



A Study Guide for
**Digital Image
Processing**

Mark J. T. Smith
Alen Docef

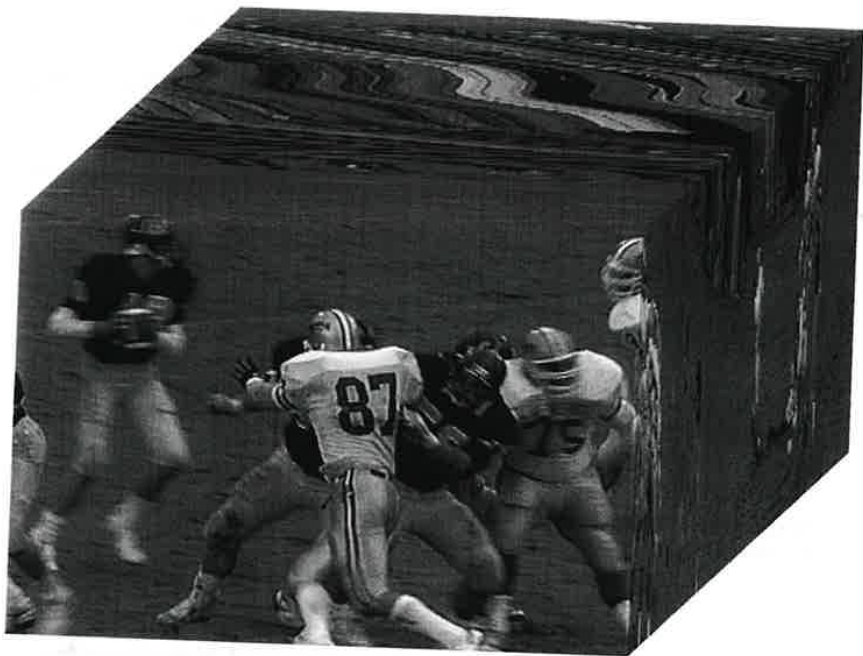


Scientific Publishers

FT 1637
GenSci

TA 1637
S64
1997
Copy 2

*A Study Guide for
Digital Image Processing*



Mark J. T. Smith

Alen Docef

Early Release Edition

This text is an early release edition, printed by the publisher for exclusive use by the authors for the course EE6418 at the Georgia Institute of Technology, during Spring quarter, 1997. The first edition is scheduled for release in January 1998.



TA 1637
.S64
1997
Copy 2

Library of Congress Cataloging-in-Publication Data

A C.I.P. Catalogue record for this text is available from the Library of Congress.



copyright © 1997 by Scientific Publishers, Inc.
Riverdale, Georgia 30274

All rights reserved. No part of this textbook may be reproduced, digitally scanned and stored or transmitted, without express permission in writing from the publisher.

*Scientific Publishers, Inc.
6404 Church Street
Suite 7
Riverdale, Georgia 30274*

Printed in the United States of America

97-186007

Contents

PREFACE	ix
Personal Acknowledgments	xi
1 Introduction	1
1.1 Digital Processing Systems	2
1.2 Topics in Image Processing	11
1.3 Prerequisites	22
1.4 Practice Exam in DSP	23
1.5 General References	31
2 Mathematical Preliminaries	33
2.1 Basic Linear Algebra	33
2.2 Review of Random Variables	49
2.3 Random Sequences	54

2.4	Power Spectral Densities	60
2.5	Filtering Random Sequences	61
2.6	Guide to Further Study	63
3	Fundamentals of Multidimensional Signal Processing	67
3.1	Signals and LSI Systems in 2-D	68
3.2	The Discrete-Time Fourier Transform	81
3.3	The 2-D z -Transform	86
3.4	Filter Design	97
3.5	2-D Sampling	99
3.6	Guide to Further Study	103
4	The Human Visual System	107
4.1	Light and Vision	108
4.2	Brightness and Contrast	115
4.3	Luminance and Color Perception	120
4.4	Spatial Frequency Characteristics of the HVS	125
4.5	Temporal Properties of Vision	129
4.6	Higher Level Processing in the HVS	130
4.7	Guide to Further Study	135
5	Transforms, Filter Banks, and Wavelets	137

CONTENTS

CONTENTS

v

5.1	Block Transforms for Images	140
5.2	The 2-D DFT	148
5.3	Image Decomposition Using Filter Banks	174
5.4	Two-Band Filter Banks	184
5.5	Wavelets and Image Processing	206
5.6	Guide to Further Study	211
6	Digital Image Compression	223
6.1	Basic Concept of Compression	224
6.2	Entropy Coding Methods	230
6.3	Quantization	240
6.4	Fundamentals of Vector Quantization	248
6.5	Differential Coding Methods	259
6.6	Transform Coders and the JPEG Standard	265
6.7	Subband Image Coding	287
6.8	Fractal Image Coding	302
6.9	Lossless Image Compression	316
6.10	Guide to Further Study	324
7	Motion Representation and Video Compression	329
7.1	Motion Representation	330
7.2	Block-Based Motion Representations	334

7.3	Video Coding	347
7.4	DPCM-Based Video Coding	348
7.5	Elements of the Video Coding Standards	352
7.6	Recommendation H.261	356
7.7	The MPEG Standards	359
7.8	Subband Video Coding	370
7.9	Guide to Further Study	375
8	Image Enhancement	379
8.1	Pixel Amplitude Modification Techniques	380
8.2	Direct Filtering Methods	391
8.3	Morphological Filters	403
8.4	Analysis-Modification-Synthesis (AMS) Techniques	409
8.5	Multiplicative Homomorphic Filtering	413
8.6	Interpolation and Zooming	416
8.7	Guide to Further Study	422
9	Image Restoration	425
9.1	Problem Statement and Inverse Solution	427
9.2	Least Mean-Square-Error Restoration	432
9.3	FIR MSE Restoration	442
9.4	Least-Squares Restoration	443

CONTENTS

CONTENTS

vii

347
348
352
356
359
370
375
379
380
391
403
409
413
416
422
425
427
432
442
443

9.5	Iterative Restoration	451
9.6	Maximum Entropy Restoration	458
9.7	Extrapolation and Superresolution	461
9.8	Guide to Further Study	467
10	Image Analysis	471
10.1	Segmentation	474
10.2	Representations and Features	483
10.3	Classical Statistical Classifiers	500
10.4	Template Matching and Correlation Methods	507
10.5	Neural Network Classifiers	510
10.6	Object Detection Methods	519
10.7	Guide to Further Study	536
11	Practical Issues in Image Processing	543
11.1	Image Sources	545
11.2	Image Storage	555
11.3	Long-Term Storage of Images	556
11.4	Image Processors	557
11.5	Image Processing Software	562
11.6	Image Display	567
11.7	Communications	573

11.8 Price List 575

11.9 Guide to Further Study 575

INDEX

587

CONTENTS

575

575

587

Preface

Image processing now is almost a household word. Public awareness of this subject area has been bolstered considerably in recent years by multimedia products that have appeared on the market, by advances in telecommunications such as fax and teleconferencing, and by the media attention given to special effects in movies, just to mention a few. One only has to read the newspaper or watch television to sense that the world is going high-tech, and images and image processing are prominent parts of this movement.

This book is devoted exclusively to the exciting and rapidly developing field of digital image processing and is intended as a text for a graduate-level course. Moreover, it was written to be an integrated part of the course structure at the Georgia Institute of Technology. As part of the educational program in the School of Electrical and Computer Engineering's Center for Signal and Image Processing (CSIP), a wide variety of courses are offered in signal processing. Basic digital signal processing as taught from the text "Discrete-time Signal Processing" by Oppenheim and Schaffer, is the base course for all others in the program. In light of the sequential course structure, this text was written specifically to provide a treatment of image processing with notation and a perspective that is consistent with the introductory DSP courses in CSIP.

Great attention has been given to the composition of the chapters, treatment of subject matter, and organization of the book. Many illustrative examples are provided to make concepts easier to understand at first glance.

The authors of this text firmly believe in the importance of computer projects as an educational instrument. The CSIP image processing course for which this book was written has a strong emphasis on computer projects. Hence, many

references are made throughout to external printed and electronic sources for detailed information on various aspects of image processing. For this reason, we view this text as a *guide* to image processing rather than a traditional reference source.

Many image processing topics have been included in this volume, some of which are traditional, and others that have just recently become popular. The text begins with a classical introduction of the general area, and a review of some of the fundamentals with which you should feel familiar, such as random processes and matrix algebra. Afterward basic signal processing theory is extended to multiple dimensions.

The human visual system is considered next, where we examine the special properties of how we perceive what we see. These properties are presented and explained in terms of frequency domain concepts. Transforms such as the DFT and discrete cosine transform are examined in the following chapter. They provide a link between the spatial domain and the frequency domain, as well as a wide variety of other domains. The remainder of the book deals with applications, among which are image and video compression, enhancement, restoration, segmentation, machine vision, and image analysis.

As authors, we hope that you will have an enjoyable experience with this textbook and welcome your feedback.

mjts@ece.gatech.edu

alen@ece.gatech.edu

Personal Acknowledgments

As with any writing project of substantial magnitude, there are many individuals who directly and indirectly contribute to the effort. I am indebted to our colleagues in the Center for Signal and Image Processing—in particular, Professors Thomas P. Barnwell, Russell M. Mersereau, Ronald W. Shafer, Monson H. Hayes, Mark C. Clements, James H. McClellan, Petros Maragos, Douglas Williams, Vijay Madisetti, and Guotong Zhou. The impact of their friendship and professional association is immeasurable.

