Computer Vision
A Modern Approach

David A. Forsyth
University of California at Berkeley

Jean Ponce
University of Illinois at Urbana-Champaign

An Alan R. Apt Book

Prentice Hall

Prentice Hall
Upper Saddle River, New Jersey 07458

Library of Congress Cataloging-in-Publication Data

CIP data on file.

Vice President and Editorial Director, ECS: *Marcia J. Horton*
Publisher: *Alan Apt*
Associate Editor: *Toni D. Holm*
Editorial Assistant: *Patrick Lindner*
Vice President and Director of Production and Manufacturing, ESM: *David W. Riccardi*
Executive Managing Editor: *Vince O'Brien*
Assistant Managing Editor: *Camille Trentacoste*
Production Editor: *Leslie Galen*
Director of Creative Services: *Paul Belfanti*
Creative Director: *Carole Anson*
Art Director: *Wanda Espana*
Art Editor: *Greg Dulles*
Manufacturing Manager: *Trudy Pisciotti*
Manufacturing Buyer: *Lynda Castillo*
Marketing Manager: *Pamela Shaffer*
Marketing Assistant: *Barrie Reinhold*

About the cover: Image courtesy of Gamma/Superstock

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-085198-1

Pearson Education Ltd., *London*
Pearson Education Australia Pty. Ltd., *Sydney*
Pearson Education Singapore, Pte. Ltd.
Pearson Education North Asia Ltd., *Hong Kong*
Pearson Education Canada, Inc., *Toronto*
Pearson Educacíon de Mexico, S.A. de C.V.
Pearson Education–Japan, *Tokyo*
Pearson Education Malaysia, Pte. Ltd.
Pearson Education, *Upper Saddle River, New Jersey*

different images of a surface in a fixed view illuminated by different sources. This method recovers the height of the surface at points corresponding to each pixel; in computer vision circles, the resulting representation is often known as a *height map*, *depth map*, or *dense depth map*.

Fix the camera and the surface in position and illuminate the surface using a point source that is far away compared with the size of the surface. We adopt a local shading model and assume that there is no ambient illumination (more about this later) so that the radiosity at a point P on the surface is

$$B(P) = \rho(P)\boldsymbol{N}(P) \cdot \boldsymbol{S}_1,$$

where $\boldsymbol{N}$ is the unit surface normal and $\boldsymbol{S}_1$ is the source vector. With our camera model, there is only one point P on the surface for each point (x, y) in the image, and we can write $B(x, y)$ for $B(P)$. Now we assume that the response of the camera is linear in the surface radiosity, so the value of a pixel at (x, y) is

$$\begin{aligned} I(x, y) &= kB(x, y) \\ &= k\rho(x, y)\boldsymbol{N}(x, y) \cdot \boldsymbol{S}_1 \\ &= \boldsymbol{g}(x, y) \cdot \boldsymbol{V}_1, \end{aligned}$$

where k is the constant connecting the camera response to the input radiance, $\boldsymbol{g}(x, y) = \rho(x, y)\boldsymbol{N}(x, y)$, and $\boldsymbol{V}_1 = k\boldsymbol{S}_1$.

In these equations, $\boldsymbol{g}(x, y)$ describes the surface and $\boldsymbol{V}_1$ is a property of the illumination and of the camera. We have a dot product between a vector field $\boldsymbol{g}(x, y)$ and a vector $\boldsymbol{V}_1$, which could be measured; with enough of these dot products, we could reconstruct $\boldsymbol{g}$ and so the surface.

5.4.1 Normal and Albedo from Many Views

Now if we have n sources, for each of which $\boldsymbol{V}_i$ is known, we stack each of these $\boldsymbol{V}_i$ into a known matrix $\mathcal{V}$, where

$$\mathcal{V} = \begin{pmatrix} \boldsymbol{V}_1^T \\ \boldsymbol{V}_2^T \\ \cdots \\ \boldsymbol{V}_n^T \end{pmatrix}.$$

For each image point, we stack the measurements into a vector

$$\boldsymbol{i}(x, y) = \{I_1(x, y), I_2(x, y), \ldots, I_n(x, y)\}^T.$$

Notice that we have one vector per image point; each vector contains all the image brightnesses observed at that point for different sources. Now we have

$$\boldsymbol{i}(x, y) = \mathcal{V}\boldsymbol{g}(x, y),$$

and $\boldsymbol{g}$ is obtained by solving this linear system—or rather, one linear system per point in the image. Typically, $n > 3$ so that a least squares solution is appropriate. This has the advantage that the residual error in the solution provides a check on our measurements.

