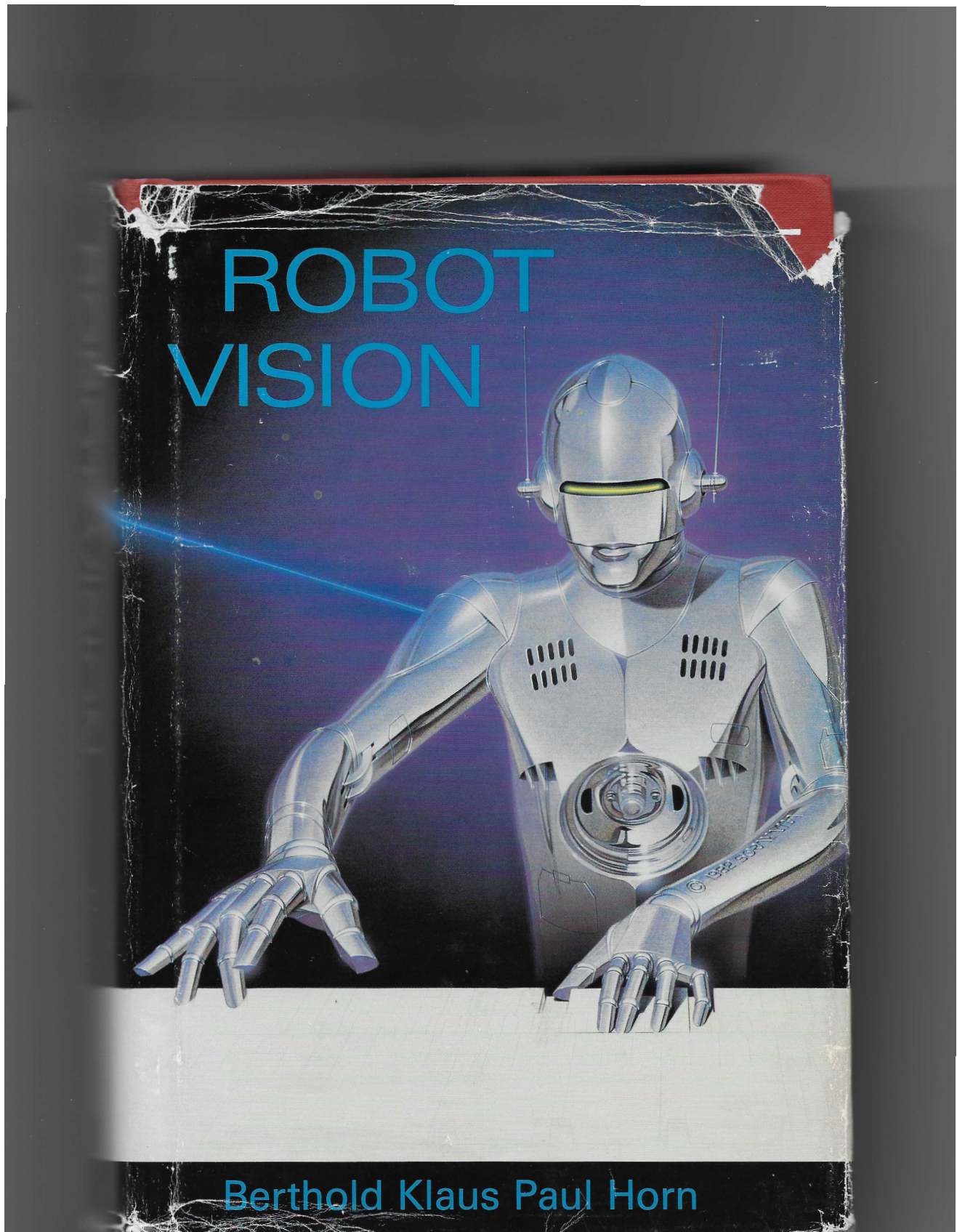




Robot Vision
by Berthold Klaus Paul Horn

"*Robot Vision* presents a coherent development, from image formation to image analysis, to scene analysis. It is a remarkable achievement of the Horn school of vision that it serves both as a personal and a textbook. Every scientist and engineer involved with computational vision should read it, carefully!"—Prof. Alan K. Mackworth, The University of British Columbia

This book presents a coherent development to the fast-moving field of computer vision, using a consistent notation and a detailed understanding of the image formation process. It covers the most recent research and will be a useful and current reference for professionals working in the fields of computer vision, image processing, and pattern recognition.