Our activities over the years in the general area of image processing have benefited from interactions with many outstanding colleagues and friends. We are indebted to Professors Tor Ramstad, P. P. Vaidyanathan, Kambiz Nayebi, Martin Vetterli, Robert Bamberger, Romain Murenzi, Jan Biemond, Rashid Ansari, Faouzi Kossentini, Bernd Girod, and Ali Akansu, and Drs. Steven Eddins, Marvin Cohen, Wilson Chung, Iraj Sodagar, Chris Barnes, and Henrique Malvar.

In addition, I am grateful for our recent collaborations with Carlos Handy, Wayne Daley, Tim Brosnan, Doug Brittan, Karim Bouyoucef, and Jean-Pierre Leduc. Very special words of thanks go to Jin-Woo Nahm, Quoc H. Pham, Sang-II Park, Hong Man, Tami Randolph, Fernando Mujica, Halûk Aydinoglu, Jose Gonzalez, and Joe Monaco, all of whom contributed examples or experimental results. I deeply appreciate the help given to us by Joseph Arrowood, Mohammad A. Khan, Charlena Williams, Robbie Armitano, Gwangcheol Park, and Ms. Dorothy C. Smith in the preparation of figures and in typing.

This project could not have been completed without the steadfast love and support of my wife Cynthia and children Stephen, Kevin, and Jennifer. Finally and most important, I thank the Lord Jesus Christ for providing the strength, inspiration, opportunity, and guidance through all phases of this project.

Mark J. T. Smith

A Study Guide for Digital Image Processing

Mark J. T. Smith ■ Alen Docef

Georgia Institute of Technology



Center for Signal & Image Processing

of Two-Dimensional Digital
Acoust. Speech and Signal

Two-Dimensional Linear Shift-
invariant Systems, CAS-24,

Represented by the Expan-
sion, Edinburgh, Section A

ence of Noise," Proc. IRE 37

Chapter 4

The Human Visual System

There are many facets to image processing as discussed in Chapter 1. Frequently, the output of the processing is an image that will be observed by a human. In this case, the true measure of image processing quality is how well the image appears to the observer. In other cases, the output may be a decision regarding whether an object of interest is contained in the picture or whether some objects in the picture are normal or defective. Humans happen to be very good at recognizing objects and making such decisions. Given these motivating factors, we devote this chapter to the physiological and cognitive properties of vision that are most relevant to image processing. The human visual system, which we call HVS for short, is very complex and is not understood well in a complete sense. However, certain aspects of its functioning can be modeled reliably and many of its properties can be identified and used to advantage. The HVS can provide clues for improving the subjective quality of image processing systems. Moreover, it is possible that clues to improving object recognition and image interpretation may also be gleaned from a better understanding of the HVS. For convenience, one can partition the HVS into two broad levels of bio-neural processing: peripheral level processing, and central level processing. Most of the useful statements we can make involve peripheral level functions and include optics, perception of brightness and color, frequency response characteristics to visual stimuli, and optical masking effects. At the central level, where interpretations are made in the brain, observations can be made but quantitative characterizations are much more difficult to make. Thus, our focus is on the former, i.e. the peripheral level, and we begin with

the rudimentary principles of light and vision.

4.1 Light and Vision

The process by which we *see* involves light from the sun or some other source reflecting off of objects into our eyes. The illumination source may be characterized by its *intensity* and *wavelength*. The visible range of the electromagnetic spectrum starts at a wavelength, λ , of about 350 nanometers (nm) and ends at about 780 nm. It is the general convention to work with wavelength instead of frequency. Frequency f (in Hertz) and wavelength (in meters) are related by

$$\lambda = c/f,$$

where $c \approx 3 \times 10^8$ meters/sec, which is the speed of light in air. Figure 4.1 shows the visible range of the electromagnetic (EM) spectrum in relation to other EM waves. Above the visible range are infrared, microwaves, and radio waves. Below it are ultraviolet, x-rays, gamma rays, and cosmic rays.

As observers, we view reflected light when we look at a scene or object. This process can be represented in terms of illumination and reflectance components. In particular, we denote $i(t_1, t_2)$ as the illumination source. The illumination source is a continuous spatial intensity with continuous horizontal and vertical spatial coordinates, t_1 and t_2 . Most illumination sources of interest provide fairly constant illumination over the spatial variables, t_1 and t_2 . When this light strikes an object or scene, it is reflected. The amount of light reflected will generally vary spatially, which defines the images we see. The reflectance may be represented by the function, $r(t_1, t_2)$. Thus, the image, $x(t_1, t_2)$, we see is

$$x(t_1, t_2) = i(t_1, t_2)r(t_1, t_2).$$

There are some obvious physical constraints on the reflectance and illumination. The reflectance can never be greater than unity, otherwise it would be amplifying the illumination, which physically is impossible. In addition, the reflectance may not be zero, as this would correspond to 100 percent absorption. The illumination is also bounded. It obviously cannot be negative and practically it must be less than infinity. Thus,

$$0 < r(t_1, t_2) < 1 \quad \text{and} \quad 0 \leq i(t_1, t_2) < \infty.$$

the sun or some other source of light in air. Figure 4.1 shows the range of the electromagnetic spectrum in relation to infrared, microwaves, and radio waves, and cosmic rays.

Figure 4.1 shows the range of the electromagnetic spectrum in relation to infrared, microwaves, and radio waves, and cosmic rays.

When this light strikes a scene or object. This light has two components: an illumination component and a reflectance component. The illumination component is the light that comes from the source. The reflectance component is the light that is reflected by the object. The amount of light reflected by the object is a function of the wavelength of the light and the reflectance of the object. The reflectance of the object is a function of the wavelength of the light and the reflectance of the object.

In addition, the amount of light reflected by the object is a function of the wavelength of the light and the reflectance of the object. The reflectance of the object is a function of the wavelength of the light and the reflectance of the object.

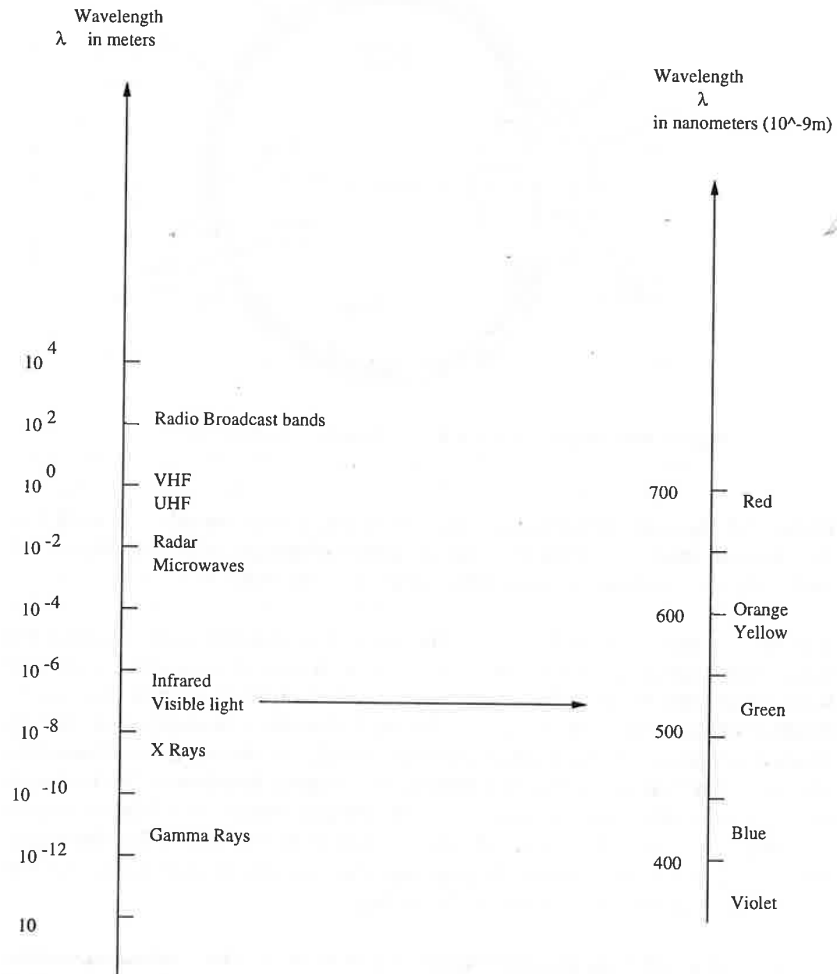


Figure 4.1: Electromagnetic spectrum as a function of wavelength.

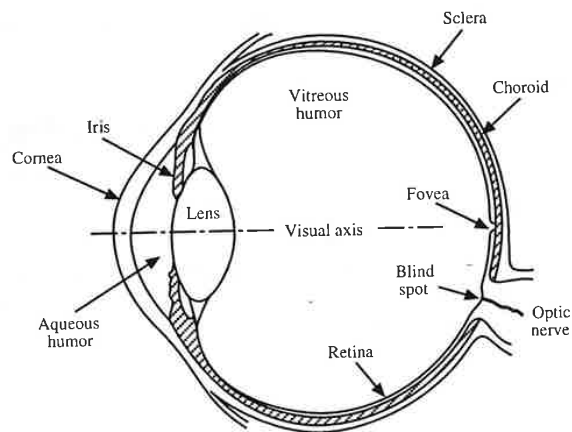
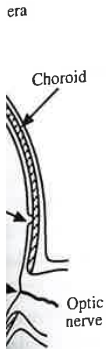


Figure 4.2: Diagram of a horizontal cross-sectional view of the eye.

