{\pi r_b^2} \exp\left(-|r|^2/(r_b^2) \right) \right]$$

$$= \sum_{j=1}^{K} \frac{b_0}{\pi r_b^2} \exp\left(-\frac{|r-r_j|^2}{r_b^2} \right) , \qquad (5.10)$$

where r_j is a random variable uniformly distributed over the object area, and K is the number of blobs in the background. The object area should be understood here as the projection of the actual object space onto the image plane such that the magnification between the object and the image planes is automatically taken into account. A detailed description of image formation was given for the case of a pinhole aperture in chapter 2.

The description of a nonuniform background as given by equation (5.10) is not sufficient, however, to yield a Gaussian autocorrelation function. The calculations of the autocorrelation function, presented in detail in appendix B, shows that the number of blobs K must be itself a random variable and more specifically it must be Poisson distributed. This can be seen clearly by simple inspection of equation (B14), where the expression for R_f reduces to a Gaussian function if the last two terms cancel out. The expression for R_f (or W_f) given by equations (B15) (or (B16) and (B17)) is an expression for the second-order statistics of the lumpy backgrounds of type I thus generated. The derivations given in appendix B lead to a measure of lumpiness given by equation (B17) as

$$W_f(0) = \frac{\overline{K}}{A_d} b_0^2 , \qquad (5.11)$$

where $\overline{K}/A_d$ is the mean number of blobs per mm^2 and b_0 is the strengh of the blobs and is

in units of counts per sec. The lumpiness is then expressed in units of counts^2/(sec^2 mm^2).

The first-order statistics can be expressed as the expected value of the background over the ensemble of images. The mean level $\overline{B}$ in the object is the sum of two terms, and its expression is given by

$$\overline{B} = B_0 + \langle\, b(\mathbf{r}) \,\rangle_{K,\mathbf{r}_j}$$

$$= B_0 + \left\langle \sum_{j=1}^{K} \frac{1}{A_d} \int d^2\mathbf{r}_j \; \frac{b_0}{\pi\, r_b{}^2} \exp\left(- \frac{|\mathbf{r}-\mathbf{r}_j|^2}{r_b{}^2} \right) \right\rangle_K$$

$$= B_0 + \frac{1}{A_d} \frac{b_0}{\pi\, r_b{}^2} \left\langle \sum_{j=1}^{K} \int d^2\mathbf{r}_j \; \exp\left(- \frac{|\mathbf{r}-\mathbf{r}_j|^2}{r_b{}^2} \right) \right\rangle_K$$

$$= B_0 + \frac{1}{A_d} \frac{b_0}{\pi\, r_b{}^2} \left\langle \sum_{j=1}^{K} (\pi\, r_b{}^2) \right\rangle_K$$

$$= B_0 + \frac{b_0}{A_d} \left\langle \sum_{j=1}^{K} 1 \right\rangle_K$$

$$= B_0 + \frac{b_0}{A_d} \langle\, K \,\rangle_K$$

$$= B_0 + \frac{b_0}{A_d} \overline{K} \quad , \tag{5.12}$$

where we assumed that A_d tends to infinity to derive line number 4. Note that B_0 is in units of counts/(sec mm^2), $\overline{K}/A_d$ is in units of counts/mm^2, and b_0 in units of counts/sec.

The mean background $\overline{B}$ is then expressed as a number of counts/(sec mm^2).

Lumpy Backgrounds of Type II

The second approach to lumpy backgrounds with a Gaussian autocorrelation function is to filter uncorrelated Gaussian noise. If the filter function is chosen to be a Gaussian function of correlation length r_b, a mathematical description of the background $b(\mathbf{r})$ generated from filtered uncorrelated Gaussian noise is best given in the discrete form

$$b(\mathbf{r}_j) = \sum_i a(\mathbf{r}_i) \frac{H(0)}{\pi r_b^2} \exp\left(- |\mathbf{r}_j - \mathbf{r}_i|^2 / r_b^2 \right)$$

$$= \frac{H(0)}{\pi r_b^2} \sum_i a(\mathbf{r}_i) \exp\left(-|\mathbf{r}_j - \mathbf{r}_i|^2 / r_b^2 \right) \quad , \tag{5.13}$$

where $a(\mathbf{r}_i)$ is a random variable normally distributed with mean value A_0 and standard deviation σ, with i specifying the i^{th} blob located at the i^{th} pixel, and $\mathbf{r}_j$ is the location of the j^{th} pixel in the background. Note that $b(\mathbf{r}_j)$ is obtained by performing a sum over all the pixels, $\mathbf{r}_i$ being a nonrandom variable, and that the randomness resides in the amplitude of the Gaussian blobs. This is a quite different approach from the lumpy backgrounds of type I where the location of the blobs was then the random variable, their amplitude being constant and nonrandom.

The first- and second-order statistics of these lumpy backgrounds $b(\mathbf{r})$ of type II, are derived in appendix C. A measure of lumpiness is now given by equation (C4) as

$$W_f(0) = W(0)\,[H(0)]^2 \quad . \tag{5.14}$$

The expression for R_f is derived in appendix C and is given by equation (C5). For stationary processes, the mean of the random process b(**r**) can be evaluated by taking the ensemble average over the random variable,which is, in this case, the amplitude of the noise impulses at each pixel location in the image. We find

$$\begin{aligned}
\overline{B} &= \langle\, b(\mathbf{r}) \,\rangle_a \\
&= \left\langle\, \frac{H(0)}{\pi\, r_b^{\,2}} \sum_i a(\mathbf{r}_i)\, \exp\left(- |\mathbf{r}_j - \mathbf{r}_i|^2 / r_b^{\,2} \right) \right\rangle_a \\
&= \frac{H(0)}{\pi\, r_b^{\,2}} \sum_i \langle\, a(\mathbf{r}_i) \,\rangle_a\, \exp\left(- |\mathbf{r}_j - \mathbf{r}_i|^2 / r_b^{\,2} \right) \\
&= \frac{H(0)}{\pi\, r_b^{\,2}} \sum_i A_0\, \exp\left(- |\mathbf{r}_j - \mathbf{r}_i|^2 / r_b^{\,2} \right) \\
&= \frac{1}{\epsilon^2} H(0)\, A_0 \quad ,
\end{aligned} \tag{5.15}$$

where ϵ^2 is the pixel area, and $\overline{B}$ has the same units as A_0, that is counts/(sec mm^2).

Task Specification

Our studies are based on two classification tasks: the detection task, in which an

observer must determine whether or not there is a known signal in the center of the image, and the Rayleigh task, where the observer must decide if an image contains a single or a double signal centered in the image. For both tasks, the signal is known exactly and is centered in the image. The uncertainty in the background is described by the autocorrelation function of the background, which is always Gaussian. The two models of lumpiness that were described in the previous section are two examples of lumpy backgrounds satisfying stationary statistics and a Gaussian autocorrelation function.

Detection Task

In the first study, the signal to be detected is a discretized version of the 2D Gaussian function s(**r**) given by equation (5.1). Under the null hypothesis H_1, the signal is absent and the object f(**r**) is an inhomogeneous lumpy background as described in the preceding section. Under hypothesis H_2, the object f(**r**) is the signal s(**r**) superimposed on the lumpy background. The width of the Gaussian blobs r_b is chosen to be three times the width of the signal r_s to allow for discrimination. If r_b is less than 3 times r_s , the background looks as if it is formed of multiple signals, and the discrimination of the signal from the background becomes very difficult even if we know the exact location of the signal in the image. Moreover, if we were to consider a random location of the signal in the image, the task of discriminating the signal from a lump would become even more difficult. Some simulations of lumpy backgrounds of different correlation lengths will be presented in chapter 6 to illustrate this point. We should also keep in mind that the detection of the signal is done at the detection plane and that both the signal and the background have to be imaged through the imaging system. For the problem at hand, we shall consider as an imaging system a pinhole gamma-ray camera forming a single projection of an object onto the detector plane. We shall look for an optimum aperture size to discriminate a signal of

size r_s superimposed on a lumpy background of correlation length r_b. Given the optimum aperture size, we shall then look at the dependence of the performance on the exposure time.

Rayleigh Discrimination Task

The second task involves the discrimination between images with a single, centered, Gaussian signal and images with two Gaussian signals separated by a distance 2d and whose mid-point is centered in the image. Under hypothesis H_1 the signal is a single Gaussian of known amplitude a_s and half width r_s as described in equation (5.1). Under hypothesis H_2 the signal is a double Gaussian, where the amplitude of each Gaussian in the pair is half that of the single Gaussian, the widths are the same as the width of the single, and the separation and orientation are known. The object under either hypothesis is the appropriate signal added to an inhomogeneous background characterized in the same manner as for the detection task. As for the detection task, we shall look for the optimal aperture size that will optimize the discrimination between the two types of signals and at the dependence of the performance on the exposure time.

Performance of Statistical Observers

We shall now investigate the performance of two statistical observers, the npw ideal and the Hotelling observers.

The Nonprewhitening Ideal Observer

The strategy of the npw ideal observer is to apply a simple template matched to the expected difference signal to form a test statistic, which is given, as shown in chapter 2, by

$$\lambda_{npw} = (\mathbf{H}\Delta\mathbf{s})^t \mathbf{g} \quad , \tag{5.16}$$

and the SNR associated with the test statistic, or detectability index, is given in the low-contrast approximation by

$$[d_{npw}]^2 = \frac{[\Delta\overline{\lambda}]^2}{[\sigma_\lambda]^2} \quad , \tag{5.17}$$

where the mean and variance are found by averaging over both the Poisson noise and the object variability due to the nonuniform background. Let us consider the numerator first. The mean of λ when hypothesis k is true is

$$\overline{\lambda_k} = \langle \langle \lambda \rangle_{n|f} \rangle_{f|k} \quad . \tag{5.18}$$

We find the average for the class by performing an average over the Poisson noise for a particular object in the class, followed by an average over all objects from the class. If we use the expression for the test statistic in equation (5.16), we find that

$$\overline{\lambda_k} = \Delta\mathbf{s}^t \, \mathbf{H}^t \, \langle \langle \mathbf{g} \rangle_{n|f} \rangle_{f|k} = \Delta\mathbf{s}^t \, \mathbf{H}^t \, \overline{\mathbf{g}_k} \quad , \tag{5.19}$$

where

$$\overline{\mathbf{g}_k} = \langle \mathbf{g} \rangle_k = \langle \overline{\mathbf{g}_f} \rangle_{f|k} = \langle \langle \mathbf{g} \rangle_{n|f} \rangle_{f|k} \quad . \tag{5.20}$$

If we denote as $\Delta\overline{\mathbf{g}}$ the difference in the means of $\mathbf{g}$ under the two hypotheses, the difference in the means of λ under the two hypotheses is then

$$\Delta\overline{\lambda} = \Delta s^t \, \mathbf{H}^t \, \Delta\overline{\mathbf{g}} \quad . \tag{5.21}$$

Moreover, since the background has equal statistics under each hypothesis, the difference in overall class mean $\Delta\overline{\mathbf{g}}$ is simply the image of the difference signal,

$$\Delta\overline{\mathbf{g}} = \mathrm{H}\,(\overline{\mathbf{f}_2}-\overline{\mathbf{f}_1}) = \mathrm{H}\,\Delta\mathbf{s} \quad , \tag{5.22}$$

so that,

$$\Delta\overline{\lambda} = \Delta s^t \, \mathbf{H}^t \, \mathbf{H} \, \Delta\mathbf{s} \quad . \tag{5.23}$$

We calculate the variance of the test statistic by cascading averages over noise and objects in a similar manner. Under the low-contrast assumption, the definition of the variance of λ is

$$[\sigma_\lambda]^2 = \langle \langle (\lambda - \overline{\lambda})^2 \rangle_{n|f} \rangle_f \quad . \tag{5.24}$$

If we add and subtract the mean of λ, averaged over Poisson noise given a particular object, and denoted $\overline{\lambda_f}$, equation (5.24) becomes

$$[\sigma_\lambda]^2 = \langle \langle (\lambda - \overline{\lambda_f} + \overline{\lambda_f} - \overline{\lambda})^2 \rangle_{n|f} \rangle_f$$

$$= \langle \langle (\lambda - \overline{\lambda_f})^2 \rangle_{n|f} \rangle_f + \langle \langle (\overline{\lambda_f} - \overline{\lambda})^2 \rangle_{n|f} \rangle_f + \text{cross term.} \tag{5.25}$$

A quick inspection will show that the cross-term is equal to zero and so we are left

with two terms to evaluate. We evaluate the first term to find:

$$\langle \langle (\lambda - \overline{\lambda_f})^2 \rangle_{n|f} \rangle_f = \langle \langle (\Delta s^t \, \mathbf{H}^t \, \mathbf{g} - \Delta s^t \, \mathbf{H}^t \, \overline{\mathbf{g}_f})^2 \rangle_{n|f} \rangle_f$$

$$= \langle \langle \Delta s^t \, \mathbf{H}^t (\mathbf{g} - \overline{\mathbf{g}_f}) (\mathbf{g} - \overline{\mathbf{g}_f})^t \, \mathbf{H} \, \Delta s \rangle_{n|f} \rangle_f$$

$$= \Delta s^t \, \mathbf{H}^t \langle \langle (\mathbf{g} - \overline{\mathbf{g}_f})(\mathbf{g} - \overline{\mathbf{g}_f})^t \rangle_{n|f} \rangle_f \, \mathbf{H} \, \Delta s$$

$$= \Delta s^t \, \mathbf{H}^t \, \mathbf{M}_1 \, \mathbf{H} \, \Delta s \quad , \tag{5.26}$$

where the diagonal matrix $\mathbf{M}_1$ characterizes the contribution of the Poisson noise to the variance and is shown in appendix A (equations (A9) and (A10)) to be given by

$$\mathbf{M}_1 = \kappa \, T \, A_{ap} \, \overline{B} \, \mathbf{I} \quad , \tag{5.27}$$

where κ is the efficiency factor and $\mathbf{I}$ is the identity matrix. The second term in equation (5.25) is given by

$$\langle \langle (\overline{\lambda_f} - \overline{\lambda})^2 \rangle_{n|f} \rangle_f = \langle \langle (\Delta s^t \, \mathbf{H}^t \, \overline{\mathbf{g}_f} - \Delta s^t \, \mathbf{H}^t \, \overline{\mathbf{g}})^2 \rangle_{n|f} \rangle_f$$

$$= \Delta s^t \, \mathbf{H}^t \langle \langle (\overline{\mathbf{g}_f} - \overline{\mathbf{g}})(\overline{\mathbf{g}_f} - \overline{\mathbf{g}})^t \rangle_{n|f} \rangle_f \, \mathbf{H} \, \Delta s$$

$$= \Delta s^t \, \mathbf{H}^t \, \mathbf{R}_g \, \mathbf{H} \, \Delta s \quad , \tag{5.28}$$

where $\mathbf{R}_g$ is the autocorrelation function of the nonuniform background in the image plane.

We can now combine the expressions for the mean and variance to write d_{npw} as

$$[d_{npw}]^2 = \frac{(\Delta s^t \mathbf{H}^t \mathbf{H} \Delta s)^2}{\Delta s^t \mathbf{H}^t (\kappa T A_{ap} \overline{B} \mathbf{I} + \mathbf{R}_g) \mathbf{H} \Delta s} \quad , \tag{5.29}$$

where $\mathbf{R}_g$ is proportional to T^2, while $\mathbf{H}$ is proportional to T. This expression may be simplified by performing the calculation in the Fourier domain. We use the discrete Fourier operator $\mathbf{F}$ defined by equation (2.45) to map the difference signal Δs into the Fourier domain as follows,

$$\mathbf{F}\Delta s = \Delta \tilde{s} \quad , \tag{5.30}$$

where $\Delta \tilde{s}(m)$ is the Fourier amplitude of Δs at frequency ρ_m . As discussed in chapter 2, the Fourier operator $\mathbf{F}$ is a unitary operator, so that $\mathbf{F}\mathbf{F}^\dagger=\mathbf{F}^\dagger\mathbf{F}=\mathbf{I}$, where $\mathbf{I}$ is the identity matrix and the dagger represents the conjugate transpose of a matrix or vector. Since the difference signal is real, we find that

$$\Delta \tilde{s}^\dagger = (\mathbf{F}\Delta s)^\dagger = \Delta s^t \mathbf{F}^\dagger \quad . \tag{5.31}$$

The system matrix $\mathbf{H}$ of a shift-invariant imaging system is approximately block circulant, and therefore it is diagonalizable using the discrete Fourier operator.