The difficulty with this approach is that substantial regions of the surface may be in shadow for one or the other light (see Figure 5.10). There is a simple trick that deals with shadows. If there really is no ambient illumination, then we can form a matrix from the image vector and

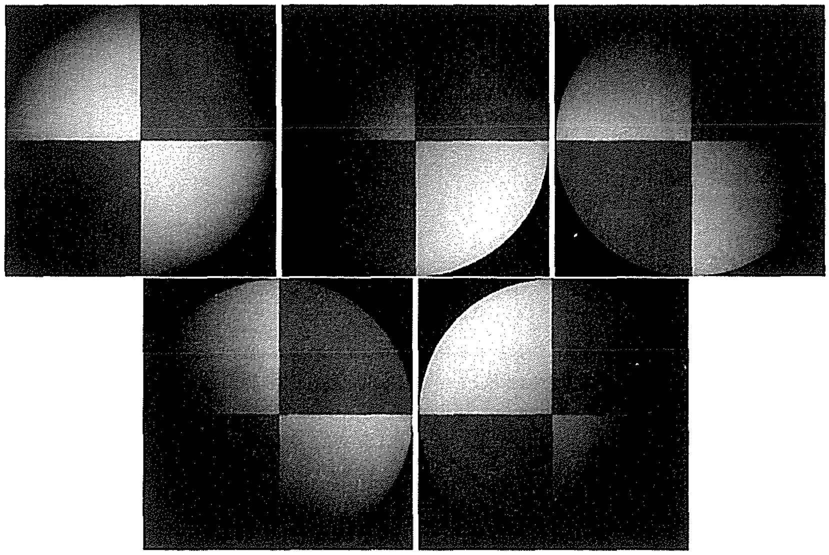


Figure 5.10 Five synthetic images of a sphere, all obtained in an orthographic view from the same viewing position. These images are shaded using a local shading model and a distant point source. This is a convex object, so the only view where there is no visible shadow occurs when the source direction is parallel to the viewing direction. The variations in brightness occuring under different sources code the shape of the surface.

We form

$$\mathcal{I}(x, y) = \begin{pmatrix} I_1(x, y) & \cdots & 0 & 0 \\ 0 & I_2(x, y) & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & I_n(x, y) \end{pmatrix}$$

and

$$\mathcal{I}\boldsymbol{i} = \mathcal{I}\mathcal{V}\boldsymbol{g}(x, y),$$

and $\mathcal{I}$ has the effect of zeroing the contributions from shadowed regions, because the relevant elements of the matrix are zero at points that are in shadow. Again, there is one linear system per point in the image; at each point, we solve this linear system to recover the $\boldsymbol{g}$ vector at that point.

Measuring Albedo We can extract the albedo from a measurement of $\boldsymbol{g}$ because $\boldsymbol{N}$ is the unit normal. This means that $|\boldsymbol{g}(x, y)| = \rho(x, y)$. This provides a check on our measurements as well. Because the albedo is in the range zero to one, any pixels where $|\boldsymbol{g}|$ is greater than one are suspect—either the pixel is not working or $\mathcal{V}$ is incorrect. Figure 5.11 shows albedo recovered using this method for the images of Figure 5.10.

Recovering Normals We can extract the surface normal from $\boldsymbol{g}$ because the normal is a unit vector

$$N(x, y) = \frac{1}{|\boldsymbol{g}(x, y)|}\boldsymbol{g}(x, y).$$

Figure 5.12 shows normal values recovered for the images of Figure 5.10.

Sec. 5.4

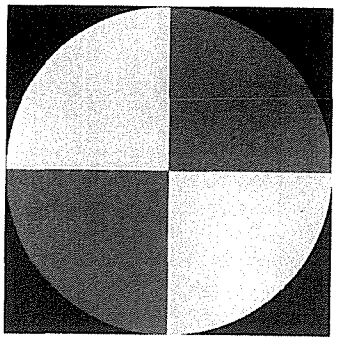


Figure 5.11 The magnitude of the vector field $g(x, y)$ recovered from the input data of Figure 5.10 represented as an image—this is the reflectance of the surface.

5.4.2 Shape from Normals

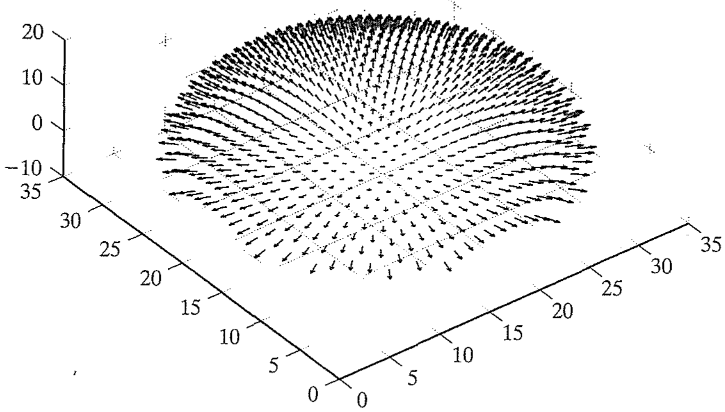
The surface is $(x, y, f(x, y))$, so the normal as a function of (x, y) is

$$N(x, y) = \frac{1}{\sqrt{1 + \frac{\partial f}{\partial x}^2 + \frac{\partial f}{\partial y}^2}} \left\{ -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\}^T$$

To recover the depth map, we need to determine $f(x, y)$ from measured values of the unit normal.

Assume that the measured value of the unit normal at some point (x, y) is $(a(x, y), b(x, y), c(x, y))$. Then

$$\frac{\partial f}{\partial x} = \frac{a(x, y)}{c(x, y)} \quad \text{and} \quad \frac{\partial f}{\partial y} = \frac{b(x, y)}{c(x, y)}.$$



DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase® Smarter legal research.