Before delving into the phenomenology of vision, a brief overview is needed on the physical structure of the eye. As an aid for discussion, examine Figure 4.2, which shows a horizontal cross-sectional view of the eye.

The average eye is about 20 mm in diameter. The exposed outer covering is a tough transparent membrane called the *cornea*. It provides the outer protection to lens and refracts the light entering the eye. In fact, most of the refraction (or bending of the light) that occurs in the eye takes place in the cornea. As can be seen in Figure 4.3, light enters the eye through the cornea and is focused by the lens. The lens in the eye is a flexible transparent membrane. Its bi-convex shape is controlled by the muscles in the ciliary body. To focus on objects far away, the lens is flattened. While for objects in the near field, the lens is conformed to become thicker. In this way, the lens and its controlling muscles provide the capability for fine-grain focusing.

Light entering the optic globe is controlled by the pupil, which is a small opening in the iris. The pupil can contract to about two millimeters and dilate to about eight millimeters on average. The pupil provides control over the amount of light that enters the optic globe. The lens focuses the incoming light, which is entering the pupil, and images it on the retina located in the back of the eye. This general imaging process is illustrated in the diagram shown in Figure 4.3. As shown, the object in view is inverted as it passes through the focal



view of the eye.

f overview is needed on
on, examine Figure 4.2,
e.

posed outer covering is a
vides the outer protection
most of the refraction (or
ce in the cornea. As can
cornea and is focused by
a membrane. Its bi-convex
body. To focus on objects
in the near field, the lens is
and its controlling muscles

which is a small opening
and dilate to about
the amount of
light, which
back of the
shown in Figure
the focal

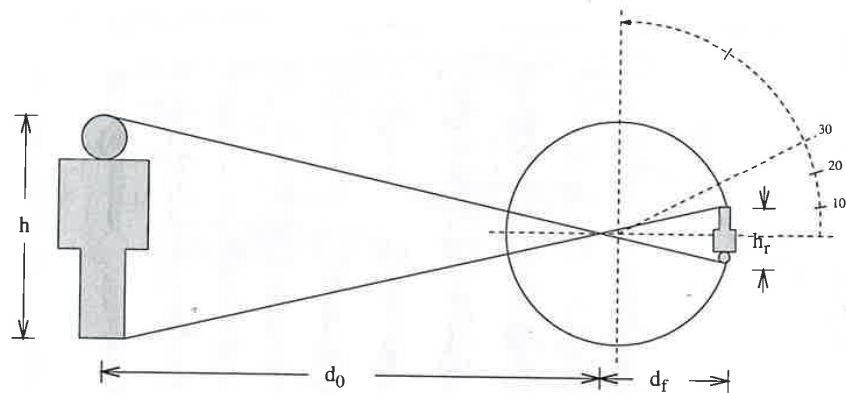


Figure 4.3: Illustration of how an object is imaged on the retina of the eye.

point of the lens and on to the retina. The geometry of the imaging process is very simple but very important to our understanding of vision. The distance between the focal center of the lens and the retina, which is denoted by d_f in the figure, varies between 14 and 17 mm for the average person. At the point of lowest refraction, d_f is about 17 mm, while at its highest refraction it is about 14 mm. Objects being viewed more than three meters away engage the lens at its lowest refractive power, the opposite being true for objects in the very near field.

With this simple understanding of the geometry of imaging, it is straightforward to calculate the size of an object on the retina. From Figure 4.3, we see the key parameters are: d_o the distance from the object to the focal point; d_f the distance from the focal point to the retina; h the height of the object being viewed; and h_r the height of the resulting retinal image. Given the geometry, the retinal image size is given by

$$h_r = \frac{d_f h}{d_o}.$$

It is evident that viewing distance is inversely proportional to the retinal image size.

Perhaps less obvious is the implication of retinal image size on perception, and how light is converted into the picture interpreted by the brain. The retina contains two major types of photoreceptors, *rods* and *cones*. A diagram is

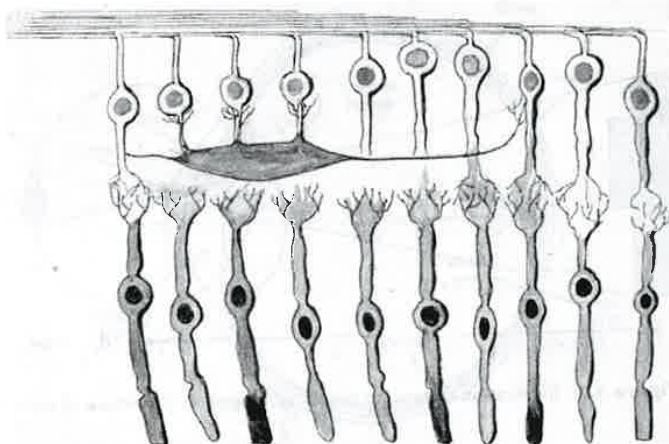
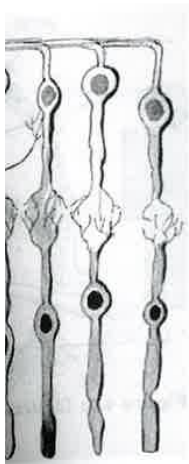


Figure 4.4: Diagram of photoreceptors, cells, and nerves in the retina.

given in Figure 4.4.

The rods are the most numerous of the two. Typically, there are about 100 million in each eye. These receptors are long, thin, and very sensitive. Several are typically connected to a nerve. Rods are responsible for what we call *scotopic* vision or low-light vision. Cones are the other type of photoreceptor. They are shorter, thicker, less sensitive, and fewer in number. As shown both rods and cones are connected to bipolar and ganglion cells along the way. Cones are responsible for our bright light or *photopic* vision. Cones also allow us to perceive color, a point upon which we will expand shortly.

In between scotopic and photopic vision is mesopic vision. This occurs in medium light and involves the use of both rods and cones simultaneously. Some insight into vision can be gained by examining the distribution of photoreceptors in the retina. For convenience, we reference circumferential position in the eye on a scale of 360 degrees, with zero being located directly in back of the eye as labeled in Figure 4.3. Spreading out the circular arc to a straight line, we obtain the axis shown in Figure 4.5. This figure shows the distribution of rods and cones as a function of degree of arc in the eye. Notice first that the cones are concentrated in the back of the eye (centered at zero degrees). This dense region of photoreceptors is called the *fovea*. The rods are distributed over a



nerves in the retina.

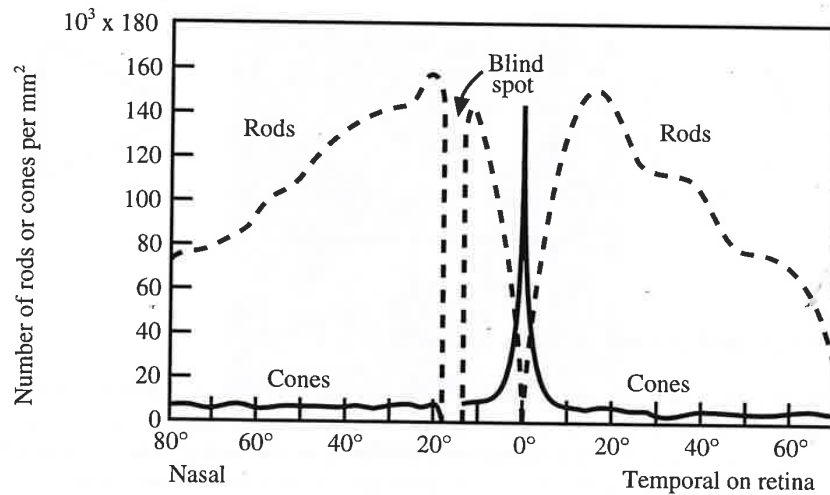


Figure 4.5: Plot of the density of rods and cones in the retina as a function of angle in degrees.

much wider region of the retina. You will also observe that slightly off axis at about 15-20 degrees, there are no photoreceptors. This is the location where the bundle of nerves from the photoreceptors collect and go to the brain and is the so-called *blind spot* in the retina. The bundle of nerves at the blind spot when viewed collectively are often called the *optic nerve*. The position of the optic nerve in the eye is shown explicitly in Figure 4.2. Light that is imaged on the optic nerve cannot be seen, because there are no photoreceptors at that location. We can experience the presence of the optic nerve by performing a simple experiment. Place the textbook about two feet in front of you and look at Figure 4.6.

Cover your left eye (i.e. right eye open) and focus on the circle located on the left. Even though you are focused on the circle, you should still see the square next to it in your peripheral vision field. Now move the book slowly toward your face while continuing to focus on the circle. At some point during the movement, the square will seem to disappear. This is because the square object imaged on your retina is passing over your optic nerve. Using the simple geometry-based analysis depicted in Figure 4.3, one can even calculate the viewing distance at



Figure 4.6: Optic nerve experiment.

which the object will no longer be visible.

With two eyes working together, we are generally not aware we have a blind spot. The blind spot, which is the size of a pen tip, is critical for sight as it is the inlet for the visual information pathway to the brain. After this point, the optic nerves pass through the geniculate bodies in both hemispheres of the brain and on to the visual cortex, which is located in the back of the head. Very complex processing occurs in the visual cortex. Unfortunately, not much is understood about these operations.

We do have some understanding about how light is received and converted within the eye to neural signals, and are able to characterize this behavior with simple models. The light intensity that we see, denoted $x(t_1, t_2)$, varies spatially in t_1 and t_2 , and is also a function of wavelength λ . Thus, we could write this intensity as an explicit function of wavelength, $x(t_1, t_2, \lambda)$. The integral of this *radiometric* quantity with respect to λ is called the irradiance or radiant flux per area. The irradiance,

$$\int x(t_1, t_2, \lambda) d\lambda,$$

is an energy per area per time, and may have units of watts per square meter, or equivalently joules/(meter)²-sec.