We can now write an equivalent expression for d_{npw} in the Fourier domain by inserting Fourier operators as shown

$$[d_{npw}]^2 = \frac{[\Delta s^t (F^\dagger F) H^t (F^\dagger F) H (F^\dagger F) \Delta s]^2}{\Delta s^t (F^\dagger F) H^t (F^\dagger F)[\kappa T A_{ap} \overline{B} I + R_g] (F^\dagger F) H (F^\dagger F) \Delta s}$$

$$= \frac{[(\Delta s^t F^\dagger) (F H^t F^\dagger) (F H F^\dagger) (F \Delta s)]^2}{(\Delta s^t F^\dagger) (F H^t F^\dagger) [\kappa T A_{ap} \overline{B} (F I F^\dagger) + F R_g F^\dagger] (F H F^\dagger) (F \Delta s)}$$

$$= \frac{[\Delta \tilde{s}^\dagger \tilde{H}^\dagger \tilde{H} \Delta \tilde{s}]^2}{\Delta \tilde{s}^\dagger \tilde{H}^\dagger [\kappa T A_{ap} \overline{B} \tilde{I} + W_g] \tilde{H} \Delta \tilde{s}} \quad . \tag{5.32}$$

For clarity, we shall now carry out the calculation in 1D and then generalize the result to the 2D problem. In 1D, the Fourier transform of **H** is then given by

$$F H F^\dagger = \tilde{H} \tag{5.33}$$

with

$$\tilde{H}(i, j) = \tilde{H}(i) \, \delta_{ij} \quad . \tag{5.34}$$

We have written the diagonal matrix $\tilde{H}$ in component form as a function of a single index that runs along the matrix diagonal. In the Fourier domain the covariance matrix **K** becomes

$$F K F^\dagger = \tilde{K} \quad , \tag{5.35}$$

which we can write in component form, using equation (A4), as

$$\begin{aligned} \tilde{K}(i, j) &= (W_g(i) + \kappa T A_{ap} \overline{B}) \, \delta_{ij} \\ &= [|\tilde{H}(i)|^2 W_f(i) + \kappa T A_{ap} \overline{B}] \, \delta_{ij} \quad . \end{aligned} \tag{5.36}$$

Note that $\mathbf{W}_g$ is the power spectrum of the nonuniform background in the image plane, while $\mathbf{W}_f$ is the power spectrum of the nonuniform background in the object plane. If we write out the summations represented implicitly in the matrix notation, we find that

$$[d_{npw}]^2 = \frac{\left[\sum_i |\Delta\tilde{s}(i)|^2 \, |\tilde{H}(i)|^2\right]^2}{\sum_i |\Delta\tilde{s}(i)|^2 \, |\tilde{H}(i)|^2 \, [\kappa \, T \, A_{ap} \, \overline{B} + W_g(i)]}$$

$$= \frac{\left[\sum_i |\Delta\tilde{s}(i)|^2 \, |\tilde{H}(i)|^2\right]^2}{\sum_i |\Delta\tilde{s}(i)|^2 \, |\tilde{H}(i)|^2 \, [\kappa \, T \, A_{ap} \, \overline{B} + |\tilde{H}(i)|^2 \, W_f(i)]} \quad . \tag{5.37}$$

$$= \frac{I_1^2}{\kappa \, T \, A_{ap} \, \overline{B} \, I_1 + I_2} \quad , \tag{5.38}$$

where

$$I_1 = \sum_i |\Delta\tilde{s}(i)|^2 \, |\tilde{H}(i)|^2 \tag{5.39}$$

and

$$I_2 = \sum_i |\Delta\tilde{s}(i)|^2 \, |\tilde{H}(i)|^4 \, W_f(i) \quad . \tag{5.40}$$

This result can now be generalized to the 2D problem. Moreover, in the limit where the area of the sampling element approaches zero, the discrete sums approximate

integrals and equations (5.39) and (5.40) become

$$I_1 = \int_{\infty} d^2\rho \; |\Delta\tilde{s}(\rho)|^2 \; |\tilde{H}(\rho)|^2 \tag{5.41}$$

and

$$I_2 = \int_{\infty} d^2\rho \; |\Delta\tilde{s}(\rho)|^2 \; |\tilde{H}(\rho)|^4 \; W_f(\rho) \quad , \tag{5.42}$$

where ρ is the 2D spatial-frequency vector. This Fourier representation is valid only in the low-contrast limit where stationary statistics apply.

The Hotelling Observer

To determine a figure of merit for the Hotelling observer, we must first determine the first- and second-order statistics on the data that lead to the scatter matrices $\mathbf{S}_1$ and $\mathbf{S}_2$. If the two classes are equally likely, the interclass scatter matrix (see equation (2.51)) is given by

$$\begin{aligned} \mathbf{S}_1 &= \frac{1}{4} (\bar{\mathbf{g}}_2 - \bar{\mathbf{g}}_1)(\bar{\mathbf{g}}_2 - \bar{\mathbf{g}}_1)^t \\ &= \frac{1}{4} (\Delta\bar{\mathbf{g}})(\Delta\bar{\mathbf{g}})^t \\ &= \frac{1}{4} (\mathbf{H}\Delta\mathbf{s})(\mathbf{H}\Delta\mathbf{s})^t \quad , \end{aligned} \tag{5.43}$$

where we used equation (5.22) to derive line 3.

The covariance matrix of the k^{th} class is defined from equation (2.52) as

$$\mathbf{K}_k = \langle \langle (\mathbf{g}-\overline{\mathbf{g}_k})(\mathbf{g}-\overline{\mathbf{g}_k})^t \rangle_{n|f} \rangle_{f|k} \quad . \tag{5.44}$$

Since the object inhomogeneity is the same under both classes, the covariance matrices are approximately equal in the low-contrast limit ($\mathbf{K}_1 \simeq \mathbf{K}_2 \simeq \mathbf{K}$). Thus, for our low-contrast SKE problem, we see from equation (2.52) that $\mathbf{S}_2$ is simply equal to $\mathbf{K}$. Appendix A gives the derivation of the covariance matrix $\mathbf{K}$ in the general case where the data present both quantum noise and structured noise. We find the value in the i^{th} row and j^{th} column to be

$$K(i,j) = R_g(i,j) + \kappa \, T \, A_{ap} \, \overline{B} \, \delta_{ij} \quad , \tag{5.45}$$

where $\mathbf{R}_g$ is the autocorrelation of the data due to the nonuniform background and A_{ap} is the area of the aperture. We see that the covariance matrix is a sum of two matrices, one containing the randomness of the object due to the lumpy background, $\mathbf{R}_g$, and the other a diagonal matrix containing the variance associated with the Poisson nature of the detected data.

If we substitute $\mathbf{K}$ for $\mathbf{S}_2$ in the expression for the Hotelling observer test statistic given by equation (2.64), we find that

$$\begin{aligned} \lambda_{Hot} &= [\langle \mathbf{g} \rangle_2 - \langle \mathbf{g} \rangle_1]^t \, \mathbf{K}^{-1} \, \mathbf{g} \\ &= \Delta \mathbf{g}^t \, \mathbf{K}^{-1} \, \mathbf{g} \;=\; (\mathbf{H}\Delta \mathbf{s})^t \, \mathbf{K}^{-1} \, \mathbf{g} \quad . \end{aligned} \tag{5.46}$$

The detectability index d_{Hot} is related to the Hotelling trace J by equation (2.69) as

$$\begin{aligned}[d_{Hot}]^2 &= 4\,J = 4\,\mathrm{tr}[\mathbf{S}_2^{-1}\mathbf{S}_1] \\ &= \mathrm{tr}[\mathbf{K}^{-1}(\mathbf{H}\Delta\mathbf{s})(\mathbf{H}\Delta\mathbf{s})^t] \\ &= \Delta\mathbf{s}^t\mathbf{H}^t\mathbf{K}^{-1}\mathbf{H}\Delta\mathbf{s} \quad . \end{aligned} \tag{5.47}$$

The full expression for the Hotelling observer's detectability for linear, shift-invariant systems and stationary covariance matrices can also be given as an integral over the Fourier domain by inserting the Fourier operator as shown

$$\begin{aligned}[d_{Hot}]^2 &= \Delta\mathbf{s}^t\mathbf{H}^t\mathbf{K}^{-1}\mathbf{H}\Delta\mathbf{s} \\ &= \Delta\mathbf{s}^t\,(\mathbf{F}^\dagger\mathbf{F})\,\mathbf{H}^t\,(\mathbf{F}^\dagger\mathbf{F})\,\mathbf{K}^{-1}\,(\mathbf{F}^\dagger\mathbf{F})\,\mathbf{H}\,(\mathbf{F}^\dagger\mathbf{F})\,\Delta\mathbf{s} \\ &= (\Delta\mathbf{s}^t\mathbf{F}^\dagger)\,(\mathbf{F}\mathbf{H}^t\mathbf{F}^\dagger)\,(\mathbf{F}\mathbf{K}^{-1}\mathbf{F}^\dagger)\,(\mathbf{F}\mathbf{H}\mathbf{F}^\dagger)\,(\mathbf{F}\Delta\mathbf{s}) \\ &= \Delta\tilde{\mathbf{s}}^\dagger\tilde{\mathbf{H}}^\dagger\tilde{\mathbf{K}}^{-1}\tilde{\mathbf{H}}\Delta\tilde{\mathbf{s}} \quad . \end{aligned} \tag{5.48}$$

If we again consider the 1D problem, equation (5.48) becomes in component form,

$$\begin{aligned}[d_{Hot}]^2 &= \sum_i\sum_j\sum_k\sum_\ell \Delta\tilde{s}^*(i)\,\tilde{H}^*(i,j)\,\tilde{K}^{-1}(j,k)\,\tilde{H}(k,\ell)\Delta\tilde{s}(\ell) \\ &= \sum_i \Delta\tilde{s}^*(i)\,\tilde{H}^*(i)\tilde{K}^{-1}(i)\,\tilde{H}(i)\Delta\tilde{s}(i) \\ &= \sum_i \frac{|\Delta\tilde{s}(i)|^2\,|\tilde{H}(i)|^2}{[\kappa\,T\,A_{ap}\;\overline{B} + |\tilde{H}(i)|^2\,W_f(i)]} \quad , \end{aligned} \tag{5.49}$$

where the raised asterisk denotes the complex conjugate and we have made use of the delta

functions in the diagonal matrices to reduce the multiple sum to a sum over a single index.

We can generalize this result to the 2D problem. Moreover, by taking the limit of the sum for a large number of fine samples, we can write d_{Hot} as a continuous integral in the Fourier domain as

$$[d_{Hot}]^2 = \int_{\infty} d^2(\rho) \frac{|\widetilde{\Delta s}(\rho)|^2 \, |\tilde{H}(\rho)|^2}{[\kappa \, T \, A_{ap} \, \overline{B} + |\tilde{H}(\rho)|^2 \, W_f(\rho)]} \quad . \tag{5.50}$$

where a tilde above a function signifies its Fourier transform. Thus $\tilde{H}(\rho)$ is the system transfer function, and $\widetilde{\Delta s}(\rho)$ is the Fourier spectrum of the difference signal $\Delta s(\mathbf{r})$.

The integrand in Eq. (5.50) may be regarded as a generalization of the frequency-dependent noise-equivalent quanta (NEQ) for the lumpy-background discrimination problem. The NEQ concept, introduced by Shaw and reviewed in Dainty and Shaw (1974) and Shaw (1978), has also been extensively used, among others, by Wagner (1983) and Wagner and Brown (1985) for SKE discrimination tasks. The generalized NEQ, first introduced by Barrett et al. (1989), is given by

$$NEQ(\rho) = \frac{|\tilde{H}(\rho)|^2}{[\kappa \, T \, A_{ap} \, \overline{B} + |\tilde{H}(\rho)|^2 \, W_f(\rho)]} \quad . \tag{5.51}$$

The usual expression for $NEQ(\rho)$ for the flat-background problem is recovered from this form by setting $W_f(\rho)$ to zero and recognizing that $\kappa \, T \, A_{ap} \, \overline{B}$ is the Poisson noise power spectral density for that case.

In order to calculate the detectabilities of the npw ideal and the Hotelling observers given by equations (5.38), (5.41), (5.42), and (5.50), respectively, we must assign numerical

values to the signal and the imaging system parameters. The Gaussian signals had a width r_s of 14.1 mm and a contrast relative to the mean background level $\overline{B}$ of 5%. This contrast value is a realistic, albeit challenging, value for nuclear medicine applications. To achieve this contrast figure, the background level $\overline{B}$ was set equal to 610 counts/(sec mm^2), which then sets the signal parameter a_s to 1.9×10^4 counts/sec. The correlation length of the lumpy background was set equal to 42.4 mm, so that $r_b = 3\ r_s$. The background level $\overline{B}$ was chosen to give 10^7 disintegrations in the given imaging time and in a 128^2 mm^2 object, so that one could detect 10^9 counts with a perfect photon collection system. For all studies the exposure time T was set to 1 sec unless otherwise specified and κ was a constant set equal to 10^{-3} mm^{-2}.

Example 5.1 Detection Task for Gaussian and Rectangular Apertures

The performance of the npw ideal and the Hotelling observers can be evaluated by substituting the appropriate expressions for $\Delta\tilde{s}(\rho)$, $\tilde{H}(\rho)$, $W_f(\rho)$ and A_{ap} into equations (5.38), (5.41), (5.42) and (5.50), respectively. For the simple two-hypotheses detection task, the expression for $\Delta\tilde{s}(\rho)$ is simply the expression for $\tilde{s}(\rho)$, which can be derived from equation (5.1), to yield

$$\tilde{s}(\rho) = a_s \ \exp(- \pi^2 r_s^2 |\rho|^2) \quad . \tag{5.52}$$

The expression for $W_f(\rho)$ is given by equation (5.9), while the expression for $H(\rho)$ takes the form

$$\tilde{H}(\rho) = \kappa T \pi r_p^2 \exp(- \pi^2 r_p^2 |\rho|^2) \quad , \tag{5.53}$$

in the case of a Gaussian aperture. For a rectangular aperture,

$$H(\rho) = \kappa\, T\, L^2 \, \frac{\sin(\pi L \xi_x)}{\pi L \xi_x} \, \frac{\sin(\pi L \xi_y)}{\pi L \xi_y} \quad , \tag{5.54}$$

where ξ_x and ξ_y denote the components of the vector ρ such that $\rho^2=\xi_x^2+\xi_y^2$. Note that the parameter A_{ap} was replaced either by πr_p^2 for a Gaussian aperture or L^2 for the rectangular aperture. A numerical integration of equations (5.38) and (5.50) yields the expression for the npw ideal and the Hotelling observers, respectively.

Figure 5.1 (a) is a plot of d_{Hot}^2 for the low-contrast detection task as a function of Gaussian aperture radius r_p for increasing levels of the background parameter $W_f(0)$. Figure 5.1 (b) is a similar plot for the npw ideal observer. The upper curve in both figures corresponds to a uniform background, in which both observers reduce to the ideal observer. Note that, in the uniform-background case where the observer performance is purely quantum-limited, the aperture should be as large as possible to collect as many photons as possible. This result is consistent with the finding of Tsui et al. (1978).

In the presence of a lumpy background the performances of both observers demonstrate a clear optimum size. We have normalized the abscissa by the signal size r_s in all figures to clarify the relationship between the optimum aperture size and the size of the signal. For all values of $W_f(0)$, we find that the optimum aperture size is generally the same size as the signal. The npw ideal observer is seen to experience a greater performance penalty than the Hotelling observer as the aperture size is increased beyond the optimum value.

Figure 5.2 repeats the calculations of Figure 5.1, when the Gaussian aperture is replaced by a square aperture of area L^2. Little difference between the Gaussian aperture and square aperture for the npw ideal observer can be observed. However, there is a marked difference in the performance of the Hotelling observer for the square aperture, for

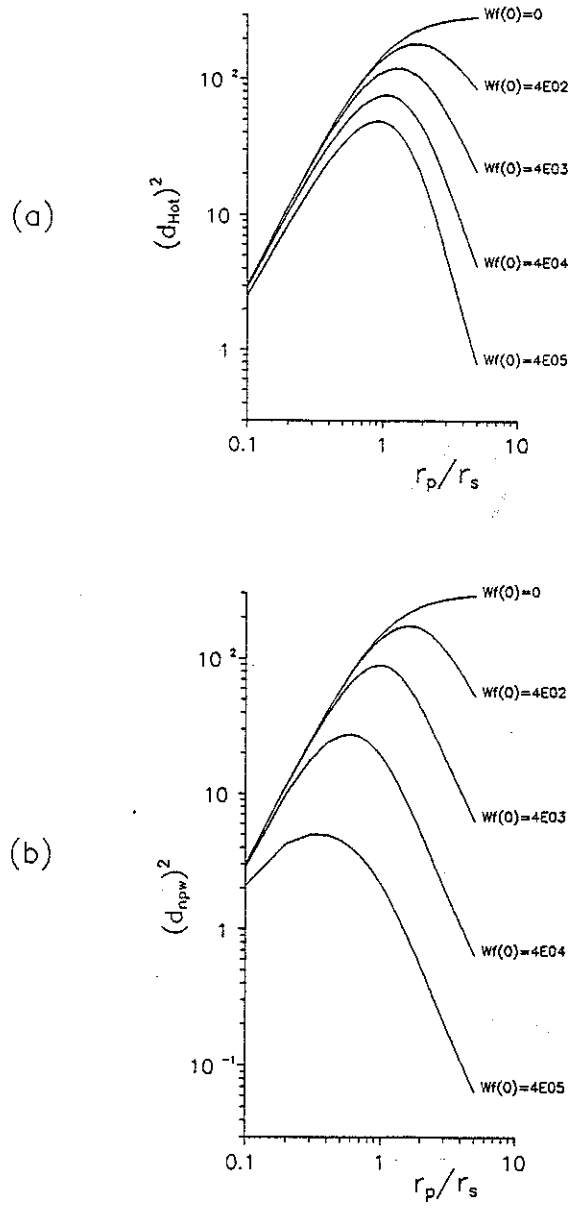


Fig. 5.1. (a) Family of Hotelling $[d_{Hot}]^2$ curves as a function of the ratio of the Gaussian aperture width to the signal width for a detection task. The plots show the effect of increasing the value of the power spectral density at all frequencies by increasing the value of $W_f(0)$ expressed in counts^2/(sec^2 mm^2). Signal width is 10 mm, the width of the autocorrelation function of the background is 30 mm, and the exposure time is 1 sec. The signal contrast is 5%. (b) Plot of the npw ideal observer $[d_{npw}]^2$ as a function of the Gaussian aperture width to the signal width for the same detection task as in (a).