An outgrowth of the author's work at MIT, *Robot Vision* presents a work for understanding existing machine vision and planning future research. The book includes a great deal of new material that is important to engineers applying machine vision methods in the real world. The chapters on binary image processing, for example, help explain how to improve the performance of commercial devices now available. The book contains material on photometric stereo, extended Gaussian image processing, and what may be the next through the commercialization of the results in

This book is one of a series of texts written by faculty of the Electrical Engineering and Computer Science Department at the Massachusetts Institute of Technology. It was edited and produced by The MIT Press under a joint production-distribution arrangement with the McGraw-Hill Book Company.

Ordering Information:

North America

Text orders should be addressed to the McGraw-Hill Book Company. All other orders should be addressed to The MIT Press.

Outside North America

All orders should be addressed to The MIT Press or its local distributor.

© 1986 by The Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

This book was set under the direction of the author using the T_EX typesetting system and was printed and bound by Halliday Lithograph in the United States of America.

Library of Congress Cataloging in Publication Data

Horn, Berthold Klaus Paul.

Robot vision.

(MIT electrical engineering and computer science series)

Bibliography: p.

Includes index.

1. Robot vision. I. Title. II. Series
TJ211.3.H67 1986 629.8'92 85-18137

ISBN 0-262-08159-8 (MIT Press)

ISBN 0-07-030349-5 (McGraw-Hill)

ated electrons. Older vacuum devices tend to have coatings with relatively low quantum efficiency, while solid-state devices are near ideal for some wavelengths. Photographic film tends to have poor quantum efficiency.

2.5.1 Sensing Color

The sensitivity of a device varies with the wavelength of the incident light. Photons with little energy tend to go right through the material, while very energetic photons may be stopped before they reach the sensitive layer. Each material has its characteristic variation of quantum efficiency with wavelength.

For a small wavelength interval $\delta\lambda$, let the flux of photons with energy equal to or greater than λ , but less than $\lambda + \delta\lambda$, be $b(\lambda)\delta\lambda$. Then the number of electrons liberated is

$$\int_{-\infty}^{\infty} b(\lambda)q(\lambda) d\lambda.$$

If we use sensors with different photosensitive materials, we obtain different images because their spectral sensitivities are different. This can be helpful in distinguishing surfaces that have similar gray-levels when imaged with one sensor, yet give rise to different gray-levels when imaged with a different sensor. Another way to achieve this effect is to use the same sensing material but place filters in front of the camera that selectively absorb different parts of the spectrum. If the transmission of the i^{th} filter is $f_i(\lambda)$, the effective quantum efficiency of the combination of that filter and the sensor is $f_i(\lambda)q(\lambda)$.

How many different filters should we use? The ability to distinguish among materials grows as more images are taken through more filters. The measurements are correlated, however, because most surfaces have a smooth variation of reflectance with wavelength. Typically, little is gained by using very many filters.

The human visual system uses three types of sensors, called *cones*, in daylight conditions. Each of these cone types has a particular spectral sensitivity, one of them peaking in the long wavelength range, one in the middle, and one in the short wavelength range of the visible spectrum, which extends from about 400 nm to about 700 nm. There is considerable overlap between the sensitivity curves. Machine vision systems often also use three images obtained through red, green, and blue filters. It should be pointed out, however, that the results have little to do with human color sensations unless the spectral response curves happen to be linear combinations of the human spectral response curves, as discussed below.

One property of a sensing system with a small number of sensor types having different spectral sensitivities is that many different spectral distri-

butions will produce the same output. The reason is that we do not measure the spectral distributions themselves, but integrals of their product with the spectral sensitivity of particular sensor types. The same applies to biological systems, of course. Colors that appear indistinguishable to a human observer are said to be *metameric*. Useful information about the spectral sensitivities of the human visual system can be gained by systematically exploring metamers. The results of a large number of color-matching experiments performed by many observers have been averaged and used to calculate the so-called *tristimulus* or *standard observer curves*. These have been published by the *Commission Internationale de l'Eclairage* (CIE) and are shown in figure 2-8. A given spectral distribution is evaluated as follows: The spectral distribution is multiplied in turn by each of the three functions $x(\lambda)$, $y(\lambda)$, and $z(\lambda)$. The products are integrated over the visible wavelength range. The three results $\bar{X}$, $\bar{Y}$, and $\bar{Z}$ are called the tristimulus values. Two spectral distributions that result in the same values for these three quantities appear indistinguishable when placed side by side under controlled conditions. (By the way, the spectral distributions used here are expressed in terms of energy per unit wavelength interval, not photon flux.)

The actual spectral response curves of the three types of cones cannot be determined in this way, however. There is some remaining ambiguity. It is known that the tristimulus curves are fixed linear transforms of these spectral response curves. The coefficients of the transformation are not known accurately.