What we see physically is $x(t_1, t_2, \lambda)$. What we perceive, however, is generally an image described in terms of *brightness*, *hue*, and *saturation*. Brightness refers to the intensity we perceive. Hue refers to the color, and saturation to the vividness of the color. These are subjective terms and are influenced in a non-trivial way by the viewing environment and past exposure to visual stimuli. In the next sections, we look at these perceptual quantities in further detail.

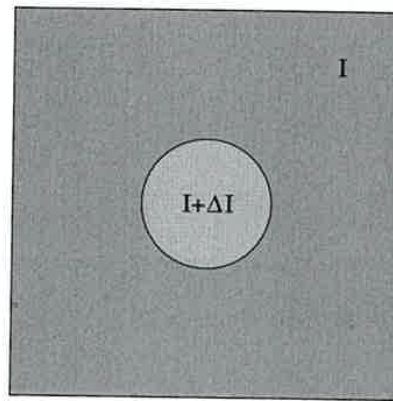


Figure 4.7: Weber's ratio experiment.

4.2 Brightness and Contrast

In addition to the HVS sensitivity, the HVS has a unique characteristic with respect to brightness and intensity. *Brightness* refers to how intense we *think* a visual stimulus is, which is not the same as its measured intensity or irradiance. For example, a visual stimulus that is doubled in intensity will not appear to be twice as bright. This is because intensity is not proportional to brightness. A functional relationship between brightness and intensity was derived from a set of experiments, commonly attributed to Weber. In particular, we can consider a set of perceptual experiments in which measured input and perceived human responses are recorded. The experimental setup is simple and consists of having a human observer view a spot on a constant background as depicted in Figure 4.7. The observer is given gain control over the differential intensity, ΔI , of the circular spot. The background intensity, I , is fixed initially to a starting value at the lowest end of the visible intensity range. The observer sets the value of ΔI initially to zero and increases it slowly. The observer is asked to indicate when the spot on the constant background becomes visible. This point is called the *just noticeable difference* because it is the differential value at which the contrast between two different intensities is just noticeable. This value of ΔI is then recorded. The procedure is repeated for different values of I over the full range.

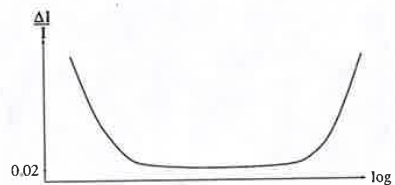


Figure 4.8: Plot of Weber's ratio experiment result.

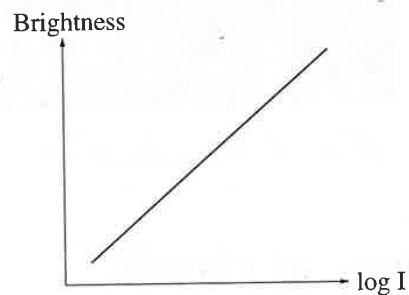


Figure 4.9: Brightness vs. intensity characteristic for the HVS.

What can be observed is that for each value of I in the intensity range, the ratio of ΔI to I is roughly a constant. In particular,

$$\frac{\Delta I}{I} \approx 0.02. \quad (4.1)$$

The quantity, $\Delta I/I$, is called the Weber ratio and is plotted in Figure 4.8. It indicates that ΔI increases as I increases. In other words, we are better able to discern differences in intensity when the image is dark than when it is light. The Weber ratio provides the key to the brightness-intensity relationship as shown in Figure 4.9.

The abscissa, I , is plotted on a log scale, while brightness, shown on the ordinate, is linear. The approximate straight-line relationship shown indicates that brightness is approximately logarithmic with intensity. One might also note that the intensity range shown is extremely large. The human visual system is quite remarkable in its ability to adapt to a wide range of light intensities. On the intensity scale of milli-Lamberts, this range covers 10 orders of magnitude.

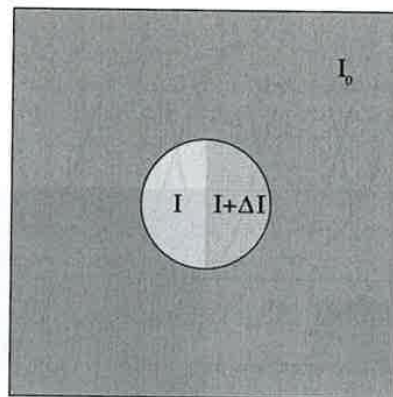


Figure 4.10: Illustration of contrast sensitivity experiment with varying background.

At the bottom-left of the figure, the light intensity is very low. Only the rods contribute to perception in this range. As the light intensity is increased, both rods and cones contribute to perception. In the right half of the intensity range, virtually all contributions to perception are coming from the cones.

It is important to interpret correctly the results of these experiments. The plots in Figures 4.9 and 4.8 do not imply that the HVS operates simultaneously over the full intensity range. The observer in the Weber ratio experiment is allowed as much time as is needed before making a decision on ΔI . During this time, the eye is able to adjust to the background. Normal viewing conditions are significantly different from this situation, however. During normal viewing of an image, we are concerned with being able to discern shading variations between objects where the average background intensity is fixed. This situation is captured by the experiment depicted in Figure 4.10.

During normal viewing, the HVS adapts to the background first and then operates in a smaller range about the mean background level. Similar to the Weber ratio experiment in Figure 4.7, the background here is fixed to a constant I_0 . The spot, however, is divided into two parts, one with intensity I , the other with intensity $I + \Delta I$. The experiment involves a double iteration over three variables. In the inner iteration loop, I_0 is held fixed. I is also fixed at an initial low value. ΔI is then increased starting at zero until the observer detects a difference in contrast between the two circle halves. The ΔI is recorded

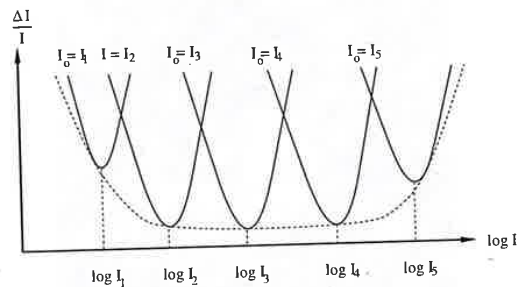


Figure 4.11: Plot of contrast sensitivity experiment with varying background.

as the just noticeable difference. The intensity of I is incrementally increased and the procedure is repeated. This forms the inner loop of the experiment. This same procedure is then repeated over the full range of background intensities, I_0 . The plot of just noticeable differences is shown in Figure 4.11. For a given background illumination, the HVS operates over a modest intensity range. The brightness characteristic remains fairly consistent over a wide range of background illuminations.

For typical monochrome images, humans can resolve about 60 to 100 distinct levels of gray. In fact, images quantized to only 16 levels can appear to be quite natural for this reason. You may recall the close visual similarity between the 8-bit and 4-bit gray-level quantized images shown in Figure 1.5 in Chapter 1, which illustrates this point. In general, we represent monochrome images with 256 levels, as this number of amplitude levels is more than sufficient.

The HVS properties related to brightness are useful in image processing and can be used to improve performance in some cases. For instance, in processing applications where noise or processing artifacts are produced (such as in image compression), subjective quality can be improved by employing strategies that force the noise to occur in the lighter amplitude regions where the HVS is less sensitive, rather than in the darker regions. In addition, these properties suggest that contrast perception can be improved by matching the intensity profile of the image to the brightness sensitivity characteristics of the HVS. This property of brightness perception can be experienced first hand with the aid of a couple of pictures. Look at the set of squares shown in Figure 4.12. Examine the two interior squares in Figure 4.12(a). These squares appear to be of different intensities, but in fact their intensities are the same. The difference

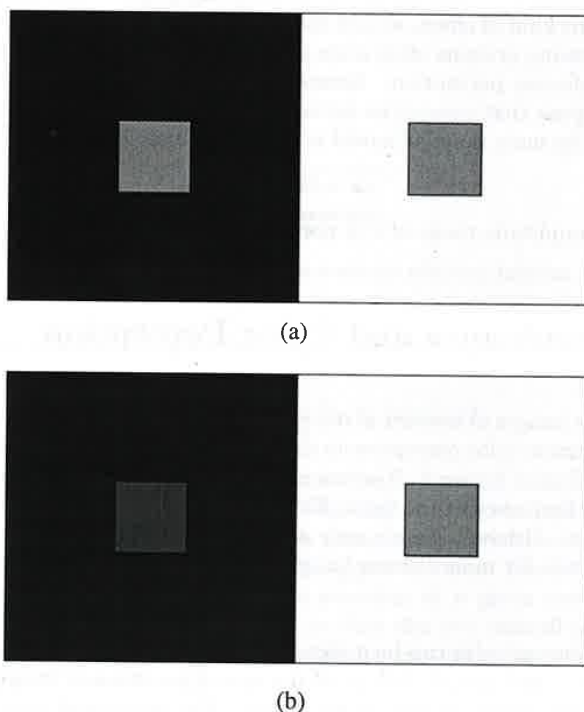


Figure 4.12: Example of brightness perception as a function of background. The interior squares in picture (a) have the same intensity but appear to be different. The interior squares in (b) have different intensities but appear to be the same.

we see is a consequence of the nonlinear brightness characteristic of the HVS. By understanding this property, intensities can be adjusted so that they appear to be the same. This is shown in Figure 4.12(b). These interior squares appear to have the same intensity, but they do not. The brightness characteristics we just discussed can be used to modify the intensity profile of the image to remove or reduce this kind of effect, should such be deemed desirable. More commonly, image processing systems often scale the input intensity image to better match human brightness perception. Several popular contrast models are available for this purpose that convert an intensity image, x , into a contrast normalized image, $\hat{x}$. The most popular model is the logarithmic law. It is given by

$$\hat{x} = 50 \log_{10} x, \quad (4.2)$$

where the amplitude range of x is normalized to $1 \leq x \leq 100$.