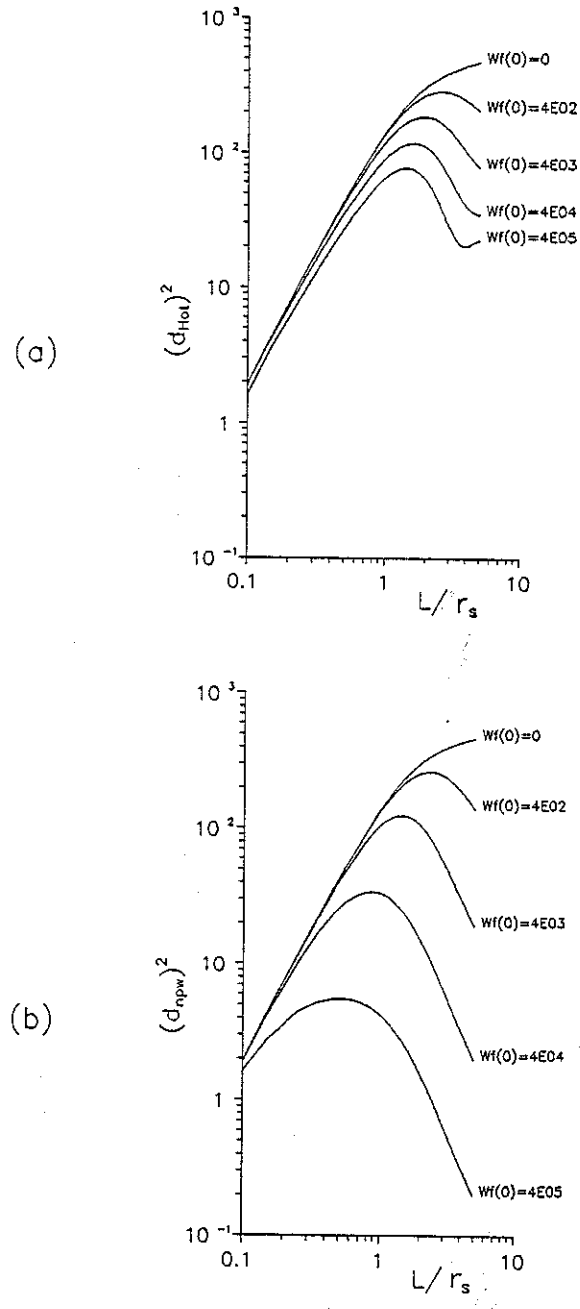


Fig. 5.2. (a) Plot of the Hotelling $[d_{Hot}]^2$ for the detection task with a square aperture of length L on a side. The abscissa is the ratio of the squar aperture length to the signal width. Signal width is 10 mm, the width of the autocorrelation function of the background is 30 mm, and the exposure time is 1 sec. The signal contrast is 5%. (b) Plot of the npw ideal observer $[d_{npw}]^2$ as a function of the ratio of the side length of the square aperture to the signal width. Signal and background parameters are the same as for (a).

we now find oscillations in the d_{Hot} graph. Some intuition regarding the presence of these oscillations can be ascertained from a frequency-domain point of view, where the Fourier transform of the square aperture becomes an oscillatory function in the integral of equation (5.50). As the square size L increases, its sinc function in the Fourier picture narrows and oscillatory lobes in the transfer function move toward zero frequency. Certain ranges of L beyond the optimum size evidently result in particular positions of the maxima and minima of the sinc that result in local improvements in the discriminability between signal and background, resulting in small oscillations in d_{Hot}.

Example 5.2 Rayleigh Task for Gaussian and Square Apertures

The only change in computation from the detection to the Rayleigh task resides in substituting a different expression for $\Delta\tilde{s}(\rho)$ in the expression for d_{npw} and d_{Hot}. A simple calculation of the substraction of the Fourier transform of a double Gaussian signal from the Fourier transform of a single Gaussian signal leads to an expression for $\Delta\tilde{s}(\rho)$ given by

$$\Delta\tilde{s}(\rho) = a_s\ (1-\cos(2\pi d\xi))\ \exp(\ -\ \pi^2\ r_s{}^2\ |\rho|^2\ )\quad , \tag{5.55}$$

where 2d is the seperation of the two Gaussian signals forming the double Gaussian signal.

Figures 5.3 exhibits representative results for the Hotelling observer for the Rayleigh discrimination task for d = 10 mm. In this case the optimum aperture size is comparable to the size of the signal even without background inhomogeneity. The figures show that only a very small performance degradation occurs for the Hotelling observer when a lumpy background is introduced for either Gaussian or square apertures. The differencing operation performed by the Hotelling observer is quite effective at distinguishing background from signal for this task, partly because the double signal does not effectively

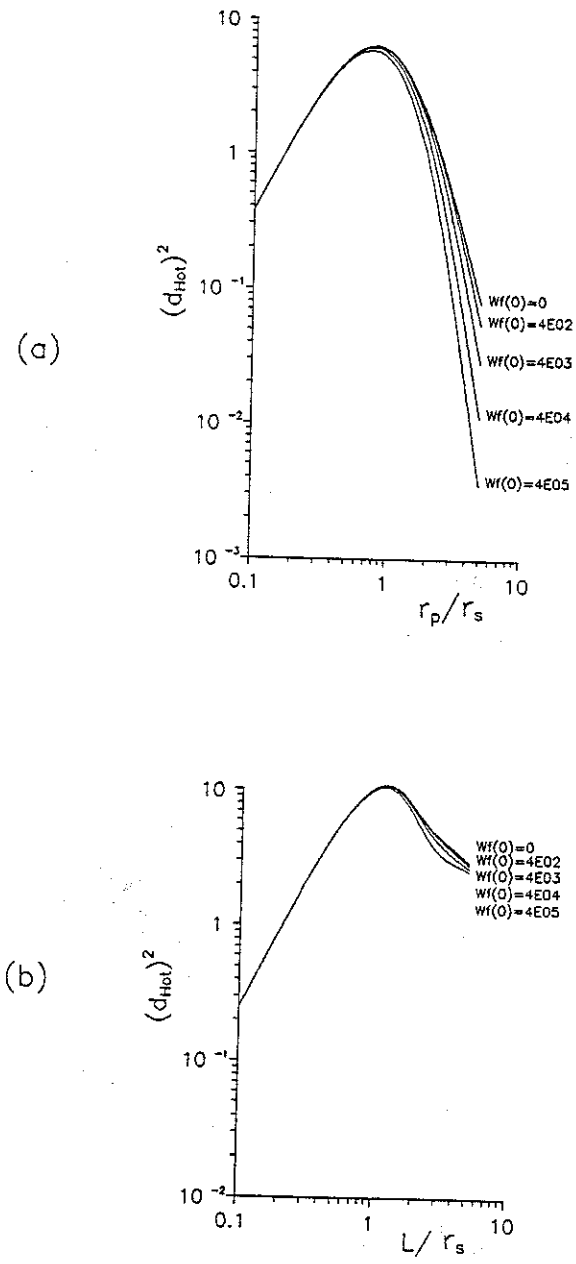


Fig. 5.3. (a) Plot of the Hotelling $[d_{Hot}]^2$ for the Rayleigh discrimination task for Gaussian apertures of varying size r_p. Source width is 10 mm, the pair separation is 20 mm, and the exposure time is 1 sec. Same contrast, background parameters as in Fig. 5.2. (b) Plot of the Hotelling $[d_{Hot}]^2$ for the Rayleigh discrimination task for square apertures of varying length L on a side. Source width is 10 mm, the pair separation is 20 mm, and the exposure time is 1 sec. Same contrast, background parameters as in Fig. 5.2.

mimic the symmetry of the background "lumps." We see some evidence of oscillations in the curves for the square aperture and the Hotelling observer, again due to the effect of an oscillatory function in the Fourier-domain integral.

Finally, Figure 5.4 shows the Rayleigh discrimination ability of the npw ideal observer for (a) Gaussian and (b) square apertures, respectively. We see that the large apertures are much more strongly affected by the presence of the lumpy background than the smaller ones regardless of aperture shape.

Example 5.3 Variation of Performance with Exposure Time

The two previous examples showed the result of aperture size on the performance of the npw ideal and the Hotelling observers. We shall now look at the effect of longer exposure time on the performance for a fixed resolution. In the case of a Gaussian aperture, r_p is set to r_s, while L is set to 2 r_s in the case of a rectangular aperture.

A striking difference between the Hotelling and the npw ideal observers is found by plotting the detection performance of each observer as a function of exposure time T. These plots are given in Figure 5.5 for increasing values of the background spectral density level $W_f(0)$. When $W_f(0)$ is zero, the background is uniform and both observers reduce to the ideal observer, whose performance increases linearly with exposure time. Figure 5.5 (a) shows a power-law increase in d_{Hot} as a function of T for all non-zero values of $W_f(0)$, with the slope approximately equal to 2/3. No evidence of saturation of the Hotelling-observer performance can be observed in the plot. Figure 5.5 (b) demonstrates a hard saturation in the performance of the npw ideal observer for the same set of tasks. The behavior of the npw ideal observer is predicted from the expression for the d_{npw} given by equations (5.38), (5.41), and (5.42). Since each system operator (and its Fourier transform)

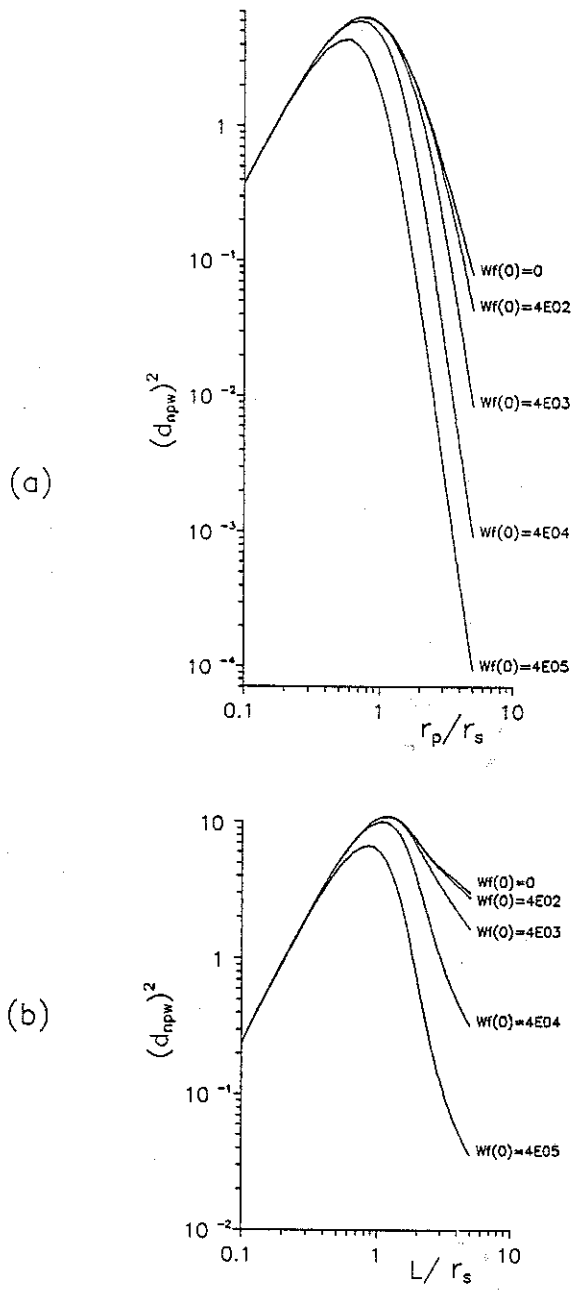


Fig. 5.4. (a) Plot of $[d_{npw}]^2$ for the npw ideal observer for the Rayleigh resolution task with a Gaussian aperture as described in Fig. 5.3. (b) Plot of $[d_{npw}]^2$ for the npw ideal observer for the Rayleigh resolution task with a square aperture as described in Fig. 5.3.

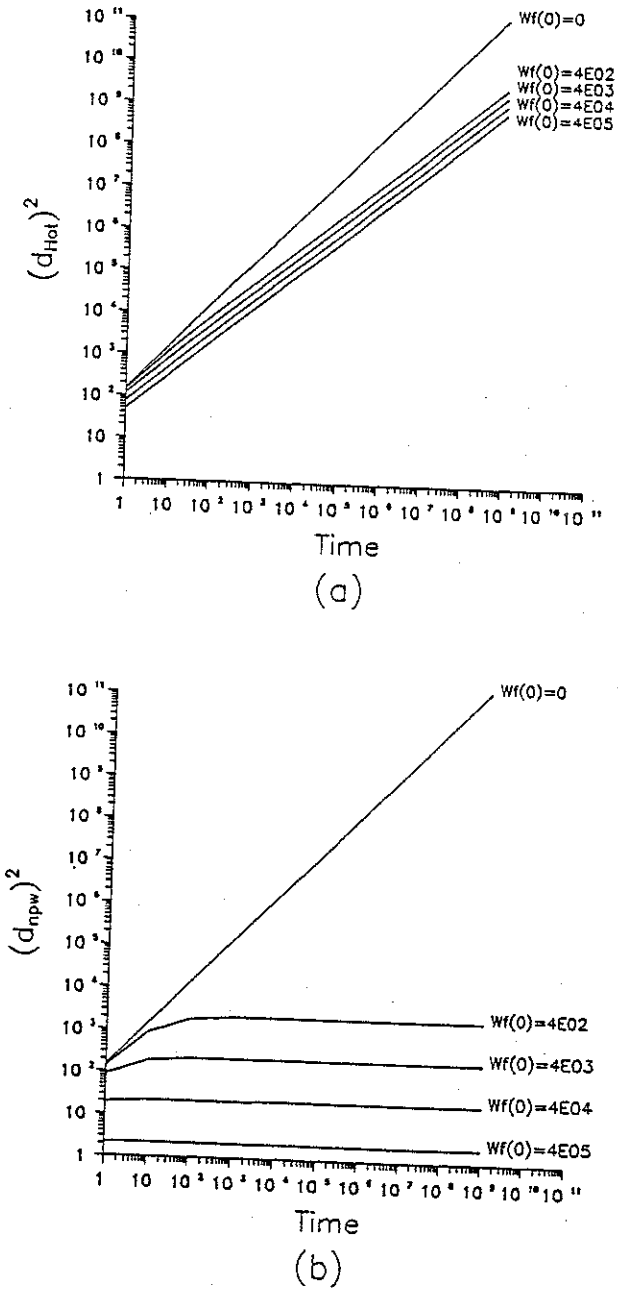


Fig. 5.5. (a) Hotelling $[d_{Hot}]^2$ as a function of exposure time expressed in sec for a detection task. The signal width r_s is 10 mm and the background autocorrelation length r_b is 30 mm. The Gaussian aperture width r_p is equal to the signal width and the signal contrast is 5%. (b) Npw ideal observer $[d_{npw}]^2$ vs. exposure time for the same detection task as in (a).

carries along the time parameter T, the numerator in the $[d_{npw}]^2$ and the background variance term are both proportional T^4, while the Poisson variance term is proportional to T^3. This leads to

$$[d_{npw}]^2 \propto \frac{\alpha T^4}{\beta T^3 + \gamma T^4} \quad . \tag{5.56}$$

As T becomes large, $[d_{npw}]^2$ goes to a constant for any aperture size. Since the magnitude of the γT^4 term increases rapidly with increasing aperture size since it goes as $A_{ap}{}^4$, the conspicuity limit sets in earlier for larger apertures.

It is difficult to get a similar expression for the Hotelling detectability as a function of exposure time, because the numerator and denominator are interacting within the same integral, but we shall discuss in the next session how to look at the Hotelling behavior.

Discussion

In all cases we have considered, plots of d^2 versus aperture size have initial slope equal to 2. This is the regime of negligible blur, so that the counts collected from the unblurred signal region just go up with the aperture area. When the blur due to the aperture becomes appreciable, the convolution operation of the imaging equation means that each elemental area of the aperture "throws" a representation of the signal onto the detector plane. Without lumpiness ($W_f(0)=0$), the Hotelling and npw ideal observers both reduce to the ideal observer. The ideal observer is a matched filter that knows where to look for the signal counts that are spread out by the aperture and gather them all up again. Thus, with

sizable blur, the counts from the signal region on the detector plane collected by the observer still grow with aperture area A_{ap}. In the uniform background case, the area of the background that contributes at any one detector element continues to grow as A_{ap} as the aperture size increases, because we have carried out the calculation as if the background is of infinite spatial extent. Since the area over which the ideal observer has to look for the signal increases by A_{ap}, the background that contributes to the noise in the ideal-observer test statistic grows as $A_{ap}{}^2$, and the standard deviation increases as A_{ap}. The performance of the ideal observer therefore saturates. While the optimum aperture size in the flat background case is infinite, the ideal observer $[d_{Hot}]^2$ or $[d_{npw}]^2$ asymptotically approaches $C^2 \kappa T \overline{B} (A_s^2/2)$, where C is the contrast ($C=s(0)/\overline{B}$) and A_s is the area of the signal (e.g. Gaussian signal area = πr_s^2).

If a spatially varying component is then added to the flat background, the slope at large aperture dimensions for the detection problem decreases further for both the Hotelling and npw ideal observers. Larger apertures collect more of the lumpy background into each detector element, and this extra variation effectively increases the noise in the observer's test statistic. We have seen, though, that the Hotelling observer can cope with the lumpiness to a much greater extent than the npw ideal observer. This can be better understood by looking at the space- and Fourier-domain pictures of the Hotelling feature operator. Figure 5.6 (a) is a one-dimensional schematic of the Hotelling feature operator $\mathbf{a}_{Hot}=\mathbf{K}^{-1}\mathbf{H}\Delta\mathbf{s}$ in the space domain for the lumpy-background detection problem with a Gaussian aperture (the two-dimensional function is rotationally symmetric for this aperture), and Figure 5.6 (b) gives the Fourier-domain filter function. We see from Figure 5.6 (a) that the Hotelling observer's strategy in the detection problem in effect is to use the area surrounding the known signal location to get an estimate of the local background level. The observer subtracts this background estimate from a measure of the counts in the signal region to determine whether

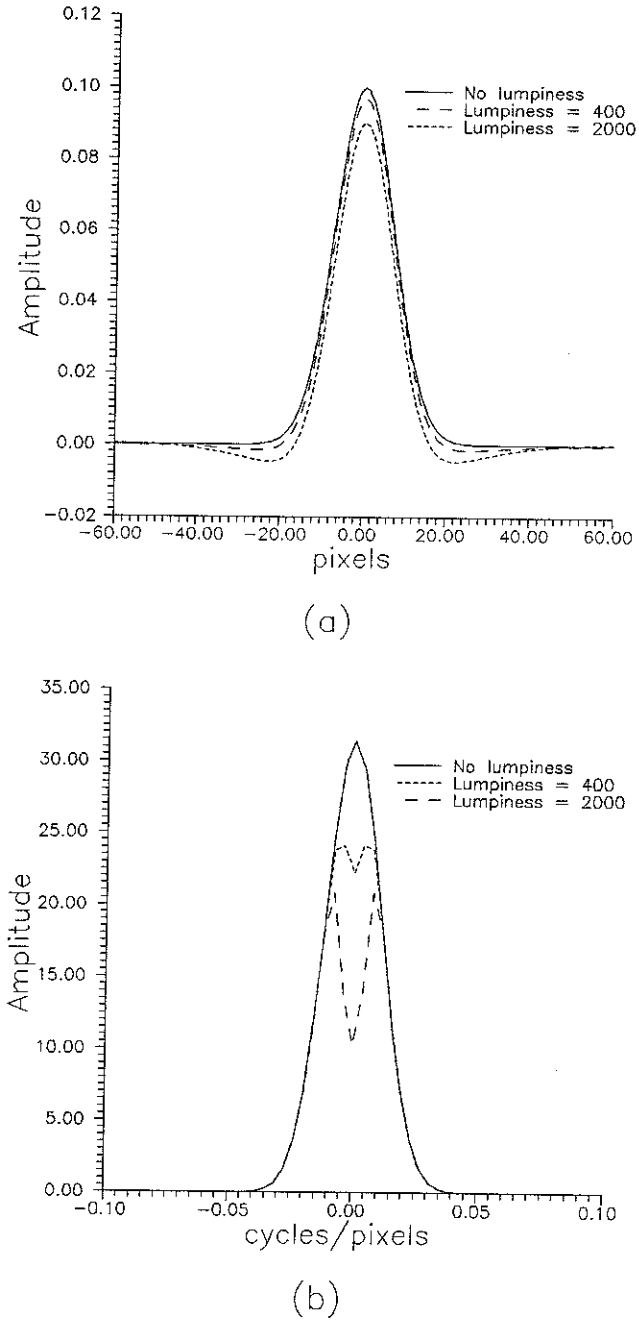


Fig. 5.6. (a) Plot of the Hotelling feature operator along a radial axis in space domain for the detection of a Gaussian signal on a uniform and two nonuniform backgrounds. "Lumpiness" is equivalent to $W_f(0)$. The mean value of the background was 610 counts/(sec mm^2). The signal width was 10 mm and the background autocorrelation length 30 mm. (b) Profile of the Fourier tranform of the feature operator shown in (a).

the count density in the signal location is high enough to be called "signal present." This differencing operation in the space domain is equivalent to a high-pass filter in the frequency domain, as seen in Figure 5.6 (b). As long as the background autocorrelation function has a length different from the signal size, the Hotelling observer can "look" in the Fourier domain for the signal in the frequency channels where the signal and background differ significantly. For these studies the signal width r_s is 1/3 the width of the background correlation length r_b. Thus in the Fourier domain the signal spectrum is wider than the spectrum of the background. The Hotelling observer makes use of this information by making the discrimination based on measuring the amount of power in the higher frequencies. We would expect this correction scheme to falter as the background correlation length approaches the size of the signal.