We show in exercise 2-14 that a machine vision system with the same color-matching properties as the human color vision system must have sensitivities that are linear transforms of the human cone response curves. This in turn implies that the sensitivities must be linear transforms of the known standard observer curves. Unfortunately, this rule has rarely been observed when color-sensing systems were designed in the past. (Note that we are not addressing the problem of color sensations; we are only interested in having the machine confuse the same colors as the standard observer.)

2.5.2 Randomness and Noise

It is difficult to make accurate measurements of image brightness. In this section we discuss the corrupting influence of noise on image sensing. In order to do this, we need to discuss random variables and the probability density distribution. We shall also take the opportunity to introduce the concept of convolution in the one-dimensional case. Later, we shall encounter convolution again, applied to two-dimensional images. The reader familiar with these concepts may want to skip this section.

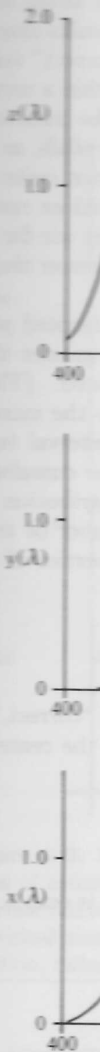


Figure 2-8. The spectral distributions will be each of the functions $x(\lambda)$, $y(\lambda)$, and $z(\lambda)$. In this way we can characterize any spectral distribution by its tristimulus values.

ot measure
duct with
lies to bio-
o a human
e spectral
ematically
atching ex-
d used to
These have
(CIE) and
ted as fol-
the three
the visible
ristimulus
s for these
side under
used here
ot photon
es cannot
ambiguity.
s of these
n are not

the same
have sen-
e curves.
ms of the
rely been
Note that
interested
server.)

In this
sing. In
obability
duce the
hall en-
e reader

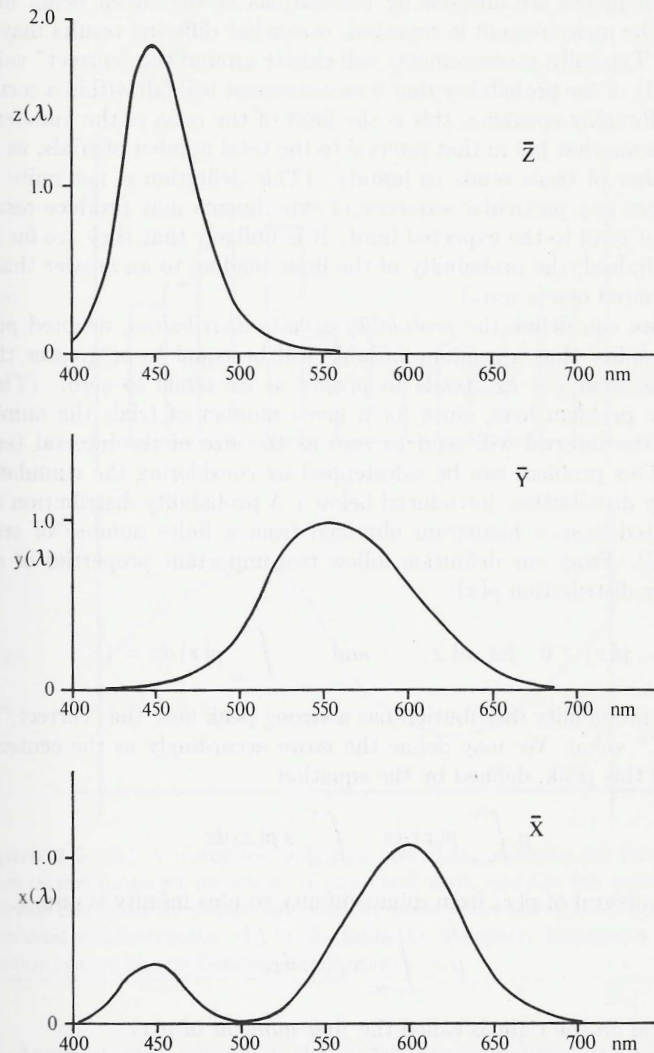


Figure 2-8. The tristimulus curves allow us to predict which spectral distributions will be indistinguishable. A given spectral distribution is multiplied by each of the functions $x(\lambda)$, $y(\lambda)$, and $z(\lambda)$, in turn, and the products integrated. In this way we obtain the tristimulus values, $\bar{X}$, $\bar{Y}$, and $\bar{Z}$, that can be used to characterize the spectral distribution. Spectral distributions that lead to the same tristimulus values appear the same when placed next to one another.