4.3 Luminance and Color Perception

Many of the images of interest in image processing are color images. In this section, we examine color perception in the HVS and attempt to understand a little of the bio-physics involved. The color we perceive is related to the electromagnetic (EM) frequency of the light. EM frequencies (i.e. colors) greatly enhance what we see. Although we are only able to distinguishing among about 60 to 100 gray levels for monochrome images, the HVS is able to discern thousands of colors.

Our perception of color can be understood to some extent in terms of the properties of light and the physiology of the eye. The intensity image, $x(t_1, t_2, \lambda)$, enters the eye and is imaged on the retina. The optics and photoreceptors in the eye have varying degrees of sensitivity to the wavelength, as we know. Retinal photoreceptors do not respond to EM wavelengths above or below 350 and 780 nanometers respectively. Not surprisingly, they have non-uniform sensitivity to light of different wavelengths. The average person has maximum visual sensitivity at $\lambda = 555$ nm, which is yellow-green light. This visual sensitivity is captured by a quantity we call *luminance*.

Luminance is a quantity that was adopted by the CIE¹ in 1948 as a measure that reflects the eyes average sensitivity to light as a function of wavelength.

¹The CIE stands for *Commission Internationale de l'Eclairage* and is an international standards body for light and color.

characteristic of the HVS. ed so that they appear interior squares appear ness characteristics we of the image to remove able. More commonly, image to better match t models are available a contrast normalized w. It is given by

(4.2)

≤ 100 .

otion

color images. In this sec-
npt to understand a little
related to the electromag-
e, colors) greatly enhance
shing among about 60 to
able to discern thousands

resent in terms of the prop-
ensity image, $x(t_1, t_2, \lambda)$,
s and photoreceptors in
length, as we know. Reti-
above or below 350 and
low non-uniform sensitiv-
has maximum visual
The visual sensitivity is

is 540 as a measure
of wavelength.
The international

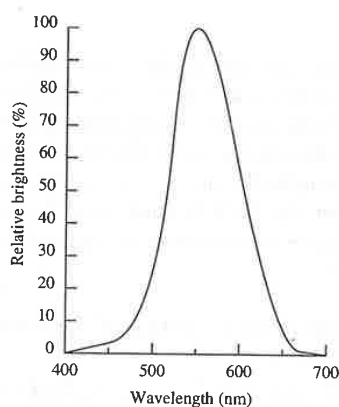


Figure 4.13: A typical relative luminous efficiency function.

The measure which is called the luminance per area is given by

$$\mathcal{L}(t_1, t_2) = \mathcal{K} \int_{\lambda=0}^{\infty} x(t_1, t_2, \lambda) v(\lambda) d\lambda,$$

where $v(t_1, t_2, \lambda)$ is the luminous efficiency function, and $\mathcal{K}$ is a constant. The notion of luminance and the luminous efficiency function is based on the following experiment. A subject is given a monochrome² visual stimulus (i.e. a spot of light) at the reference wavelength $\lambda_{\text{ref}} = 555$ nm. The subject is given gain control of another monochromatic stimulus at a given test wavelength λ_{test} , and is asked to adjust the gain so that the two stimuli appear to be of the same intensity. This test is repeated for the full range of wavelengths, $350 < \lambda_{\text{test}} < 750$ nm. The ratio

$$v(\lambda) = \frac{\lambda_{\text{ref}}}{\lambda_{\text{test}}}$$

is computed, thereby defining the function we call *the luminous efficiency function* for monochrome light.

This function is approximately independent of the value used for λ_{ref} under normal viewing conditions. One should be aware that the characteristics of $v(\lambda)$ vary depending on whether scotopic or photopic conditions are in play.

²Monochrome refers to a light at a single wavelength.

In addition, the function does vary slightly from individual to individual. The CIE defines a standard curve, which was derived from a set of observers, as shown in Figure 4.13. Note that its maximum value is unity and it decreases monotonically in both directions. This profile should be consistent with our expectation since the amplitude given to each $x(\lambda_{\text{test}})$ must be greater than $x(\lambda_{\text{ref}})$. This is because the HVS is most sensitive to the wavelength, λ_{test} . Stimuli at all other frequencies must require higher intensities to be seen at the same perceived intensity.

These measurements lead to the definition of the *luminance per area* given by

$$\mathcal{L}(t_1, t_2) = \mathcal{K} \int_{\lambda=0}^{\infty} x(t_1, t_2, \lambda) v(\lambda) d\lambda, \quad (4.3)$$

where $\mathcal{K} = 685$ lumens/watt, $x(t_1, t_2, \lambda)$ is in watts/m², and $v(\lambda)$ is dimensionless. The units of $\mathcal{L}$ are lumens/m². Luminance brings us one step closer to characterizing our perception of color.

The colored light that we see is generally not monochromatic, that is, it does not typically contain only one EM frequency. Rather, it is a rich composite of many frequencies. The more monochromatic the light is, the more it appears as a pure color to the eye. As the bandwidth is broadened to encompass more frequencies, the color is perceived to be washed out. This property is described as *saturation*. Pure colors are said to be saturated, while broad spectrum light is said to be unsaturated.

The study of color is an interesting topic and has received a fair amount of attention over the years. Thomas Young [3] might be considered the grandfather of *color representation*. He introduced the notion that any color may be decomposed into and represented by three primary colors. Later work focusing on the HVS has established that color is perceived by three different classes of cones, each of which has a different peak absorption characteristic. This is illustrated in Figure 4.14.

The first set of cones peak in the yellow-green range of the spectra at a wavelength of about 450 nanometers (nm). The second set peaks in the green range at about 560 nm, and the third at about 625 nm, which is in the blue range. Color sensation in the HVS can be modeled by

$$\beta_i(C) = \int_{\lambda_{\min}}^{\lambda_{\max}} S_i(\lambda) c(\lambda) d\lambda \quad (4.4)$$

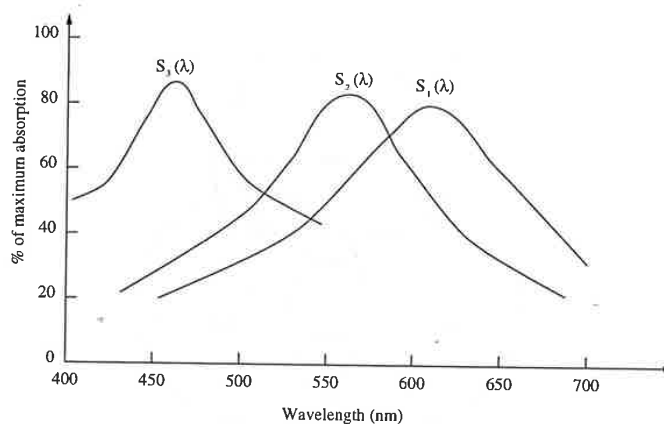


Figure 4.14: Absorption characteristics for the s_1, s_2 , and s_3 cones in the three receptor color model.

where $i = 1, 2, 3$ and $x(t_1, t_2, \lambda)$ is the EM energy distribution of the colored light being observed. The minimum and maximum wavelengths over which the integration is being performed is approximately 380 to 780 nm. The implication of this model is interesting. It says that two colors with different spectral characteristics can be perceived as being the same. Such would occur for two colors, C_1 and C_2 , if

$$\int_{\lambda_{\min}}^{\lambda_{\max}} S_i(\lambda) C_1(\lambda) d\lambda = \int_{\lambda_{\min}}^{\lambda_{\max}} S_i(\lambda) C_2(\lambda) d\lambda.$$

4.3.1 The Additive and Subtractive Color Systems

Red, green, and blue are called the primary colors of light. With these monochromatic colors, a wide variety of colors can be produced by additive combination. The CIE defines red to be monochromatic light with $\lambda = 700$ nm, green to be light at $\lambda = 546.1$ nm, and blue to be light at $\lambda = 435.8$ nm. The additive color system is depicted in Figure 4.15.

If equal portions of red and green are combined, yellow is produced. Equal amounts of red and blue produce magenta, and equal amounts of blue and green produce cyan. Yellow, magenta, and cyan are called the secondary colors

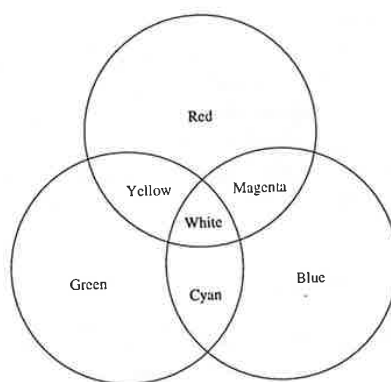


Figure 4.15: Diagram of the additive color system.

in the additive color system. When equal amounts of red, green, and blue light are combined, we get white. A spectrum of colors result from mixing red, green, and blue in different proportions. Here each color component is added, which means that the resulting light is the sum of EM waves of different frequencies. White is the most spectrally rich of colors.