The approach of looking at the feature operator can also be used to understand the difference in behavior between the npw ideal observer and the Hotelling ideal observer as a function of exposure time. As the exposure time increases, the quality of information in the high frequencies of the data improves, and the Hotelling performance never saturates. On the other hand, the npw ideal observer uses a simple template matched to the difference signal. Even though increasing the exposure time reduces the quantum fluctuations in the data, the observer is limited by an inability to compensate for the varying background. Thus, the npw ideal observer becomes "conspicuity limited" in the limit of large exposure times

The template used by the ideal observer when the background is uniform is simply the image of the difference signal $\mathbf{H}\Delta\mathbf{s}$. In the detection task, the template is therefore just the image of the Gaussian signal itself. In the Rayleigh task, the feature operator for the uniform-background case is the image of the double Gaussian subtracted from the image of the single Gaussian and is shown in Figure 5.7. This feature operator looks much like the

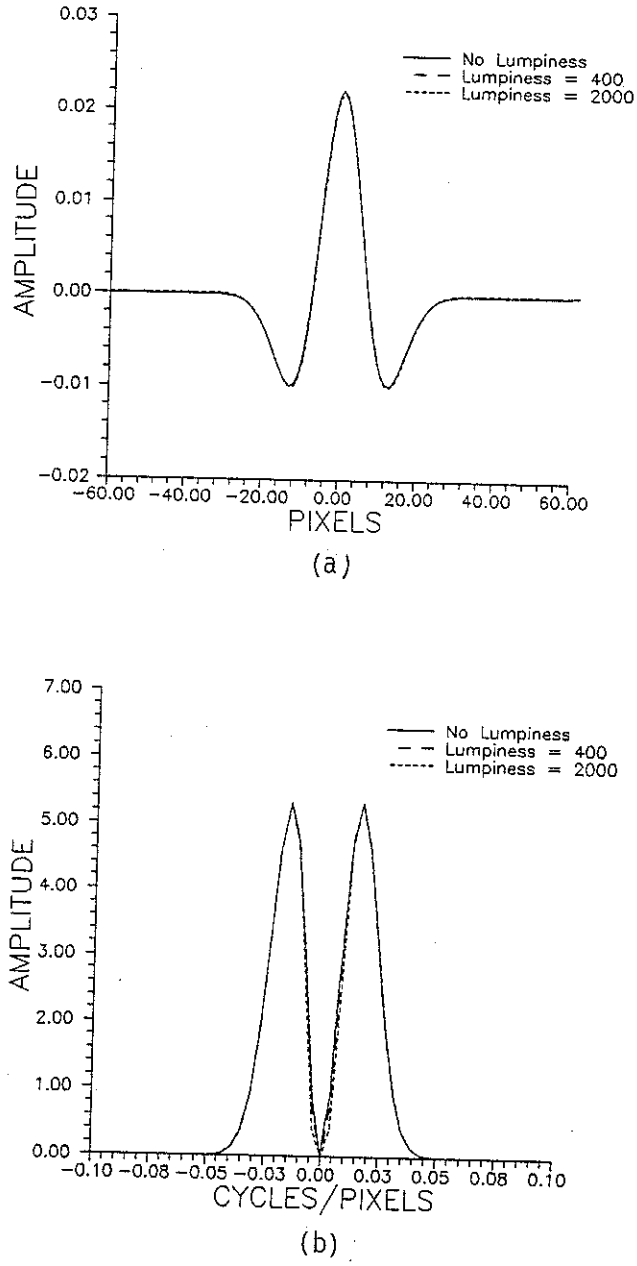


Fig. 5.7. (a) Plot of the Hotelling feature operator along a radial axis in space domain for the Rayleigh discrimination task for Gaussian aperture. The mean value of the background was 610 counts/(sec mm²). The signal width is 10 mm, the separation is 20 mm, and the background autocorrelation length is 30 mm. The signal contrast is 5%. (b) Profile of the Fourier tranform of the feature operator shown in (a).

one shown in Figure 5.6 for the Hotelling observer for detection in the presence of a lumpy background. The Hotelling observer uses a high-frequency filter to perform both tasks. The filters for the two problems differ slightly in that the flat-background Rayleigh task yields a filter that is zero at ρ=0 (because the mean number of counts in the image gives no discrimination) while the lumpy-background detection task has a small but nonzero frequency-domain filter value at zero spatial frequency to attempt to discriminate on the basis of the signal energy.

In the Rayleigh task at low contrast, the optimum aperture size is comparable to the size of the signal even without background inhomogeneity. The fact that we find an optimal aperture for the discrimination task even for uniform backgrounds agrees with the findings of Tsui et al. (1982). When the background is spatially varying the npw ideal observer pays an even larger performance penalty for aperture sizes beyond the optimum, while the Hotelling observer is affected only slightly. Myers et al. (1989), have reported that at high contrasts the performance of the ideal observer in the flat-background Rayleigh task apparently does not saturate with increasing aperture size. More investigations into the transition from the low- to high-contrast regime are needed to explain this further.

For low-contrast tasks, interpretations in the frequency domain are valid and can give good insight, as we have seen particularly in understanding the strategy of the Hotelling observer. Caution must be exercised in attempting to extend this insight to the high-contrast limit, where the frequency-domain picture is not appropriate. At high contrasts a space-domain calculation must be employed due to the non-stationary nature of the covariance matrices and the correlations between Fourier components that result.

A remaining question is the relationship between these results and the discrimination ability of the human observer for the same tasks. Tasks that lead to strikingly different results for the Hotelling and the npw ideal observer, such as the

exposure-time study, allow us to test to see which observer model is a better predictor of human performance for that task. Results of psychophysical studies of human performance on the same tasks as those investigated in this chapter are given in chapter 6 of this work.

This work also has implications for machine readers, that is, computer algorithms designed to do classification tasks. In cases where there is a great difference in performance between the two model observers, simple npw templates should be avoided. Instead, automated algorithms should be designed to incorporate statistics of the background whenever possible. The Hotelling framework provides a straightforward method for calculating the best linear template for performing discrimination tasks in a varying background.

CHAPTER 6

EXPERIMENTAL METHODS AND RESULTS FOR LUMPY BACKGROUNDS

As strongly stressed in this work, the assessment of image quality in medical imaging using the predictions of mathematical observers requires the validation of human observer studies. In this chapter we shall investigate the role of nonuniform or lumpy backgrounds, as compared to uniform backgrounds, in human detection. More specifically we shall study the effect of aperture size and exposure time on human performance. This study will bring some insights into the trade-off between resolution and sensitivity that comes in play in any imaging procedure.

We shall first show some examples of lumpy backgrounds of type I and type II as mathematically described in chapter 5. Since these two types of lumpiness are described by the same Gaussian autocorrelation function, the performance predicted by the Hotelling and the npw ideal observers for the tasks described in chapter 5 (Detection vs. Rayleigh for both rectangular and Gaussian aperture), are the same for these two types of lumpiness. To investigate if the proposed mathematical models can have any validity for the assessment of image quality in medical imaging, we carried out two psychophysical studies based on lumpy backgounds of type I. The first study shows the effect of aperture size on human detection when uniform and nonuniform backgrounds are present, while the second study investigates the dependence of human performance on exposure time in a detection experiment.

Examples of Computer-Simulated Lumpy Backgrounds

We shall first illustrate the two types of lumpy backgrounds resulting from the mathematical descriptions of lumpiness that were given in chapter 5. Figure 6.1 shows some

simulations of these two types of lumpy backgrounds as a function of increasing lumpiness $W_f(0)$ given a mean background $\overline{B}$, as we go from top to bottom, and as a function of increasing mean background given a value of the lumpiness, as we go from left to right. Typical histograms of lumpy backgrounds before noise is introduced, are shown in Figure 6.2.

We then illustrate the effect of imaging the object of interest through a pinhole aperture of size r_p. The object can be described as the superposition of a Gaussian signal of half width r_s at 1/e of the peak intensity and a lumpy background of correlation length r_b. Given the signal size, which we set equal to 11.3 pixels in diameter, and an aperture size that matches in size the signal to be detected, $r_p = r_s = 5.66$ pixels, Figure 6.3 shows the impact of the background correlation length on the appearance of the image. We show with this image that, if the location of the signal was unknown, it would be difficult to discriminate the signal from the background lumps when r_b was less or equal to twice the size of the signal. We set r_b equal to three times the size of the signal for all of the computer-simulated images, so that the correlation in the background does not become the limiting factor in the task to perform.

Other parameters such as the size of the aperture and the exposure time can then be investigated as well. Given the signal size and the background correlation length, the impact of aperture size on detection is illustrated in Figure 6.4, where the size of the aperture, relative to the signal size, increases from left to right as the amount of lumpiness increases from top to bottom. Finally, we are highly interested in the study of the impact of the exposure time on human performance in a simple detection task, and this is illustrated in Figure 6.5. The exposure time increases from left to right, while the amount of lumpiness again increases from top to bottom.

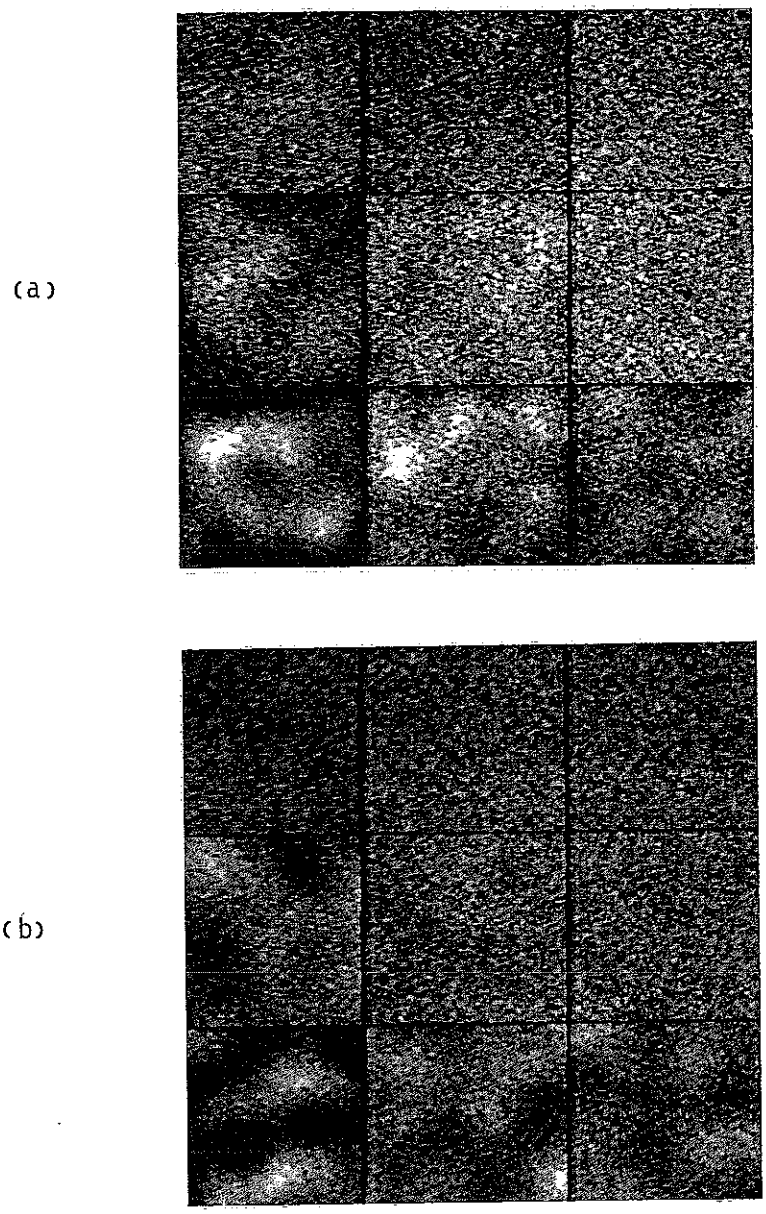


Fig. 6.1. (a) Lumpy backgrounds of type I: the mean background value $\overline{B}$ increases from left to right with $\overline{B}$ equal 384,1664, 3200 counts/(sec pixels) and the lumpiness $W_f(0)$ increases from top to bottom with $W_f(0)$ equal 10^4, 10^5, and 10^6 counts^2/(sec^2 pixels). As the lumpiness increases, the strength of the blobs b_0 increases with b_0 equal 1280, 4047.7, and 12800 counts/sec, while the mean number of blobs is 100. (b) Lumpy backgrounds of type II: the values of $\overline{B}$ and $W_f(0)$ are the same as described in (a). As the lumpiness increases, the variance of the noise σ increases and σ equal 33.33, 105.41, and 333.33 pixels. As $\overline{B}$ increases, the strengh of the filter takes the values 3, 13, and 25, respectively.

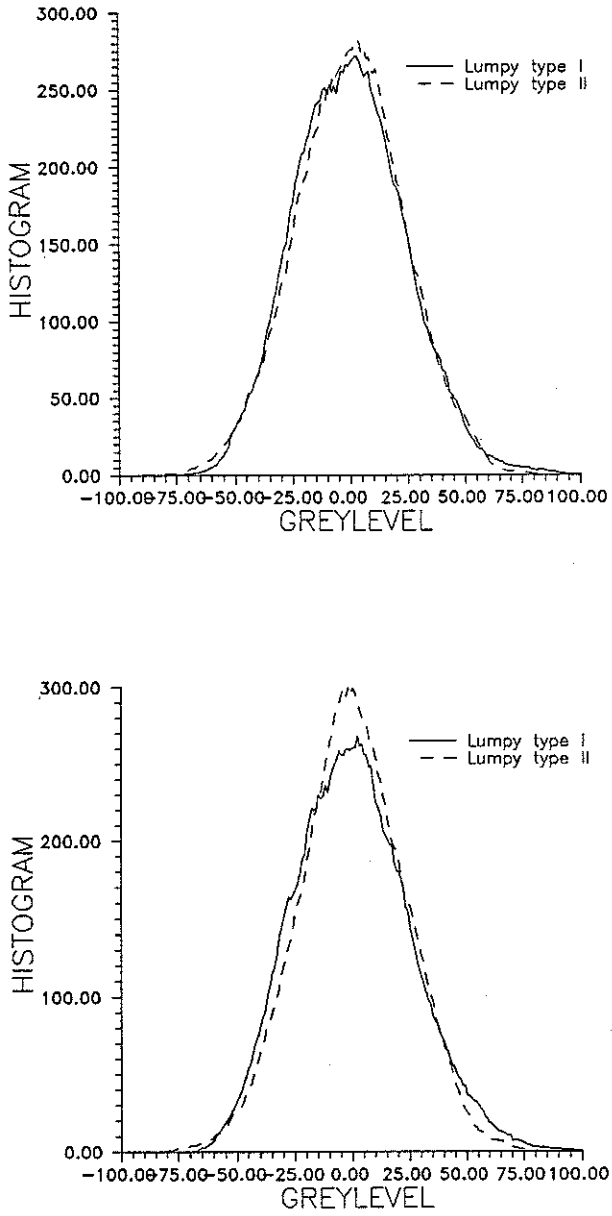


Fig. 6.2. Typical histograms of lumpy backgrounds of type I and II that were shown in Figure 6.1 for (a) $W_f(0)$ equal 10^6 counts^2/(sec^2 pixels) and $\overline{B}$ equal 384 counts/(sec pixels) (b) $W_f(0)$ equal 10^6 counts^2/(sec^2 pixels) and $\overline{B}$ equal 3200 counts/(sec pixels). The histograms were generated before Poisson noise was added to the backgrounds. Each histogram is an average of the histograms of about 50 images generated using either type I or type II lumpy backgrounds.

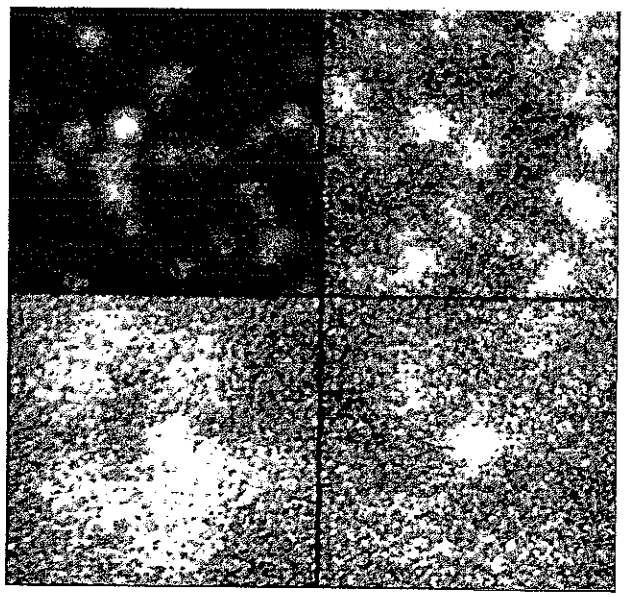


Fig. 6.3. This figure shows how detection of a signal decreases as the value of the correlation length of the background relative to the signal to detect decreases. The signal width r_s is 5.66 pixels, the width of the pinhole aperture r_p equal r_s, the mean background $\bar{B}$ equal 230 counts/(sec pixels), and the contrast of the signal on the background 50%. The lumpiness is constant for all the images, $W_f(0) = 1.3 \times 10^5$ counts/(sec^2 pixels), the strength of the blobs being 6553.6 counts/sec and the mean number of blobs 50. The exposure time is 3 sec. The correlation length r_b is 0.5 r_s in the upper left corner, 1.0 r_s in the upper right corner, 2.0 r_s in the lower left corner and 3.0 r_s in the lower right corner.

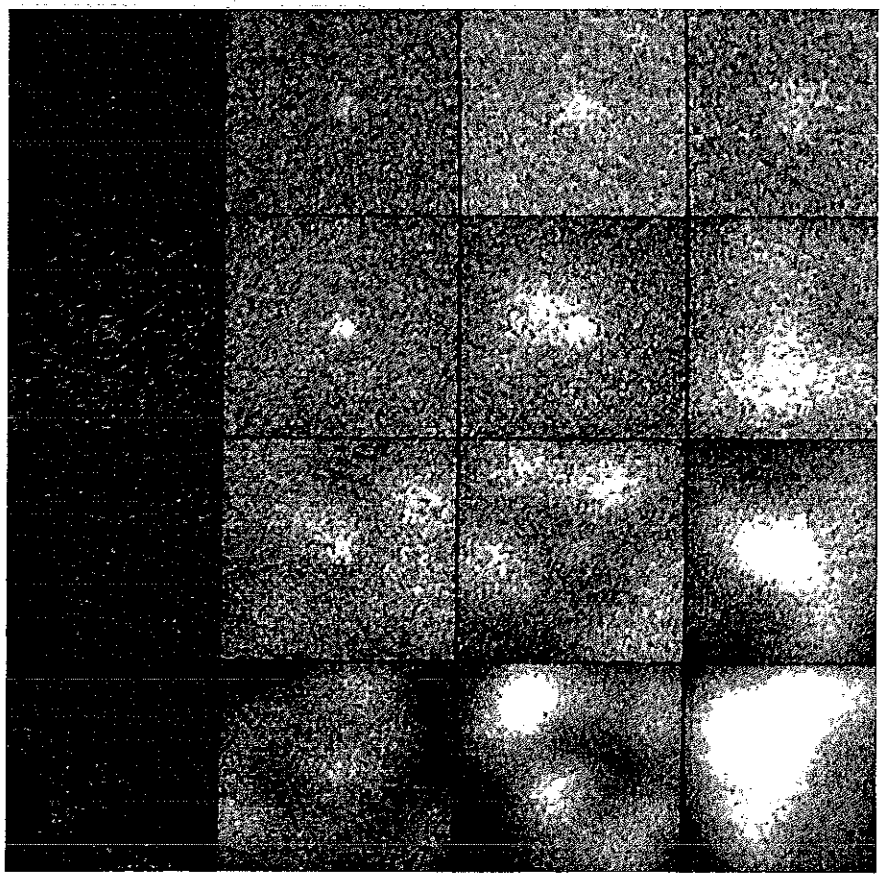


Fig 6.4. Illustrative images of a signal on various backgrounds. The left column shows Gaussian-pinhole (r_p = 0.4 r_s) images of a Gaussian signal (50% contrast, r_s=14.14 pixels, centered in the object array) superimposed on lumpy backgrounds with correlation length of 42.43 pixels for $W_f(0)$ = 0, 1.3×10^5, 8.2×10^5 and 3.3×10^6 counts^2/(sec^2 pixels) from top to bottom. The second column shows the same objects imaged by a Gaussian aperture with r_p= 0.8 r_s. The third and fourth columns show the same objects imaged by an aperture size r_p equal 1.6 r_s and 3.4 r_s respectively.

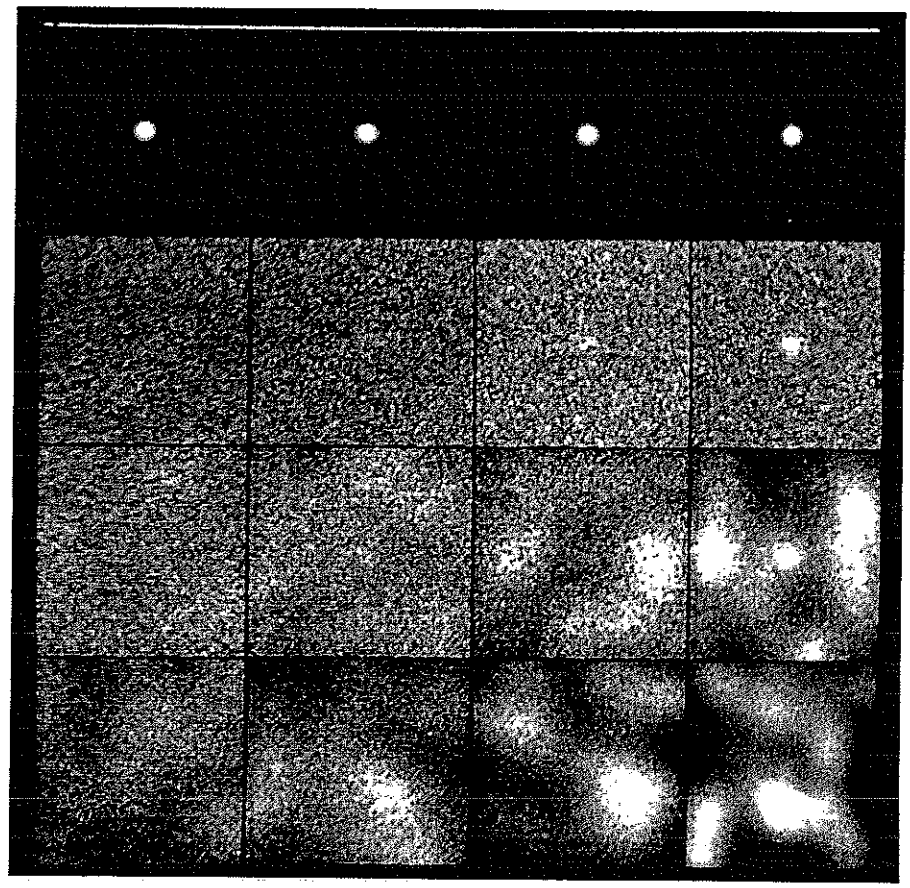


Fig 6.5. Illustrative images of a signal on various backgrounds and imaged with various exposure times. The top row shows the signal (r_s = 5.66 pixels) that is present in the images below the signal. The exposure time increases from left to right with T equal 1, 3, 10, and 50 sec, while the lumpiness increases from top to bottom with $W_f(0)$ equal 0, 1.3×10^5, and 8.2×10^5 counts^2/(sec^2 pixels). In each image, the mean background is 250 counts/(sec pixels) and the contrast of the signal on the background is 10%.

Psychophysical Studies

We shall present two psychophysical studies, one to investigate the importance of aperture size on human observer performance and a second to look at the effect of increasing the exposure time. According to the theory of lumpy backgrounds developed in chapter 5, we have a large choice of parameters to vary, such as the shape and size of the aperture and the type of task such as detection vs. discrimination. We shall present here the results of the psychophysical studies carried out for a simple two hypotheses detection task using a Gaussian aperture as the imaging device.

Parameter Description

In the first study, the signal was known exactly, and the assigned parameter values were $a_s = 6x10^4$ counts/sec and $r_s = 5.66$ pixels or 2.4 mm. The background was specified by a DC level of $3x10^3$ counts/(sec pixels) and the mean number of blobs $\overline{K}$, which was actually the mean of a Poisson random process, was assigned the values 0, 50, and 100 for $W_f(0)$ equal to 0, $1.2x10^8$, and 2.4 10^8 counts^2/(sec^2 pixels), respectively. The strength of the blobs b_0, on the other hand, was kept constant and equal to $2x10^5$ counts/sec. The value of the mean background was then slightly changing with the amount of lumpiness as given by equation (5.12) as well as the signal contrast which was 19.9%, 16.1%, and 14.1% as $W_f(0)$ increased. The exposure time of the imaging system was set to be 1 sec, while the efficiency factor κ was set to 10^{-3} pixels^{-1}.

In the second study, where we studied the exposure time dependence, we assigned a value of $2.5x10^3$ counts/sec to a_s and 2.4 mm to r_s. We set the mean number of blobs to be a constant of value 50 , and the strength of the blobs b_0 to be the variable parameter with b_0 equal 0, $6.5x10^3$, and $1.6x10^4$ counts/sec when $W_f(0)$ equal 0, $1.3x10^5$, and $8.2x10^5$ counts^2/(sec^2 pixels), respectively. The contribution of the dc level to the mean background

level was 250, 230 and 200 counts/(sec pixels) as the lumpiness increased from 0 to 8.2×10^5 counts/(sec^2 pixels). The mean background level was then constant of value 250 counts/(sec pixels), and the contrast of the signal on the background was 10% for the three values of lumpiness.

Protocol

The protocol used to carry out the psychophysical studies and described here is similar to the protocol described in chapter 3, but presents some new features. We designed 12 and 11 experiments within the first and the second study, respectively. Within each experiment, 70 images were generated; 35 images simulated a bright tumor, located in the center of the image, and superimposed on either a uniform background or a lumpy background. Among the 70 images of each experimental set, 6 images were extracted to constitute a training set for the observers. For the first study, the training set was thus composed of 72 images (12 times 6), while 66 images (11 times 6) were used to train the observers for the second study. The remaining 64 images within each set were then used as the test sample to evaluate the human observer performance. For each experiment, we studied the effect of the aperture size or the exposure time for the case of a uniform background and two different values of lumpiness. Once Poisson noise was added to the computer-simulated lumpy images, eight bits of greylevel were used to display the images. The images were displayed as 128x128-pixel array using the greylevel-to-brightness transfer curve shown in Figure 4.12 (b) before linearization . One image was displayed at a time and covered 5.4 cm x 5.4 cm on the display, and each pixel measured 0.42 x 0.42 mm^2. The light in the room was dimmed down to the same value for all the observers. The contrast and brightness levels of the display monitor were fixed during the study, and the observers were not allowed to vary them to optimize their performance. Each observer viewed the

images binocularly at a comfortable distance and was required to wear his/her usual correction lenses.

Ten different observers ran each study, each performing either 11 experiments or 12 experiments according to study I or II, respectively. The observers did not have any information on the number of occurrences of the signal. Within each study, the images from the experiments were displayed in a random order. The observer was presented a single stimulus on each trial, that is a noisy image with or without the signal. The observers were instructed to rate their certainty on a six-point scale given in Table 4.1, and the responses were recorded and analysed using ROC analysis.

Results

Variation With Aperture Size

Given the set of parameters described above, we computed the performance of the Hotelling and the npw ideal observers as a function of the relative pinhole size r_p/r_s, which varied from 0.2 to 4. The predictions of these two observers are given in Figure 6.6. The top curve represents the performance of these two observers for the case of a uniform background where quantum noise is the limitating factor. The performance of these two observers is, in the case of a uniform background, equivalent to the performance of the ideal observer, and therefore the two resulting theoretical curves perfectly overlap as shown. The two middle curves correspond to the Hotelling performance while the two lower curves correspond to the npw ideal observer. Given this set of theoretical curves that indicate very different predictions for the Hotelling and the npw ideal observers, we are interested in finding if one of these two models is a good predictor of human performance.

The psychophysical study was based on the same set of parameters as described

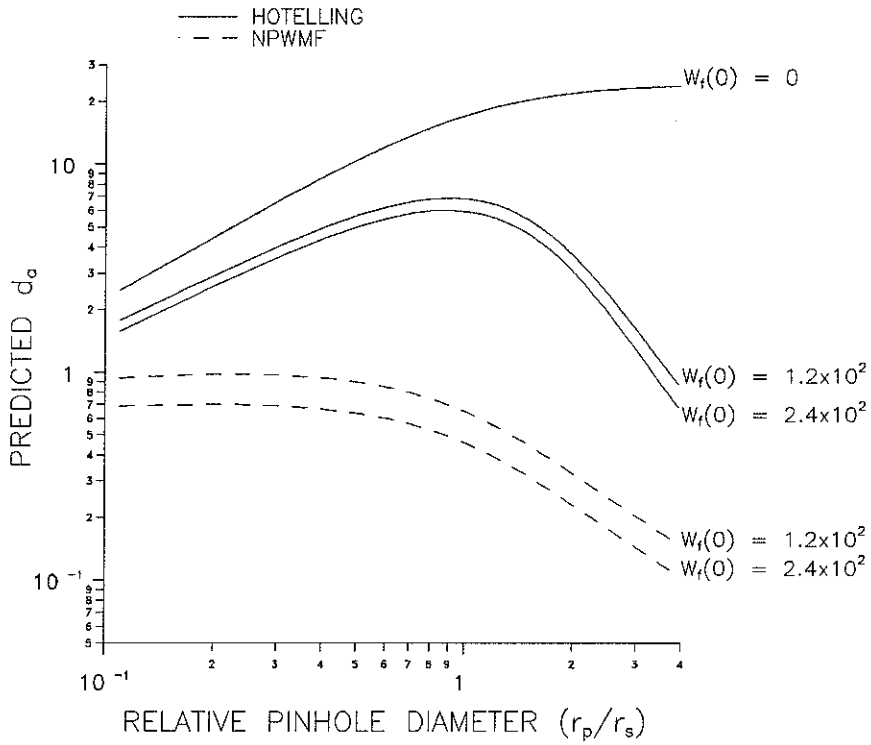


Fig. 6.6 Plot of the detectabilities predicted by the Hotelling and the npw ideal observers for the detection of a low-contrast signal on uniform ($W_f(0) = 0$) and nonuniform ($W_f(0) \neq 0$) backgrounds as a function of the size of the pinhole aperture r_p. The width of the signal r_s is 5.66 pixels, and the contrast of the signal is 19.9%, 16.5%, and 14.1% as lumpiness increases, since the dc background level is kept constant (3000 counts/(sec pixels)) as we increase lumpiness, but the mean background level is a function of both the dc background level and the lumpiness (see equation (5.12)). The mean numbers of blobs are 0, 50 and 100 as $W_f(0)$ increases, and the strength of the blobs b_0 is constant (2×10^5 counts/sec).

above, but with r_p/r_s limited to four values, 0.2, 0.8, 1.5, and 3.4. The results of the study are given in Figure 6.7. The evaluation is reported after performing an average of the performance of each individual observer, as it has been done for all the studies presented in this work. Figure 6.7 plots the average detectability index d_a vs. the four values of r_p/r_s. The estimated standard deviation at each point is given by the mean rms error between the observers.

We first note that, for uniform backgrounds, the performance of the human observer does not saturate as predicted by the ideal observer; rather, it presents an optimum when r_p is approximately matched to the size of the signal r_s. As the lumpiness increases, the performance of the human observer decreases, and the optimum aperture size decreases. Moreover, if we overlap these experimental results on the theoretical curves with an upward shift to take into account the efficiency of the human observer, we find a very good agreement of the data with the Hotelling observer predictions, while the npw ideal observer strongly fails to predict human performance for the detection of a signal in a nonuniform background.

Variation With Exposure Time

Given the set of parameters described above for this study, we set the aperture size r_p to 5.66 pixels, so that $r_p = r_s$. We then let the counting time or exposure time vary from 1 to 100, and we computed the predictions of our two theoretical observers, the Hotelling observer and the npw ideal observer. The predictions are reported in Figure 6.8, where we plot the detectability index d_a as a function of exposure time. The predictions show an increase of the Hotelling performance as a function of time, for both uniform and nonuniform backgrounds. The predicted values of the slopes are about 0.5 and 0.4 for uniform and nonuniform backgrounds, respectively. The performance of the npw ideal

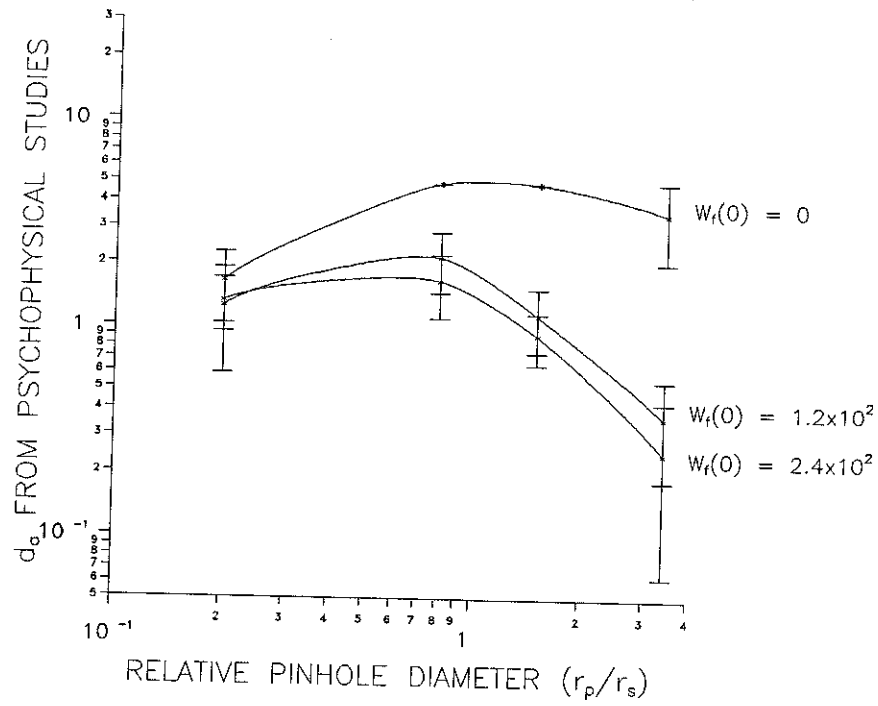


Fig. 6.7 Values of detectabilities obtained from the psychophysical studies for the detection of a low-contrast signal on uniform ($W_f(0) = 0$) and nonuniform ($W_f(0) \neq 0$) backgrounds as a function of the size of the pinhole aperture. The values of the parameters used to generate the computer-simulated images are the same as those given in Fig. 6.6. Four values of the size of the pinhole aperture r_p relative to the signal size r_s are chosen in this case, r_p/r_s equal 0.2, 0.8, 1.5, and 3.4.