The dual of the additive system is the subtractive color system. It is based on the notion of pigments as opposed to light. When an illumination light source reflects off an object with a certain pigment, some of the light energy is absorbed and only certain frequencies are returned. A particular pigment is defined by the frequencies it absorbs and the level of absorption. The subtractive color system comes into play when we consider mixing paints or inks or any substances off of which light will reflect. As we continue to mix more and more absorbing substances, the reflected light will eventually return no frequencies in the visible range. In this case we see black. Figure 4.16 depicts the subtractive color system. The primary colors in this case are yellow, magenta, and cyan. The secondary colors are red, green, and blue. The equal combination of all primary colors in the subtractive system yields black.

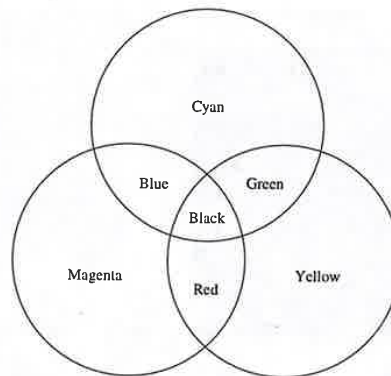


Figure 4.16: Diagram of the subtractive color system.

4.4 Spatial Frequency Characteristics of the HVS

As engineers who have been trained for many years in the application of Fourier principles to natural processes, we may feel compelled to view the HVS in linear shift-invariant terms. We might be compelled to ask, what is the frequency response of the HVS. As it turns out, this is not an unreasonable thing to do. But we must keep in mind that treating the HVS as an LSI system is just an assumption we invoke for convenience and it does break down as we will see in the next section.

As a first step, we make an input/output measurement and then compute the LSI system function associated with the HVS. Let the input be the image in Figure 4.17. Look at this image closely. It is a dark gray rectangle next to a light gray rectangle. Although the shades of gray are constant, you may observe that near the transition boundary it appears as if the shades are more pronounced (the dark shade looks darker, the light shade looks lighter). This is a phenomenon of the HVS and is called the *Mach band effect*. The same effect can be observed in the gray-level bar chart shown in Figure 4.18. Each shade is constant. However, to your eye it may appear that the regions near the transition borders are darker or brighter than they actually are. We can plot the brightness we perceive for both the examples given in Figures 4.17 and 4.18. These are shown in Figure 4.19. Examining Figure 4.19(a), which is the simpler of the two plots, we see a step function with overshoot and undershoot.

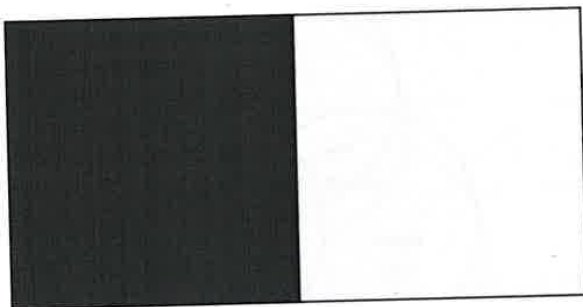


Figure 4.17: A two-tone (black/white) picture.

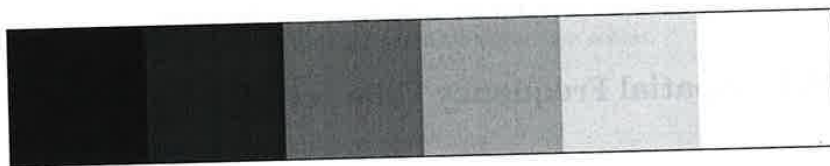


Figure 4.18: A bar chart of gray levels.

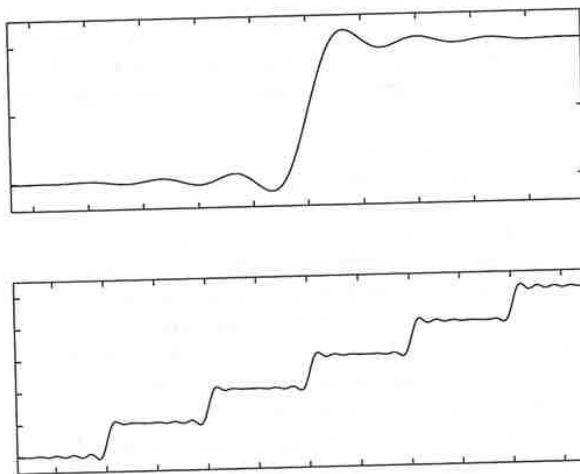


Figure 4.19: Plot of perceived brightness and intensity for examples given in Figures 4.17 and 4.18.

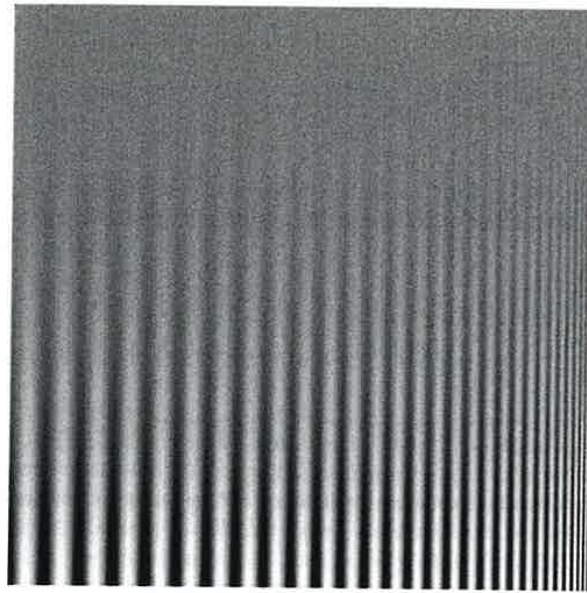


Figure 4.20: Sinusoidal grating with increasing frequency in the abscissa and decreasing amplitude in the ordinate.

The actual input to the HVS is a true step. The overshoot and undershoot is associated with the visual phenomenon known as *lateral inhibition*. From our knowledge of DSP, we know that a lowpass filter will remove the high frequency signal components in a step and result in a new step-like function with overshoot and undershoot. Consequently, we might describe the behavior of the HVS as being a lowpass filter. This simple analysis does not help us pin down the frequency response characteristics of the HVS though. To do this, we can consider another visual experiment involving the picture shown in Figure 4.20. This picture shows a grating with sinusoidal spatial variation. The frequency of the sinusoid is increasing as we move from the left to the right on the abscissa. In the lower region of the picture (on the ordinate axis) the contrast is high. The contrast (and amplitude) are reduced linearly to zero as we move up the ordinate.

Place the picture a few feet from your eyes. This is your viewing distance. Start at the bottom of the picture and move your eye up to the top slowly.

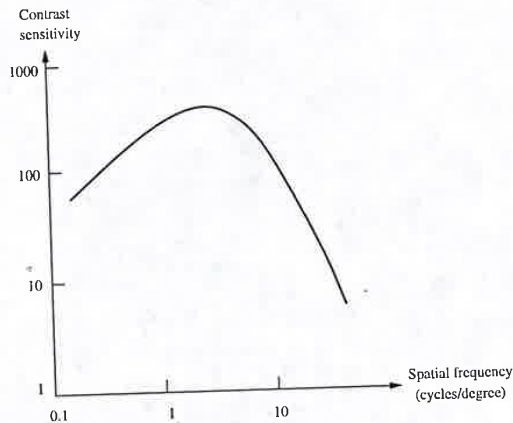


Figure 4.21: Frequency sensitivity of the eye to a sinusoidal grating.

Mark a line (lightly in pencil) where you just begin to lose the ability to see the sinusoidal variation. For most people, that line will be the arched curve shown in Figure 4.21.

Understanding how our eye works is crucial here in interpreting what we see in this picture. As you observe the grating, the picture is imaged onto your retina as illustrated in Figure 4.3. Thus, a range of spatial frequency sinusoids are being observed. The spatial frequencies can be characterized unambiguously by the number of cycles per degree of arc in the retinal plane, i.e. the abscissa in Figure 4.5. Thus, we have the concept of spatial frequency. Some frequencies we can see better than others. The fact that we see the curved profile in Figure 4.21 indicates that our eye is attenuating high and low spatial frequencies. This curve can therefore be interpreted as a frequency response, implying that the eye behaves as a lowpass or more precisely a bandpass filter. The peak response is somewhere between 3 and 10 cycles per degree retinal arc. It will vary somewhat with the viewer. Consequently, we see that this experiment confirms our earlier observation in the Mach band experiment that the eye behaves like a lowpass/bandpass filter. This sinusoidal grating experiment, however, allows us to better characterize the HVS system response. Mannos and Sakrison [2] in particular have derived a formula for the HVS response by fitting a function to the measured data. Their result is a 2-D frequency response, $H(\omega_1, \omega_2)$, which

is expressed in polar coordinate form:

$$H(\omega_1, \omega_2) = H(\rho) = A \left[B + \frac{\rho}{\rho_0} \right] e^{-(\rho/\rho_0)^\beta}, \quad (4.5)$$

where $\rho = \sqrt{\omega_1^2 + \omega_2^2}$ cycles/degree, and A , B , β , and ρ_0 are constants. For their application to image coding [2], $A = 2.6$, $B = 0.0192$, $\rho_0 = 8.772$ and $\beta = 1.1$ was suggested.