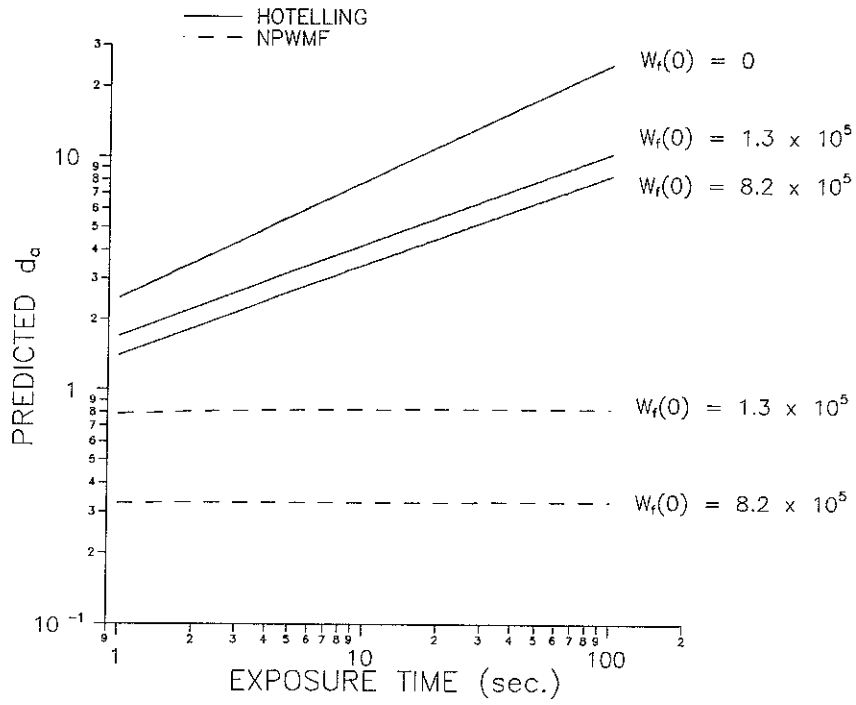


Fig. 6.8 Plot of the detectabilities predicted by the Hotelling and the npw ideal observers for the detection of a low-contrast signal on uniform ($W_f(0) = 0$) and nonuniform ($W_f(0) \neq 0$) backgrounds as a function of the exposure time T. The width of the signal is 5.66 pixels, and the contrast of the signal is 10%. As the amount of lumpiness increases, the dc background levels are 250, 230, and 200 counts/(sec pixels) such that the mean background level is constant as lumpiness inceases, $\bar{B}$ = 250 counts/(sec pixels). The mean number of blobs is 50, and the strengths of the blobs are 0, 6.55×10^3, and 1.64×10^4 counts/sec as $W_f(0)$ increases.

observer, on the other hand, is constant with increasing counting time. This indicate that the performance is conspicuity limited for the chosen values of lumpiness and exposure times. This fact can be better seen if we refer back to the theoretical curve shown in chapter 5. Figure 5.2 shows clearly the transition between quantum-limited and conspicuity-limited performance for $W_f(0) = 400$. In this experiment we would as well see this transition if we were to start the calculations for smaller values of T such as 10^{-2} sec.

The results of the exposure time study are presented in Figure 6.9. The parameters chosen to simulate the images were identical to those used in the theoretical predictions just presented, where the exposure time values were 1, 3, 10, and 20 sec for $W_f(0) = 0$; The values were 1, 3, 10, and 50 for $W_f(0)$ equal 1.3×10^5, and 1, 3, 10, and 100 for $W_f(0)$ equal 8.2×10^5 counts^2/(sec^2 pixels). Results show an increase in performance as a function of exposure time regardless of the amount of lumpiness. These results show that the npw ideal observer does not predict human performance for the detection of a SKE in a nonuniform background. Moreover, if we overlap the experimental results with the theoretical predictions of the Hotelling observer, we see that the Hotelling observer is again a good predictor of human performance within the efficiency factor of the human observer. Given that we have 30 or 40 data points per regression lines (10 observers with 3 or 4 values of T), we performed a t-test to test the null hypothesis that the slope of the experimental regression lines shown in Figure 6.9 are equal to the slopes of the theoretical lines predicted by the Hotelling observer. We note from Figure 6.9 that, for a given lumpiness value, the variances at each point on the regression line are different as the time varies. To test our hypothesis of equality of the slopes of the regression lines using the t-test statistic, we need to transform the data in such a way that the variances become equal. In a recent paper, Deaton et al. (1983) proposed several methods to deal with this problem, and one of them was to divide the values of the data points by the variances so that the data are transformed

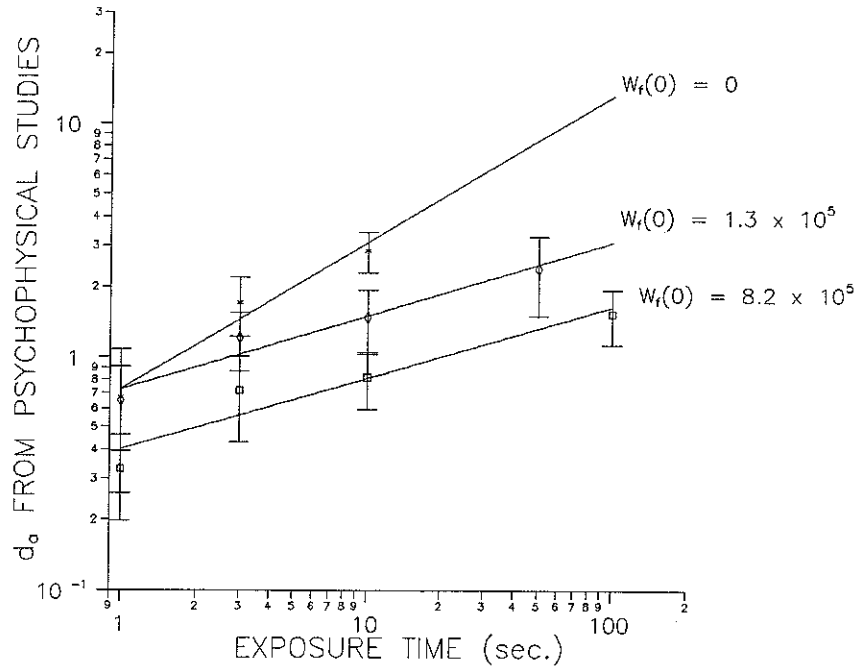


Fig. 6.9 Values of detectabilities obtained from the psychophysical studies for the detection of a low-contrast signal on uniform ($W_f(0) = 0$) and nonunifrom ($W_f(0) \neq 0$) backgrounds as a function of the exposure time T. The values of the parameters used to generate the computer-simulated images are the same as those given in Fig. 6.8. Four values of T are chosen, T equal 1, 3, 10, 50, and 100 sec as shown on the graph.

to unit variances. We then used a standard method to compute the test statistic. The t values were 2.369, 0.312, and 0.236 for the lumpiness values of 0, 1.3 x 10^5, and 8.2 x 10^5 counts^2/(sec^2 pixels), respectively. We can reject the null hypothesis for the nonuniform backgrounds at a significance level of 6.3%; in other words, 6.3% of the time we reject the null hypothesis when the null hypothesis is actually true. For the uniform background, on the other hand, the experimental and theoretical values are further apart; we can reject the null hypothesis at the significant level of 98%.

Discussion

These two psychophysical studies were designed to study the impact of resolution and sensitivity on human performance when the task is the detection of a known signal in a nonuniform background. The results of these studies show a good agreement of the experimental data with the theoretical predictions and thus show that human performance can be well predicted by the Hotelling observer.

There is, however, one deviation of the experimental data from the theoretical data obtained in the case of a uniform background. If we refer again to Figure 6.7, the results show that in a uniform background the performance is optimal when r_p takes a value that is between roughly 0.8 and 1.5 times the size of the signal. For larger apertures, the performance starts to decrease slowly instead of saturating as predicted by the Hotelling and the npw ideal observers, which are then equivalent to the ideal observer (see Figure 6.6).

Such an observation was also pointed out by Tsui et. al. (1982) where the task was the discrimination between two tumor activity distributions embedded in a uniform background. For their experimental data to agree with the theoretical predictions of the ideal observer, Tsui et. al. made the assumption that the efficiency of the human observer itself was a linear function of the aperture size. They then found good agreement of the

data with the predicted values by taking into account this varying efficiency factor. We could also take such an approach to match our experimental data to the theoretical data in the case of uniform backgrounds. However, as the complexity in the background increases, the efficiency of the human observer may be not only a function of the aperture size, but also a function of the lumpiness in the background, and the problem becomes quite complex. Note, however, that even under the assumption that the efficiency of the human observer is constant as a function of aperture size and as a function of lumpiness in the background, our experimental results are in very good agreement with the predictions of the Hotelling observer. A similar drop in efficiency was also reported by White et al. (1989) in a human performance study designed to select the optimum set of parameters to describe a collimator. The task that they proposed was the detection of tumors of varying sizes and locations in computer-simulated liver scans. They found a significant drop in efficiency for the human observer at large aperture sizes.

According to our experience in the long-tailed PSF problem, we postulate that such a drop due to the blurring effect resulting from using a large aperture (compared to the size of the signal to detect) can perhaps be corrected for by processing the images with a high-pass filter before display.

Another small deviation of the experimental data from the theoretical data was found when looking at the effect of exposure time on the performance of the human observer, again for the case of a uniform background. Further investigations should be carried out to fully understand the reason for this discrepancy. Despite this observation, the results show a good agreement of the experimental data with the theoretical data predicted by the Hotelling observer as the exposure time and the lumpiness increase.

CHAPTER 7

CONCLUSIONS

In this chapter we summarize the main findings reported in this dissertation. We then highlight several questions that arose from this work and the potential for further investigations.

Summary

The goal of this research was to contribute to the objective assessment of image quality in medical imaging. First, for the quality to relate to clinical efficacy, it must be assessed in terms of the task to perform. Then, given a well-specified task, statistical decision theory can be used to calculate the performance of some statistical observers, whose performances should predict the performance of the observer who actually performs the task, that is the human observer. Throughout the literature, the most commonly used statistical observer is the ideal observer where the proposed tasks are very stylized as, for example, the detection of a disc signal superimposed on a uniform background and embedded in white Gaussian noise. The reason for considering such stylized tasks is to be able to compute the performance of the ideal observer and gain insight into simple problems before working with more complex images such as clinical images.

Our contribution to the assessment of image quality was to depart from those stylized images and assess image quality for more complex images. Our approach was to look at a simple two-hypothesis detection task, where the signal was superimposed on a nonuniform or lumpy, yet stationary, background instead of a uniform background as for the stylized images. The statistical observers that we considered were the nonprewhitening

(npw) ideal and the Hotelling observers, since they had proven to predict well human-observer performance in previous studies carried out within our research group; those observers were also readily calculable for the task at hand, as defined.

We found a good agreement between the performances of the human and the Hotelling observers, while we showed unambigously that the npw ideal observer was inadequate for the lumpy background problem. To go one step further, the npw ideal observer will certainly be inadequate for the assessment of the quality of most of the clinical images; the Hotelling observer, on the other hand, has given us promising results for the lumpy background problem, and we may expect it to predict human performance for most real clinical images, as well.

In allowing for spatially varying backgrounds, we have also provided new insight into the problem of system optimization in emission imaging in the presence of both quantum limitations and conspicuity limitations. We showed that the choice of task is very important in assessing and optimizing imaging systems. If the task is detection of a known signal on a known, uniform background, the pinhole size should be as large as possible. Obviously, this conclusion is not in accord with clinical experience where some degree of spatial resolution is needed to perform realistic clinical tasks. Stationary nonuniform backgrounds are more realistic than uniform backgounds and lead to the intuitively appealing conclusion that the aperture should be matched to the signals if they are to be detected or discriminated. If the aperture is substantially smaller than the signal, photon collection suffers unnecessarily, while if it is much larger, the spatial resolution is inadequate for reliable discrimination between signal and background.

While we have shown that the Hotelling observer is a very good predictor of human performance for the detection of a signal in complex background, we have, however, found a discrepancy between the Hotelling and the human observer when the task was the

detection of a signal on a smooth background imaged using a long-tailed PSF. In this case, both the Hotelling and the npw ideal observers reduced to the ideal observer.

Our first motivation for studying the long-tailed PSF problem was that such PSFs are commonly encountered in medical imaging and more generally in image science when scatter has not been rejected or effectively corrected. The second motivation for this study came from the property of the ideal observer, that the ideal observer was invariant under linear filtering. It was then interesting to find out whether high-pass filtering of the images would help the human in the detection task, in spite of the quite noisy nature of the simulated images.

We showed that, in the case of long-tailed PSFs, a pseudoideal observer called the nlf (no-low-frequency) ideal observer was a better predictor of human performance than the ideal observer, although human efficiency relative to this observer was still quite low, on the order of 10% before filtering. The nlf ideal observer had the property of suppressing low frequencies in the image, which was also a feature of the so-called channelized ideal observer. The nlf ideal observer could, in this respect, be referred to as a two-channel ideal observer. We also evaluated the performance of the npw ideal observer on the images after deconvolution; the npw ideal observer was found to yield the best performance of all observers investigated, while it failed badly to predict human performance using undeconvolved images.

To show unambigously that deconvolution did help the human observer, we studied the effect of two display parameters on the images: the use of the full range of available greylevels to display each image to the human observer and the effect of photometrically linearizing the greylevel-to brightness transfer curve. Although we found that the display of the images using eight bits of greylevels or less and using either a linear or a nonlinear transfer curve had a strong impact on human performance, we also showed

that deconvolution itself did help the human observer.

This work has graphically demonstrated the limitations of too narrowly specifying the task in model calculations of the performance of imaging systems. Systems that work well, even optimally, on stylized tasks where the background and signal are both known exactly may fail badly with slightly more realistic tasks that include background variability, as shown, for example by the use of the npw ideal observer.

Potential for Future Investigations

Although the Hotelling observer performance has proven to predict well human performance for the detection of a signal in nonuniform backgrounds, it has also failed in the long-tailed PSF problem. In spite of the fact that the assessment of image quality is a strong function of the task to perform, and that the figure of merit must be chosen accordingly, a modified Hotelling observer such as a channelized Hotelling observer could be defined and its performance investigated to bring about a more unified theory.

By studying the long-tailed PSF problem, we have investigated only a quite extreme set of parameters to describe the long-tailed PSFs; we suggest that this model be investigated further starting, for example, with no tails and looking at the degradation introduced by the tails as they extend farther or increase in amplitude.

The author's strongest wish is that further investigations of the role of the display in assessing human performance be pursued. We already know that there is no unique solution to this problem, since it depends on the types of images that we are working with. For example, there is no reason to assume that quantum-limited limited images should be displayed the same way as conspicuity-limited images. However, the optimization (and its definition) of the display of the images should be investigated at least as a function of the types of images.

While the ideal observer is generally a powerful tool to assess image quality for stylized and simple tasks, the Hotelling observer seems to be the appropriate tool for more complex images that approach the complexity found in clinical images. In terms of the lumpy background problem, our psychophysical investigations covered only a small subset of the theory developed in chapter 5. For example, we suggest studying in greater detail the impact of the correlation length of the background relative to the signal size on the performance, both theoretically and experimentally. We also assumed that second-order statistics are sufficient to describe the performance of the human observer. While we know that first- and second-order statistics are sufficient to describe either the npw or the Hotelling observer, no experimental data have supported the fact that it is also true for the human observer. Greater understanding of resolution could also be gained by looking at the effect of a square aperture versus a Gaussian aperture as used in our human studies, since theoretical predictions brought about interesting features.

While we considered only simple open apertures of varying sizes, low-contrast performance with other apertures could also be investigated for the lumpy background problem. More complicated apertures, such as coded apertures, for example, could be used instead, and the performance of some statistical observers could be assessed and compared to the performance of the same observers using simple open apertures. Similarly, performance assessment based on other signal-known-exactly tasks (sine-wave detection, etc.) can be easily achieved through appropriate manipulation of the difference-signal term in the expressions for the detectabilities.

These studies evaluated the ability of model observers to perform simple detection and discrimination tasks. While these studies have included a spatially varying background, further investigations are needed to determine the ability of the statistical observers to perform still more complex tasks, including tasks where there is uncertainty in signal

parameters such as size, amplitude, and location, to further simulate more realistic tasks. Also, all performance measures derived in this work were based on using the raw data to perform the appropriate task. It is not clear whether humans would be able to effectively use raw data for complicated tasks, so studies comparing statistical-observer performance on raw and reconstructed data vs. human performance on both raw and recontructed data sets are also warranted.

While the present studies were based on simulated objects, this work has implications for phantom design as well. We expect that some systems that work well on stylized phantoms where the background is known to be flat might demonstrate performance degradations with more realistic objects that have spatially varying backgrounds.

We have shown on one hand that background or structured noise interfered with the detection of signals or "tumors" and, on the other hand, that image processing did improve the performance of the human observer for the case of a signal on a smooth background imaged using a long-tailed PSF. An interesting area of research would be to study the effect of image processing on more complex images and to assess performance for detection tasks such as the ones defined, for example, for the lumpy backgrounds.

Since many questions remain, we strongly believe that further investigations of some of the work presented will bring more insights not only into medical imaging but also image science.

APPENDIX A

DERIVATION OF THE COVARIANCE MATRIX

In this appendix we derive the covariance matrix of the data vector **g**. We are concerned with a two-dimensional (planar) imaging system with planar objects and images. Nevertheless, we can represent two-dimensional objects and images as one-dimensional column vectors by lexicographically ordering their discretized values, so that the data vector **g** is related to the object vector **f** in the following way:

$$\mathbf{g} = \mathbf{H}\,\mathbf{f} + \mathbf{n} \quad . \tag{A1}$$

If the data vector **g** and noise vector **n** are Nx1 vectors, and the object vector **f** is an Mx1 vector, then the system operator **H** is an NxM matrix. The system operator includes the aperture transmission function as well as the imaging geometry and exposure dependence (see equation (2.6)).

The covariance matrix for the data vector **g** when class k (k=1,2) is true is defined as

$$\mathbf{K}_k = \langle\, \langle\, (\mathbf{g} - \overline{\mathbf{g}_k})\,(\mathbf{g} - \overline{\mathbf{g}_k})^t \,\rangle_{n|f}\,\rangle_{f|k} \quad , \tag{A2}$$

where $\overline{\mathbf{g}_k}$ is the mean image from class k, averaged over both the noise random variable **n**, given an object **f**, and the ensemble of objects **f** that belong to class k. The covariance matrix $\mathbf{K}_k$ characterizes the variation in the data due to the spatially varying object and the Poisson noise. The inhomogeneous background in the object is assumed to be stationary and

independent of the signal. For low-contrast signals, then, $\mathbf{K}_k$ characterizes the same nonuniformity in the background and Poisson noise regardless of the class, so we can drop the subscript k from the covariance matrix ($\mathbf{K}_1 \simeq \mathbf{K}_2 \simeq \mathbf{K}$).

To determine $\mathbf{K}$ we add and subtract $\overline{\mathbf{g}_f}$, the mean data vector averaged over the Poisson noise for a given object $\mathbf{f}$, from each factor in equation (A2) such as

$$\mathbf{K} = \langle \langle (\mathbf{g} - \overline{\mathbf{g}_f} + \overline{\mathbf{g}_f} - \overline{\mathbf{g}}) (\mathbf{g} - \overline{\mathbf{g}_f} + \overline{\mathbf{g}_f} - \overline{\mathbf{g}})^t \rangle_{n|f} \rangle_f \quad . \tag{A3}$$

Equation (A3) may be rewritten as a sum of four matrices $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$, and $\mathbf{M}_4$, with

$$\mathbf{M}_1 = \langle \langle (\mathbf{g} - \overline{\mathbf{g}_f}) (\mathbf{g} - \overline{\mathbf{g}_f})^t \rangle_{n|f} \rangle_f \tag{A4}$$

$$\mathbf{M}_2 = \langle \langle (\overline{\mathbf{g}} - \overline{\mathbf{g}_f}) (\overline{\mathbf{g}} - \overline{\mathbf{g}_f})^t \rangle_{n|f} \rangle_f \tag{A5}$$

$$\mathbf{M}_3 = - \langle \langle (\overline{\mathbf{g}} - \overline{\mathbf{g}_f}) (\mathbf{g} - \overline{\mathbf{g}_f})^t \rangle_{n|f} \rangle_f \tag{A6}$$

$$\mathbf{M}_4 = - \langle \langle (\mathbf{g} - \overline{\mathbf{g}_f}) (\overline{\mathbf{g}} - \overline{\mathbf{g}_f})^t \rangle_{n|f} \rangle_f \quad . \tag{A7}$$

To evaluate each of these matrices, let us call $M_\ell(i,j)$ (with $\ell = 1,2,3,4$) the $(i,j)^{th}$ matrix element of $\mathbf{M}_\ell$. In addition, let us write $\langle g(i) \rangle_{n|f}$, the average data value at the i^{th} pixel given the object $\mathbf{f}$, as $\overline{g_f}(i)$. In component form equation (A4) becomes

$$M_1(i,j) = \langle \langle (g(i) - \overline{g_f}(i)) (g(j) - \overline{g_f}(j)) \rangle_{n|f} \rangle_f \quad , \tag{A8}$$

which, for Poisson noise, and for $i \neq j$, becomes

$$\begin{aligned} M_1(i,j) &= \langle \langle g(i) - \overline{g_f}(i) \rangle_{n|f} \langle g(j) - \overline{g_f}(j) \rangle_{n|f} \rangle_f \\ &= 0 \quad , \; i \neq j \; . \end{aligned} \tag{A9}$$

For i=j, the matrix element $M_1(i,j) = M_1(i,i)$ is given by

$$\begin{aligned} M_1(i,i) &= \langle \langle (g(i) - \overline{g_f}(i))^2 \rangle_{n|f} \rangle_f \\ &= \langle \overline{g_f}(i) \rangle_f \\ &= \overline{g(i)} \\ &= \kappa \, T \, A_{ap} \overline{B} \quad . \end{aligned} \tag{A10}$$

We have again made the low-contrast assumption, so that the stationary background gives a mean pixel value in the image that is independent of position. The value $\overline{B}$ is the mean background level in the object, κ is an efficiency factor, T is the imaging time, and A_{ap} is the area of the aperture ($A_{ap}=L^2$ for the square and $\pi \, r_p^{\,2}$ for the Gaussian aperture).

If we denote the autocorrelation function of the nonuniform background after the imaging process by $\mathbf{R}_g$, $M_2(i,j)$, $M_3(i,j)$, and $M_4(i,j)$ are given for any (i, j) by

$$\begin{aligned} M_2(i,j) &= \langle \langle (\overline{g(i)} - \overline{g_f}(i)) (\overline{g(j)} - \overline{g_f}(j)) \rangle_{n|f} \rangle_f \\ &= \langle (\overline{g(i)} - \overline{g_f}(i)) (\overline{g(j)} - \overline{g_f}(j)) \rangle_f \\ &= R_g(i,j) \quad , \end{aligned} \tag{A11}$$

$$\begin{aligned} M_3(i,j) &= \langle \langle (\overline{g(i)} - \overline{g_f}(i)) (g(j) - \overline{g_f}(j)) \rangle_{n|f} \rangle_f \\ &= \langle (\overline{g(i)} - \overline{g_f}(i)) \langle (g(j) - \overline{g_f}(j)) \rangle_{n|f} \rangle_f \\ &= 0 \quad , \end{aligned} \tag{A12}$$

$$\begin{aligned} M_4(i,j) &= \langle \langle (g(i) - \overline{g_f}(i)) (\overline{g(j)} - \overline{g_f}(j)) \rangle_{n|f} \rangle_f \\ &= \langle \langle (g(i) - \overline{g_f}(i)) \rangle_{n|f} (\overline{g(j)} - \overline{g_f}(j)) \rangle_f \\ &= 0 \quad . \end{aligned} \tag{A13}$$

We can now combine equations (A9) through (A13) to write the $(i, j)^{th}$ element of the covariance matrix defined by equation (A2) as a sum of two terms

$$K(i, j) = R_g(i, j) + \kappa \, T \, A_{ap} \, \overline{B} \, \delta_{ij} \quad , \tag{A14}$$

where δ_{ij} is the Kronecker delta function. The autocorrelation function of the nonuniform background in the image plane $\mathbf{R_g}$ is related to the autocorrelation function of the nonuniform background in the object plane $\mathbf{R_f}$ through a convolution with the aperture function. Equation (A.14) shows that the contributions of the Poisson noise and the nonuniform-background randomness to the covariance matrix of the data are additive.

APPENDIX B

DERIVATION OF THE EXPRESSION FOR THE AUTOCORRELATION FUNCTION OF LUMPY BACKGROUNDS OF TYPE I

For the case of lumpy backgrounds of type I introduced in chapter 5, the random variable used to describe the lumpiness in the background is given by

$$b(\mathbf{r}) = \left[\sum_{i=1}^{K} \delta(\mathbf{r}-\mathbf{r}_i) \right] * y(\mathbf{r}) \quad , \tag{B1}$$

with

$$y(\mathbf{r}) = \frac{b_0}{\pi r_b^{\,2}} \exp\left(-|\mathbf{r}|^2/(r_b^{\,2}) \right) \quad , \tag{B2}$$

where K, the number of Gaussian blobs in the background, is assumed to be a random variable, and $\mathbf{r}_i$ is a uniformly distributed random variable that specifies the location of the i^{th} blob. Given $b(\mathbf{r})$, the expression for the autocovariance of the background is given by

$$R_f(\mathbf{r}-\mathbf{r}'') = \langle \left(b(\mathbf{r})-\langle b(\mathbf{r}) \rangle_f\right) \left(b(\mathbf{r}'')-\langle b(\mathbf{r}'') \rangle_f\right) \rangle_f \tag{B3}$$

where $\mathbf{r}$ and $\mathbf{r}''$ are 2D position vectors and $\langle b(\mathbf{r}) \rangle_f$ is the expectation value of $b(\mathbf{r})$ averaged over the ensemble of images that constitute the lumpy backgrounds. The expression for R_f given by equation (B3) can then be expressed as the sum of four terms

$$R_f = \langle b(\mathbf{r})\, b(\mathbf{r}'') \rangle_f + \langle \langle b(\mathbf{r}) \rangle_f \langle b(\mathbf{r}'') \rangle_f \rangle_f - \langle b(\mathbf{r}) \langle b(\mathbf{r}'') \rangle_f \rangle_f - \langle \langle b(\mathbf{r}) \rangle_f\, b(\mathbf{r}'') \rangle_f \tag{B4}$$

which reduces after simple examination to only two terms

$$R_f = \langle\, b(\mathbf{r})\, b(\mathbf{r}'')\,\rangle_f \;-\; \langle\, b(\mathbf{r})\,\rangle_f \,\langle\, b(\mathbf{r}'')\,\rangle_f \quad . \tag{B5}$$

Let us first look at the first term of equation (B5). It can be explicitly written as

$$\langle\, b(\mathbf{r})\, b(\mathbf{r}'')\,\rangle_f = \Big\langle \sum_{i=1}^{K} [\delta(\mathbf{r}-\mathbf{r}_i) * y(\mathbf{r})] \sum_{j=1}^{K} [\delta(\mathbf{r}''-\mathbf{r}_j) * y(\mathbf{r}'')] \Big\rangle_f$$

$$= \Big\langle \sum_{i=1}^{K} y(\mathbf{r}-\mathbf{r}_i) \sum_{j=1}^{K} y(\mathbf{r}''-\mathbf{r}_j) \Big\rangle_f$$

$$= \sum_{K} P(K) \int d^2\mathbf{r}_1\, pr(\mathbf{r}_1) \ldots \int d^2\mathbf{r}_K\, pr(\mathbf{r}_K) \sum_{i=1}^{K} y(\mathbf{r}-\mathbf{r}_i) \sum_{j=1}^{K} y(\mathbf{r}''-\mathbf{r}_j)\ , \tag{B6}$$

where $pr(\mathbf{r}_i)$ is the probability density associated with the random variable $\mathbf{r}_i$, and P(K) is the probality of having K Gaussian blobs . If i=j, we shall show that each term in the sums over i and j contributes in the same fashion to the left-hand side of equation (B6). For example, the term for i=j=7, denoted as term1(7, 7), becomes

$$\text{term1}(7,7) = \int d^2\mathbf{r}_1\, pr(\mathbf{r}_1) \ldots \int d^2\mathbf{r}_7\, pr(\mathbf{r}_7)\, y(\mathbf{r}-\mathbf{r}_7)\, y(\mathbf{r}''-\mathbf{r}_7) \ldots . \int d^2\mathbf{r}_K\, pr(\mathbf{r}_K)$$

$$= \int d^2\mathbf{r}_7\, pr(\mathbf{r}_7)\, y(\mathbf{r}-\mathbf{r}_7)\, y(\mathbf{r}''-\mathbf{r}_7) \tag{B7}$$

Since $\mathbf{r}_7$ is uniformly distributed, $pr(\mathbf{r}_7)$ is given by $1/A_f$ with A_f being the area of the object. In the particular case where the distances from object to aperture and from aperture to detector are equal, the geometrical magnification is one and A_f equals A_d, the area of the detector. If we now define $\mathbf{r}''$ to be $\mathbf{r} + \mathbf{r}_0$, and if we denote $\mathbf{r}-\mathbf{r}_7$ as $\mathbf{r}'$, we can write equation (B7) as a convolution operation,

$$\text{term1}(7,7) = \frac{1}{A_d} \int d^2\mathbf{r}'\ y(\mathbf{r}')\ y(\mathbf{r}' + \mathbf{r}_0)$$

$$= \frac{1}{A_d}\ (\ y(\mathbf{r}_0) * y(\mathbf{r}_0)\ ) \quad . \tag{B8}$$

The convolution of two Gaussians being a Gaussian, the final expression for the contribution of i=j=7 to the first term of equation (B5) can be computed by substituting the expression for $y(\mathbf{r})$ given by equation (B2) into equation (B8),

$$\text{term1}(7,7) = \frac{1}{A_d}\ \frac{b_0^2}{2\pi r_b^2}\ \exp\left(\ -|\mathbf{r}|^2/(\ 2\ r_b^2)\ \right) \quad , \tag{B9}$$

which is independent of the choice for i and j (assuming i=j). Since there are in average $\overline{K}$ terms such that i=j, the first term of equation (B5) becomes for i=j,

$$\text{term1}(i=j) = \frac{\overline{K}}{A_d}\ \frac{b_0^2}{2\pi r_b^2}\ \exp\left(\ -|\mathbf{r}|^2/(2\ r_b^2)\ \right) \quad . \tag{B10}$$

If $i \neq j$, let us consider the example of i=6 and j=7. The contribution to the first term of equation (B5) is then,

$$\text{term1}(6,7) = \int d^2\mathbf{r}_1\ \text{pr}(\mathbf{r}_1) \ldots \int d^2\mathbf{r}_6\ \text{pr}(\mathbf{r}_6)\ y(\mathbf{r}-\mathbf{r}_6) \int d^2\mathbf{r}_7\ \text{pr}(\mathbf{r}_7)\ y(\mathbf{r}''-\mathbf{r}_7)$$

$$\ldots \int d^2\mathbf{r}_K\ \text{pr}(\mathbf{r}_K)$$

$$= \int d^2\mathbf{r}_6\ \text{pr}(\mathbf{r}_6)\ y(\mathbf{r}-\mathbf{r}_6) \int d^2\mathbf{r}_7\ \text{pr}(\mathbf{r}_7)\ y(\mathbf{r}-\mathbf{r}_7)$$

$$= \frac{1}{A_d{}^2}\ b_0{}^2 \quad . \tag{B11}$$

Since there is an average of $\overline{K^2-K}$ terms with $i \neq j$, the contribution of the terms $i \neq j$ to the first term of equation (B5) is given by

$$\text{term1}(i \neq j) = (\overline{K^2-K})\ \frac{b_0{}^2}{A_d{}^2} \quad . \tag{B12}$$

If we now examine the second term of equation (B5), it is given by

$$\text{term2} = \langle\ b(\mathbf{r})\ \rangle_f\ \langle\ b(\mathbf{r}'')\ \rangle_f$$

$$= \langle \sum_{i=1}^{K} y(\mathbf{r}-\mathbf{r}_i)\ \rangle_f\ \langle \sum_{i=1}^{K} y(\mathbf{r}-\mathbf{r}_i)\ \rangle_f$$

$$= (\overline{K})^2\ \frac{b_0{}^2}{A_d{}^2} \quad . \tag{B13}$$

Finally the expression for the autocorrelation function, that is the sum of term1(i=j), term1(i≠j) and term2 given by equations (B10), (B12) and (B13), respectively, becomes

$$R_f = \frac{\overline{K}}{A_d} \frac{b_0^2}{2\pi r_b^2} \exp\left(-|r|^2/(2\, r_b^2) \right) + (\overline{K^2}-\overline{K}) \frac{b_0^2}{A_d^2} - (\overline{K})^2 \frac{b_0^2}{A_d^2} \quad . \tag{B14}$$

This expression is quite interesting since the autocorrelation function is not found to be Gaussian unless K is itself a Poisson random variable. In this case, $\overline{K}^2-(\overline{K})^2 = \overline{K}$ and the second term of equation (B14) cancels with the third term of the equation. Thus we choose K to be a Poisson random variable and the autocorrelation function reduces to

$$R_f(r) = \frac{\overline{K}}{A_d} \frac{b_0^2}{2\pi r_b^2} \exp\left(-|r|^2/(2\, r_b^2) \right) \quad . \tag{B15}$$

The power spectrum, which is defined as the Fourier transform of the autocorrelation function for stationary random process, is then given by

$$W_f(\rho) = \frac{\overline{K}}{A_d} b_0^2 \exp\left(-2\, \pi^2\, r_b^2\, |\rho|^2 \right)$$

$$= W_f(0) \exp\left(-2\, \pi^2\, r_b^2\, |\rho|^2 \right) \quad , \tag{B16}$$

with

$$W_f(0) = \frac{\overline{K}}{A_d} b_0^2 \quad . \tag{B17}$$

We shall use $W_f(0)$ as the measure of the lumpiness in the background. Note that $W_f(0)$ is a measure of counts^2/(sec^2 mm^2).

APPENDIX C

DERIVATION OF THE EXPRESSION FOR THE AUTOCORRELATION FUNCTION OF LUMPY BACKGROUNDS OF TYPE II

Lumpy backgrounds of type II are generated by filtering uncorrelated Gaussian noise. By definition, the power spectrum of uncorrelated noise $W(\rho)$ is constant, and if we denote by $H(\rho)$ the Gaussian filter used to filter Gaussian noise, the resulting power spectrum is given by

$$W_f(\rho) = W(\rho) \mid H(\rho) \mid^2 \quad . \tag{C1}$$

To generate lumpy backgrounds of type II with the same autocorrelation function as lumpy backgrounds of type I, we should define $H(\rho)$ as

$$H(\rho) = H(0) \exp(- \pi^2 r_b^2 |\rho|^2) \quad , \tag{C2}$$

where $H(0)$ is the amplitude of the filter and r_b is the correlation length of the resulting autocorrelation function. The filtered power spectrum is then given by

$$W_f(\rho) = W(0) [H(0)]^2 \exp(- 2 \pi^2 r_b^2 |\rho|^2)$$

$$= W_f(0) \exp(- 2 \pi^2 r_b^2 |\rho|^2) \tag{C3}$$

with

$$W_f(0) = W(0) [H(0)]^2 \quad . \tag{C4}$$

The expression for the autocorrelation function is then given by the inverse Fourier transform of the power spectrum as

$$R_f(\mathbf{r}) = \frac{W(0)\,[H(0)]^2}{2\,\pi\,r_b^{\,2}}\ \exp\left(-|\mathbf{r}|^2/2\,r_b^{\,2}\right) \quad . \tag{C5}$$

The measure of lumpiness is now given by $W_f(0) = W(0)\,[H(0)]^2$, where H(0) is simply a number and W(0) is in units of counts per unit time and per unit area. Therefore, $W_f(0)$ is again in units of counts^2/(sec^2 mm^2).

REFERENCES

Andrew, B.W., and D.A. Pollen, "Relationship between spatial frequency selectivity and receptive field profile of simple cells," J. physiol. **287**, 163-176 (1979).