4.5 Temporal Properties of Vision

Temporal properties refer to the behavior of the HVS when moving objects and images are considered. Generally speaking, it refers to properties that are related to time changes in visual stimuli as oppose to stationary stimuli. The first property we consider is *Bloch's Law*. Assume that we are looking at a flashing light. If the duration of one flash is very short (say 1/10 second) and the duration of the second flash is very long (say one second), distinguishing the long from the short flash is trivial. The difference in duration is immediately obvious and we are able to tell which is longer and which is shorter. Now assume the individual durations of both are decreased to a few milliseconds and that the total energy in the flashes has been normalized so they are both the same. At some point as these durations decrease, it becomes impossible to distinguish between the two flashes. Bloch's Law states for average illumination conditions, energy normalized flashes are indistinguishable if the durations are less than about 30 milliseconds. This is called the *critical duration*. It changes with average illumination. In low-light illumination situations, the critical duration is even greater.

The next temporal property is the *critical fusion frequency*. This property addresses the perception of a flashing light appearing to be a continuous illumination source. If a light flashes above the critical fusion frequency, it appears as a constant light with the same average intensity as the flashing light. In general, the critical fusion frequency is about 50 to 60 Hz. This is a property that comes into play every time we watch television. The interlaced television picture fields we see³ are flashed at our eyes 60 times per second. Because the

³In North America, the industry standard specifies 60 fields per second. In Japan, the industry standard is 50. This rate is chosen to be the frequency used for electrical power

rate of display occurs at 60 Hz, which is above the critical fusion frequency, we perceive the sequence of image fields as being continuous. Our sensitivity to flicker has a small dependence on spatial frequency. At high spatial frequencies, we tend to be more sensitive to flicker than at lower spatial frequencies.

4.6 Higher Level Processing in the HVS

The notion that the HVS is a linear shift invariant system is useful but limited. In terms of spatial frequency characteristics, it does have behavior that can be modeled as a lowpass/bandpass filter with a characteristic modeled by equation (4.5). However, there are many exceptions we can point to where the LSI assumption is clearly not valid. For balance, we point to several examples where our perception of what we see is clearly non-linear.

To begin, we may recall that brightness perception is logarithmically and not linearly related to intensity. In addition, there are many other interesting oddities. Consider first Figure 4.22, which shows vertical lines with slanted cross bars going through them. To the eye the vertical lines may not appear to be parallel. This is an optical illusion. The perception that the lines are not perfectly parallel results from the slanted cross bars that are located in close proximity. Thus, our perception of parallelism is affected by the surrounding background.

Next examine Figure 4.23. It shows a line drawing of a face. When we see the drawing in this way, we recognize the face as being sad. Now turn the book upside down. The face now appears to be happy. In both cases the figure is exactly the same. Yet a rotation by 180 degrees had the effect of changing our perception from one of *sad* to *happy*. This illustrates that rotation can have a nonlinear effect on our perception. Our interpretation of the mood of the drawn character went from sad to happy.

Now examine Figure 4.24. This figure shows an array of black squares separated by a white border. Focus at the very center of the image. You should observe light gray spots everywhere where the white border lines intersect. The only exception is the center point where you are focusing. Now try to focus on one of the light gray spots. You should observe that the spot disappears when you

transmission to avoid interference. More on the topic of television and video standards is given in Chapter 7.

on frequency, we
our sensitivity to
atial frequencies,
equencies.

IVS

seful but limited.
avior that can be
deled by equation
o where the LSI
cal examples were

hmically and not
r interesting odd-
with slanted cross
ay not appear to
t the lines are not
re located in close
y the surrounding

a. When we see the
Now turn the book
cases the figure is
effect of changing our
rotation can have
of the mood of the

black squares separated
You should observe
The only
focus on one
when you
standards is

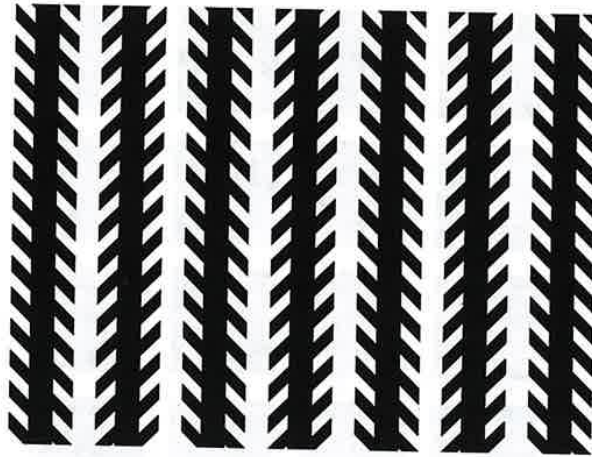


Figure 4.22: Vertical lines with slanted cross bars superimposed.

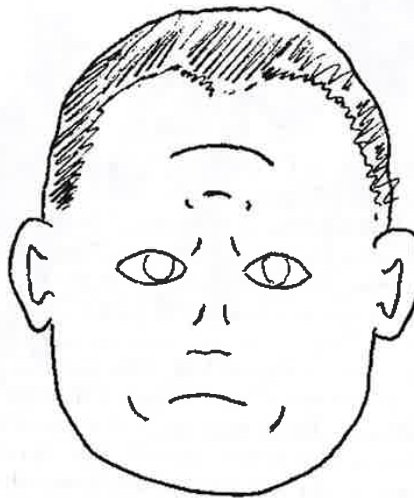


Figure 4.23: The face has a different appearance when turned upside down.

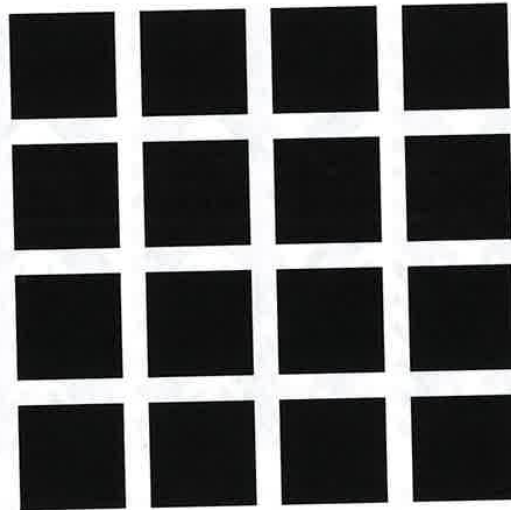


Figure 4.24: Picture of squares to illustrate changing perception with varying point of focus.

try to bring it into your field of focus. This is a dynamic effect which is also nonlinear.

In the next example, look at Figure 4.25. It shows three silhouettes of male figures walking in a room. Which of these men appears to be the tallest. To the eye of most, the leading man on the right of the page will appear taller. In actuality, all three figures are the same height. The perception that the man on the right is taller than the others results from the surrounding background. Thus, the perception of size is a function of background.

As our final example, look at Figure 4.26. This figure shows a triangle with a dot in the interior. To the eye, it should appear that the dot is located closer to the apex of the triangle than it is to the base of the triangle. You may want to measure the distances to convince yourself. But, the distances between the dot and the apex and the dot and the base are identical. This illustrates that our perception of distance is also a function of the background.

In each of the examples above, we have illustrated an optical illusion. These are cases in which the eye perceives an image in an unexpected way. These are



Figure 4.25: Silhouettes of three men in a room. The man on the right appears taller than the others. All are the same height.

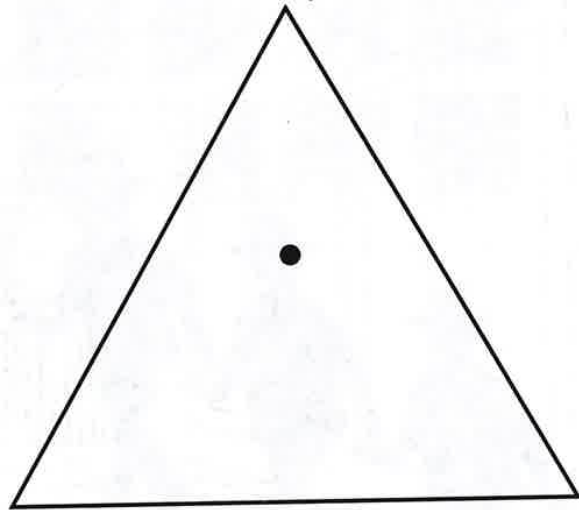


Figure 4.26: The figure shows a dot inside a triangle. The dot is equidistant from the apex of the triangle and the base.

nonlinear effects and are not well modeled by any known convenient relations.

4.7 Guide to Further Study

In this chapter, we have examined human vision and properties of visual perception. In some cases, an understanding of vision and perception can be useful in image processing. For example, addressing perception issues is important (a) in the display of motion video, where temporal fusion properties should be considered, (b) in image compression, where brightness and spatial frequency properties can be used to improve subjective quality, and (c) in image enhancement and restoration where contrast properties and other nonlinear properties can be used to advantage.

Many fine references are available on this subject. We recommend [1] in particular for discussions on perception. Other useful sources for general discussions on vision and perception are [10], [11], [12], [13], and [14].

- [1] T. N. Cornsweet, *Visual Perception*, New York: Academic Press, 1971.
- [2] J. L. Mannos and D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," *IEEE Trans. on Info. Theory*, IT-20, No. 4, July 1974, pp. 525-536.
- [3] T. Young, "On the Theory of Light and Colors," *Philosophical Transactions of the Royal Society of London*, 92, 1802, pp. 20-71.
- [4] J. C. Maxwell, "On the Theory of Three Primary Colours," Lectures delivered in 1961. W.D. Nevin (ed.), *Sci. Papers 1*, Cambridge University Press, London, 1890, pp. 445-450.
- [5] D. L. MacAdam, *Sources of Color Science*, Cambridge, Mass., MIT Press, 1970.
- [6] G. W. Wyzecki and W. S. Stiles, *Color Science*, John Wiley, New York, 1967.
- [7] *Munsell Book of Color*, Munsell Color Co., 2441 North Calvert St., Baltimore, Md.