Barlow, H.B., "Increment thresholds at low intensities considered as signal/noise discriminations," J. Physiol. **136**, 469-488 (1957).

Barlow, H.B., "Temporal and spatial summation in human vision at different background intensities," J. Physiol. Lond. **141**, 337-350 (1958).

Barlow, H.B., "The efficiency of detecting changes in density of random dot patterns," Vision Res. **18**, 637-650 (1978).

Barnard, T.W., "Image evaluation by means of target recognition," Photographic Science and Engineering **16**, 144-150 (1972).

Barrett, H.H., and W. Swindell, Radiological Imaging: The Theory of Image Formation, Detection, and Processing Vols. I and II, Academic press, New York (1981).

Barrett, H.H., W.E. Smith, K.J. Myers, T.D. Milster, and R.D. Fiete, "Quantifying the performance of imaging systems," Proc. SPIE **535**, 65-69 (1985).

Barrett, H.H., K.J. Myers, and R.F. Wagner, "Beyond signal-detection theory," Proc. SPIE **626**, 231-239 (1986).

Barrett, H.H., J.P. Rolland, R.F. Wagner, K.J. Myers, "Detection and discrimination of known signals in inhomogeneous, random backgrounds," SPIE **1090**, 176-182 (1989).

Barrett, H.H., "Objective assessment of image quality: effects of quantum noise and object variability," J. Opt. Soc. Am. A (1990).

Bibbo, M., E. Alenghat, G.F. Bahr, P.H. Bartels, H.E. Dytch, A.L. Herbst, C.M. Keebler, F. Pishotta, and G.L. Wied, "A quality control procedure on cervical lesions for the comparison of cytology and histology," J. Reproductive Medicine **28**, 811-822 (1983).

Biberman, L.M., Perception of Displayed information, Plenum, New York (1973).

Blackwell, H.R., "Contrast thresholds of human eye," J. Opt. Soc. Am. **36**, 624-643 (1946).

Bracewell, R.N., The Fourier Transform and its Applications, 2nd edition, Mc-Graw-Hill, New York (1986).

Brown, D.G., M.F. Insana, "Decision function efficiency for higher order imaging tasks," SPIE **914**, 76-79 (1988).

Burgess, A.E., K. Humphrey, R.F. Wagner, "Detection of bars and discs in quantum noise," SPIE, 173, 34-40 (1979).

Burgess, A.E., R.F. Wagner, R.J. Jennings, and H.B. Barlow, "Efficiency of human visual discrimination," Science **214**, 93-94 (1981).

Burgess, A.E., R.J. Jennings, and R.F. Wagner, "Statistical efficiency: A measure of human visual signal-detection," J. Appl. Photo. Eng. **8**, 76-78 (1982).

Burgess, A.E., R.F., Wagner, and R.J., Jennings, "Human signal detection performance for noisy medical images," in Proc. IEEE ComSoc Int. Workshop Med. Imag., Asilomar, CA, IEEE Cat. 82CH1751-7 (1982).

Burgess, A.E., "Statistical efficiency of perceptual decisions," SPIE **454**, 18-26 (1984).

Burgess, A.E., "Detection and identification efficiency: an update," SPIE **535**, 50-56 (1985).

Burgess, A.E., "On observer internal noise," SPIE **626**, 208-213 (1986).

Burgess, A.E., personal communications (1988).

Campbell, F.W., and D.G. Green, "Optical and retina factors afecting visual resolution," J. Physiol. **181**, 576-593 (1965).

Campbell, F.W., and J.G. Robson, "Application of Fourier analysis to the visibility of gratings," J. Physiol. Lond. **197**, 551-566 (1968).

Campbell, F.W., R. Carpenter, and J. Levinson, "Visibility of aperiodic patterns compared with cortical cells compared with sinusoidal gratings," J. Physiol. **204**, 283-298 (1969).

Cattell, J.M., "On errors of observation," Am. J. Psych. **5**, 285-293 (1893).

Clough, A.V., A mathematical model of single-photon emission computed tomography, Ph.D. dissertation (1986).

Cormack, J., B.F., Hutton, "Minimization of data transfer losses in the display of digitized scintigraphic images," Phys. Med. Biol. **25**, 271-282 (1980).

Dainty, D.C., and R. Shaw, Image Science, Academic Press, London (1974).

Deaton, M.L., "Estimation and Hypothesis testing in regression in the presence of nonhomogeneous error variances," Commun. Satist.-Simula. Computa. **12**, 45-66 (1983).

DePalma, J.J., and E.M. Lowry, "Sine-wave response of the visual system," J. Opt. Soc. Am. **228**, 328-335 (1962).

De Valois, K., R.L. De Valois, and E.W. Yund, J. Physiol. **291**, 483-505 (1979).

Dorfman, D.D., and Alf. E., Jr., "Maximum likelihood estimation of parameters of signal detection theory and determination of confidence intervals-rating method data," J. Math. Psych. **6**, 487-496 (1969).

Egan, J.P., Signal detection theory and ROC analysis, Academic Press, New York (1975).

Fahimi, H., and A. Macovski, "Reducing the effects of scattered photons in x-ray projection imaging," IEEE Trans. Med. Imag. **8**, 56-63 (1989).

Fiete, R.D., H.H. Barrett, W.E. Smith, and K.J. Myers, "The Hotelling trace criterion and its correlation with human observer performance," J. Opt. Soc. Am. A **4**, 945-953 (1987).

Fisher, R.A., "The use of multiple measurements in taxonomic problems," Ann. Eugenics **7**, 179-188 (1936).

Fukunaga, K., Introduction to Statistical Pattern Recognition, Academic Press, New York (1972).

Gaskill, J.D., Linear Systems, Fourier Transforms, and Optics, Wiley, New York (1978).

Geisler, W.S., "Sequential ideal-observer analysis of visual discriminations," Psychol. Review **96**, 267-313 (1989).

Georgeson, M., "Mechanisms of visual image processing : studies of pattern interaction and selective channels in human vision," Ph.D. dissertation, Univ. Sussex, Brighton, England (1975).

Giger, M.L., and K. Doi, "Investigation of basic imaging properties in digital radiography. 3. Effect of pixel size on SNR and threshold contrast," Med. Phys. **12**, 201-208 (1985).

Giger, M.L., K. Ohara, K. Doi, "investigation of basic imaging properties in digital radiography. 9. Effect of displayed grey levels on signal detection," Med. Phys. **13**, 312-318 (1986).

Goodenough, D.J., C.E. Metz, and L.B. Lusted, "Caveat on the use of the parameter d' for evaluation of observer performance," Radiology **106**, 565-566 (1973).

Goodenough, D.J., and C.E. Metz, "Implications of a "noisy" observer to data processing techniques," in Raynaud C., Todd-Pokropek A.E., eds. Information processing in scintigraphy. Orsay, France: CEA (1975).

Goodman, J.W., Introduction to Fourier Optics, Mc-Graw-Hill, New York (1988).

Green, D.B., and J.A. Swets, Signal Detection Theory and Psychophysics, Wiley, New York (1966). (Reprint with corrections: Huntington, New York, 1974).

Grettenberg, T.L., "Signal selection in communication and radar systems," IEEE Trans. Inf. theory, 265-275 (1963).

Griffiths, J.W., and N.S. Nagaraja, "Visual detection in intensity modulated displays," J. Brit. IRE **25**, 225-240 (1963).

Gu, Z.H., and S. Lee, "Optical implementation of the Hotelling trace criterion for image classification," Opt. Eng. **23**, 727-731 (1984).

Harris, J.L., "Resolving power and decision theory," J. Opt. Soc. Am. **54**, 606-611 (1964).

Hanley, J.A., and B.J. Neil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," Radiology **143**, 29-36 (1982).

Hanson, K.M., "Detectability in computed tomographic images," Med. Phys. **6**, 441-451 (1979).

Hanson, K.M., "Variations in task and the ideal observer," SPIE **419**, 60-67 (1983).

Heuscher, D.J., and C.J. Brunnett, "Signal-to-noise frequency analysis of CT systems," SPIE **914** (1989).

Hotelling, H., "The generalization of Student's ratio," Ann. Math. Stat. **2**, 360 (1931).

Hubel, D.H., And T.N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," J. Physiol. **160** (1962).

Ishida, M., K. Doi, L.L. Loo, C.E. Metz, and J.L. Lehr, "Digital image processing : Effect on detectability of simulated low-contrast radiographic patterns," Radiology **150**, 569-575 (1984).

Jaszczak, R., K. Greer, C. Floyd, C. Harris, R. Coleman, "Improved spect quantification using compensation for scattered photons," J. Nucl. Med. **25**, 893-900 (1984).

Judy, P.F., and R.G. Swensson, "Detection of small focal lesions in CT images : effects of reconstruction filters and visual display windows," The British Journal of Radiology **58**, 137-145 (1985).

Judy, P.F., and R.G. Swensson, "Display thresholding of images and observer detection performance," JOSA. A. **4**, 954-965 (1987).

Judy, P.F, R.G. Swensson, and M.F. Kijewski, "Observer efficiency and feature polarity," SPIE **767**, 310-316 (1987).

Kaufman, L., D. Shosa, L. Crooks, and J. Ewins, "Technology needs in medical imaging," IEEE Trans. Med. Imaging **MI-1**, 11-16 (1882).

Kennett, T.J., and Prestwich, W.V., "On deconvolution of exponential response functions," Phys. Med. Biol. **24**, 1107-1122 (1979).

Ker, M., G.W. Seeley, M.O. Stempski, and D. Patton, "A protocol for verifying truth of diagnosis," Invest. Radiology **23**, 485-487 (1988).

Kulikowski, J.J., and P.E. King-Smith, "Orientation selectivity of grating and line detectors in human vision," Vis. Res. **13**, 1455-1478 (1973).

Legge, G.E., D. Kersten, A.E. Burgess, "Contrast discrimination in noise," J. Opt. Soc. Am. A **4**, 391-404 (1987).

Lusted, L.B., "Logical analysis in roentgen diagnosis," Radiology **74**, 178-193 (1960).

Lusted, L.B., "Signal detectability and medical decision-making," Science **171**, 1217-1219 (1971).

Maffei, L., and A. Fiorentini, "The unresponsive regions of visula cortical receptive fields," Vis. Res. **16**, 1131-1139 (1976).

Maffei, L. C. Morrone, M. Pirchio, and G. Sandini, J. Physiol. **296**, 24-47 (1979).

Metz, C.E., "Basic principles of ROC analysis," Sem. Nucl. Med. **8**, 283-298 (1978).

Metz, C.E., "ROC methodology in radiologic imaging," Invest. Radiol. **21**, 720-733 (1986).

Metz, C.E., "Some practical issues of experimental design and data analysis in radiological studies," Invest. Radiology **24**, 234-245 (1989).

Msaki, P., B. Axelsson, C.M. Dahl, and S.A. Larsson, "Generalized scatter correction method in SPECT using point scatter distribution functions," J. Nucl. Med. **26**, 1861-1869 (1987).

Myers, K.J., H.H. Barrett, M.C. Borgstrom, D.D. Patton, and G.W. Seeley, "Effect of noise correlation on detectability of disk signals in medical imaging," J. Opt. Soc. Am. A **2**, 1752-1759 (1985).

Myers, K.J., and H.H. Barrett, "Addition of a channel mechanism to the ideal-observer model," J. Opt. Soc. Am. A4, 2447-2457 (1987).

Myers, K.J., and R.F. Wagner, "Detection and estimation: Human vs. ideal performance as a function of information," SPIE 1090, 291-297 (1988).

Myers, K.J., R.F. Wagner, D.G. Brown, and H.H. Barrett, "Efficient utilization of aperture and detector by optimal coding," Proc. SPIE **1090**, 164-175 (1989).

Nachmias, J., and A. Weber, "Discrimination of simple and complex gratings," Vis. Res. **15**, 217-223 (1975).

Nagaraja, N.S., "Effect of luminance noise on contrast thresholds," J. Opt. Soc. Am. **54**, 950-955 (1964).

Panel on Discriminant Analysis, Classification, and Clustering, "Discriminant Analysis and Clustering," Stat. Science **4**, 34-69 (1989).

Patton, D.D, "Introduction to clinical decision making," Sem. Nucl. Med. **8**, 273-282 (1978).

Pelli, D.G., "Uncertainty explains many aspects of visual detection and discrimination," J. Opt. Soc. Am. A2, 1508-1532 (1985).

Peterson, W.W., T.G. Birdsall, and W.C. Fox, "The theory of signal detectability," IRE Trans. PGIT **4**, 171-212 (1954).

Pizer, S.M., "Intensity mappings to linearized display devices," Computer graphics and image processing **17**, 262-268 (1981).

Ratliff, F., Mach Bands Holden-Day, San Franscisco (1965).

Revesz, G., H. L. Kundel, and M. A. Graber, "The influence of structured noise on detection of radiologic abnormalities," Invest. Radiology **9**, 479-486 (1974).

Roehrig, H., W.J. Dallas, T.L. Ji, R.D. Lamoreaux, D. Oikawa, R. Vercillo, and D. Yocky, "Physical evaluation of CRT's for use in digital radiography," SPIE, Medical Imaging III 1091 (1989).

Sachs, M., J. Nachmias, and J. Robson, "Spatial-frequency channels in human vision," J. Opt. Soc. Am. **61**, 1176-1186 (1971).

Seeley, G.W., M.C. Borgstrom, and J. Mazzeo, "A general interactive computer program for running signal detection experiments," Behavior Res. Methods Instrum. **4**, 555-556 (1982).

Seibert, J.A., and J.M. Boone, "X-ray scatter removal by deconvolution," Med. Phys. **15**, 567-575 (1988).

Seltzer, S.E., R.G. Swensson, P.F. Judy, and R.D. Nawfel, "Size discrimination in computed tomographic images, effect of feature contrast and display window," Invest. Radiology **23**, 455-462 (1988).

Shaw, R., "Evaluating the efficiency of imaging processes," Rep. Prog. Phys. **41** 1103-1155 (U.K., 1978).

Shosa, D., and L. Kaufman, "Methods for evaluation of diagnosis imaging instrumentation," Phys. Med. Bio. **26**, 101-112 (1981).

Simpson A.J., and M.J. Fitter, "What is the best index of detectability?," Psych. Bull. **80**, 481-488 (1973).

Smith, W.E., and H.H. Barrett, "The Hotelling trace criterion as a figure of merit for the optimization of imaging systems," J. Opt. Soc. Am **3**, 717-725 (1986).

Swensson, R.G., and P.F. Judy, "Detection of noisy visual targets: models for the effect of spatial uncertainty and signal-to-noise ratio," Perception and Psychophysics **9**, 521-534 (1981).

Swets, J.A., W.P. Tanner, and T.G. Birdsall, "Decision processes in perception," Psychophysical Review **68**, 301-340 (1961).

Swets, J.A., "ROC analysis applied to the evaluation of medical imaging techniques," Invest Radiol. **14**, 109-121 (1979).

Swets, J.A., and R.M. Pickett, Evaluation of Diagnostic Systems, Academic Press, New York (1982).

Swets, J.A., "Form of empirical ROCs in discrimination and diagnosis tasks: implications for theory and measurements of performance," Psych. Bull. **99**, 181-198 (1986).

Swets, J.A., "Measuring the accuracy of diagnosis systems," Science **240**, 1285-1293 (1988).

Tanner, W.P., and T.G. Birdsall, "Definition of d′ and η as psychophysical measures," J. Acoust. Soc. Am. **30**, 922-928 (1958).

Tsui, B.M.W, C.E. Metz, F.B. Atkiss, S.J. Starr, and R.N. Beck, "A comparison of optimum detector spatial resolution in nuclear imaging based on statistical theory and on observer performance," Phys. Med. Bio. **23**, 654-676 (1978).

Tsui, B.M.W, C.E. Metz, and R.N. Beck, "Optimum detector spatial resolution for discriminating between tumor uptake distributions in scintigraphy," Phys. Med. Bio., 775-788 (1982).

Van Meter, D., and D. Middleton, "Modern statistical approaches to reception in communication theory," IRE Trans. PGIT **4**, 119-141 (1954).

Van Trees, H.L., Detection, Estimation, and Modulation Theory, Vols. I-III, John Wiley, New York (1968).

Watson, A.B., H.B. Barlow, and J.G. Robson, "What does the eye see best?," Nature **302**, 419-422 (1983).

Wagner, R.F., "Toward a unified view of radiological imaging systems. Part II: Noisy images," Med. Phys. **4**, 279-296 (1977).

Wagner, R.F., D.G. Brown, and C.E. Metz, "On the multiplex advantage of coded source/aperture photon imaging," Proc. SPIE **314**, 72-76 (1981).

Wagner, R.F., and D.G., Brown, "Overview of a unified SNR analysis of medical imaging systems," IEEE Trans. Med. Imag. **MI-1**, 210-214 (1982).

Wagner, R.F., "Low contrast sensitivity of radiologic, CT, nuclear medicine, and ultrasound medical imaging systems," IEEE Trans. Med. Imag., MI-2 **3**, 105-121 (1983).

Wagner, R.F., and D.G. Brown, "Unified SNR analysis of medical imaging systems," Phys. Med. Biol. **30**, 489-518 (1985).

Wagner, R.F., and H.H. Barrett, "Quadratic tasks and the ideal observer," SPIE **767**, 306-309 (1987).

Wagner, R.F., K.J. Myers, D.G. Brown, M.J. Tapiovaara, and A.E. Burgess, "Higher-order tasks: human vs. machine performance," Proc. SPIE **1090**, 183-194 (1989).

Wagner, R.F., K.J. Myers, M.J. Tapiovaara, D.G. Brown, A.E. Burgess, "Maximum a posteriori detection and figures of merit for detection under uncertainty," SPIE 1231 (1990).

Whalen, A.D., Detection of Signals in Noise, Academic Press, New York (1971).

White, T.A., H.H. Barrett, E.B. Cargill, R.D. Fiete, and M. Ker, "The use of the Hotelling trace to optimize collimator performance," J. Nucl. Med. **30**, 892 (abstract) (1989).

Wickelgren, W.A., "Unidimensional strength theory and component analysis of noise in absolute and comparative judgments," J. Math. Psych. **5**, 102-122 (1968).

Yanch, J.C., A.T. Irvine, S. Webb, and M.A. Flower, "Deconvolution of emission tomographic data: a clinical evaluation," The British Journal of Radiology **61**, 221-225 (1988).