- [8] H. G. Grassman, *Theory of Compound Colours*," *Philosophic Magazine* 4, No. 7, 1954, pp. 254-264.
- [9] "Colorimetry Proposal for Study of Color Spaces," (technical note), *J. Optical Society of America*, 64, August 1973, pp. 934-939.
- [10] K. R. Castleman, *Digital Image Processing*, Englewood Cliffs, N.J.: Prentice-Hall, 1996.
- [11] M. P. Ekstrom (Ed.), *Digital Image Processing Techniques*, New York: Academic Press, 1984.
- [12] R. C. Gonzalez and P. Wintz, *Digital Image Processing*, 2nd ed. Reading, Mass.: Addison-Wesley, 1987.
- [13] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1986.
- [14] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall, Englewood Cliffs, NJ, 1990.
- [15] W. K. Pratt, *Digital Image Processing*, New York: Wiley and Sons, Second Edition, 1991.
- [16] A. Rosenfeld and A. C. Kak, *Digital Image Processing*, Vols I and II. New York: Academic Press, 1982.

Index

- 2-D DFT 149
- absolute value 500
- additive color system 123
- adjoint matrices 42
- admissibility condition 207
- analog-to-digital converters 2
- analysis problems 19
- analysis-synthesis filter banks 182
- autocorrelation 56, 499
- autocovariance 56
- back-propagation algorithm 515
- Bartlett window 440, 441
- Bayes' risk 504
- Bayesian classifiers 504
- binary images 6
- binary document compression 13
- blind spot 113
- Bloch's law 129
- block transforms 137, 140
- brightness 114, 115, 118
- cameras 2, 545, 549
- CCD cameras 2
- central moments 491
- CFAR 527
- chain codes 486
- CID cameras 2
- circulant matrices 40
- circular matrix convolution 40
- closing 406
- cofactors 42, 43
- color images 6
- color perception 120
- color sensation 122
- compass gradient filters 479
- completeness 144
- compression problem 11
- computer tomography 18
- conditioning number 44
- cones 112
- conjugate quadrature filters 188
- conjugate gradient restoration 453
- conjugate transpose 36, 41
- connectivity 477
- constant false alarm rate 527
- constrained least-squares restoration 447
- contouring 6
- contrast box method 524
- contrast enhancement 14
- contrast models 120
- contrast stretching 389
- cornea 110
- correlation 56
 - properties 58
- covariance 56, 499
 - properties 58
- covariance matrices 56
- critical fusion frequency 129
- critically sampled filter banks 182
- cross-correlation 57
- cross-covariance 56
- cumulative distribution 386

- deconvolution 429
- detection problems 19
- determinants 42
- differentiators 477
- digitized images 4
- dilation 404
- directional decompositions 412
- directional filters 396, 479
- DCT 155
 - basis functions 157
- DFT
 - magnitude 152
 - phase 152
- DHT 173
- discrete Hartley transform 173
- discrete random processes 55
- discrete sine transform 163
- discrete wavelet transform 206
- discrete-time Fourier transform (DTFT)
 - 81
- downsamplers 181
- downsampling matrix 204
- DST 163
- DTFT pairs 83
- DWT 206
- dynamic range adjustment 382
- edge detection 477
- edge enhancement filters 392
- edge sharpening 13
- eigenvalues 44, 46, 172
- eigenvectors 44, 46, 172
- eight-connected 477
- electromagnetic spectrum 108
- energy 500
- enhancement 379
- entropy 500
- ergodicity 58
- erosion 403
- exam 23
- expectation 52
- extrapolation 461
- extrapolation problem 17
- eye 110
- fast DCT 160
- features 483
- filter banks
 - design 188
 - interpretation 139
 - for enhancement 410
- filtering random sequences 61
- FIR restoration filters 442
- first-order histogram 498
- food processing problems 19
- four-connected 477
- Fourier descriptors 494
- fovea 113
- fractal interpolation 417
- frame grabber 9, 545
- frequency response 81
- Gaussians
 - pdfs 52
 - random sequences 60
 - circularly rotated 154
- gradient operators 479
- gray-scale resolution 6
- grayscale closing 408
- grayscale dilation 408
- grayscale erosion 408
- grayscale opening 408
- HVS 107
- Hadamard matrix 165
- Hadamard transform 165
- Hamming window 440, 441
- hanning window 440
- Hermitian symmetric matrices 39, 154
- histogram 474
 - equalization 384
 - features 497
 - modification 380, 384

blem 17

n 139

ient 410

sequences 61

lters 442

am 498

problems 19

'7

rs 494

ion 417

545

se 81

ences 60

tated 154

s 479

ion 6

408

1 408

408

3 408

x 165

form 165

w 440, 441

440

metric matrices 39,

384

1 380, 384

hue 114

human visual system 107

ideal interpolation filter 417

identity matrices 38

IIR restoration filters 441

ill conditioned matrices 42, 44

illumination 108, 415

image processing platforms 9, 543

image enhancement problem 13

image processing systems 2, 543

image storage requirements 9, 555, 556

inertia 500

infinite loop (see recursion)

infrared images 22, 519

inner product 38

intensity 108

interpolation 18, 416

inverse DTFT 83

inverse filter solution 429

involutory matrix 141

irradiance 114

iterative constrained successive approximation 454

iterative restoration 451

joint distributions 50

just noticeable difference 115

K distribution 530

Karhunen-Loeve transform 168

Kronecker product, application 148, 165

k-means algorithm 509

Lagrange multipliers 171

Laplacian filters 481

lateral inhibition 127

least means-square-error restoration 432

least-squares restoration 432

light 108

linear interpolation 417

logarithmic laws 382

lossless compression 11

luminance 120, 122

luminance images 9

M-band filter banks 182

Mach band effect 125

magnetic resonance image 20

Markov random sequences 60

mathematical morphology 403

matrix convolution 39

matrix inversion 42, 43

matrix product 37

matrix-vector transforms 141

maximally decimated filter banks 182

maximum entropy restoration 458

maximum likelihood 502

mean of random sequence 55

mean 52

median filters 398

mid-riser 242

mid-tread 242

minors 43

modulated filter banks 198

moment representations 488

moments of random sequences 55

monochrome images 6

morphological algorithms 531

morphology 403

mother wavelet 207

motion blur 17

mu-law 383

multilayer perceptrons 510

multiplicative homomorphic filtering 413

multivariate distributions 50

neural networks 510

non-separable filter banks 203

object characterization 19

object boundary features 484

object detection 519

- object recognition problem 20
- opening 406
- optic nerve 113
- optical illusions 130, 132
- order statistic CFAR 528
- orthogonal matrices 41
- orthogonal vectors 41
- orthonormality 144
- outer product 37
- Parseval's relation 83
- parts of the eye 110
- pels 2
- photopic vision 112
- photoreceptors 112
- pixelation 4
- pixels 2
- point-spread function 425
- polyphase filters 186
- positive definite matrices 41
- positive semidefinite matrices 41
- power spectral densities 60
- practice DSP exam 23
- prerequisites 22
- probability density functions 50
- probability distributions 50
- pseudo matrix inverses 43
- PSF 425
- QMFs 186
- quad-tree representations 485
- quadrature mirror filters 186
- quantization error 241
- quantizing the gray scale 6
- radiant flux 114
- random processes 54
- random sequences 54
- random variables 49
- rank 43
- rank of a matrix 43
- rank-order filters 398
- raster scanning 4
- receiver operator characteristics 526
- reconstruction 99
- reconstruction from projections 18
- recursion (see infinite loop)
- recursive filter banks 200
- reflectance 108, 415
- resampling matrix 204
- restoration 425
- restoration problem 17
- retina 111
- RGB color 6
- ROC curves 526
- rods 112
- run-length representations 484
- runlength 271
- sampling 99
- SAR images 14, 520
- saturation 114, 122, 243
- scotopic vision 112
- second-order histogram 499
- segmentation 474
- selective averaging 397
- separable transforms 145
- sequency 167
- silhouette matching 509
- single layer perceptrons 510
- singular matrices 42, 44
- singular value decomposition 46, 47, 48
- sliding window 138
- Smith-Barnwell filters 188
- smoothing filters 394
- spatial frequencies 125
- speckle noise 13
- spectral shaping 409
- stationarity 57
- statistical classifiers 500
- steepest descent iterations 451
- strict-sense stationarity 57
- subband filter decompositions 174

- subtractive color system 124
- superresolution 461
- SVD matrix inversion 48
- symmetric matrices 38
- synthetic aperture radar 520
- template matching 507
- temporal properties of vision 129
- thresholding 474
- Toeplitz matrices 39
- trace 38
- triangular window 441
- two-band filter banks 177, 179, 184
- ultrasound images 15
- uncertainty principle 138
- undersampled images 4
- uniform pdfs 52
- unitary matrices 41
- unitary DFT 149
- unitary matrix 141
- upsamplers 181
- variance 52
- variance normalized distance 504
- variance of a random sequence 56
- video compression problem 12
- vidicon tubes 2
- viewing distance 111
- vision 108
- Von Hann window 440
- Walsh Hadamard transform 165
- wavelength 108
- wavelets 206
- Weber's ratio 115
- Weibull distribution 529
- white noise 13
- Wiener filter 437
- wide-sense stationarity 57
- Zernike moments 493
- zero-mean cross-correlation 509
- zero-order interpolation 417
- zonal sampling 175
- zooming 18, 416