Depth From Defocus:
A Real Aperture Imaging Approach

Springer Science+Business Media, LLC

Subhasis Chaudhuri
A.N. Rajagopalan

Depth From Defocus: A Real Aperture Imaging Approach

With a Foreword by Alex (Sandy) Pentland

With 58 Illustrations

Springer

Dr. Subhasis Chaudhuri
Dr. A.N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology
Powai, Bombay, 400 076
India

Library of Congress Cataloging-in-Publication Data
Chaudhuri, Subhasis.
Depth from defocus: A real aperture imaging approach / Subhasis Chaudhuri, A.N. Rajagopalan.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4612-7164-2 ISBN 978-1-4612-1490-8 (eBook)
DOI 10.1007/978-1-4612-1490-8
1. Images, Photographic. 2. Depth of field (Photography)
3. Image reconstruction. I. Rajagopalan, A.N. II. Title.
TR222.C49 1998
771—dc21 98-45797

Printed on acid-free paper.

Production managed by MaryAnn Cottone; manufacturing supervised by Joe Quatela.
Camera-ready copy prepared from the authors' LaTeX files.

9 8 7 6 5 4 3 2 1

ISBN 978-1-4612-7164-2

To

Dearest Sucharita and Ushasi.

– SC

Dearest Dad, Mom, Lakshmi, and (late) Shridhar.

– ANR

Foreword

It is now more than fifteen years since I first noticed that defocus in an image could be an important source of depth information. Although this period is brief when compared to the history of stereo or structured light range methods, already depth from defocus has taken its place as an important sensing method. This rapid progress has been because of the intrinsic simplicity of the method: there is no expensive searching for corresponding points in two images, and because there is no correspondence problem there are none of the seemingly-inevitable large errors that come from bad matches.

The main limiting factors for depth from defocus are camera quality, and the sophistication of the signal processing. Progress in cameras has been extremely rapid over the last few years: you can now buy megapixel cameras with an honest 14 bits of intensity resolution. The great accuracy of modern cameras makes possible great accuracy in depth from defocus.

However good cameras alone are not enough. The signal processing that extracts depth must also be extremely sophisticated, and that is the subject of this book. The authors have done a superb job of describing and comparing the various depth-from-defocus methods that have been proposed, an important topic for anyone who has been following this topic. They then move substantially beyond the previous state-of-the-art to develop (and experimentally compare) sophisticated Maximum Likelihood (ML) and Markov Random Field (MRF) methods for depth extraction. In doing so they have made a very significant contribution to the mathematical underpinning of depth-from-defocus, and have in addition made it a much more accurate and practical method.

In summary, I would like to congratulate the authors on their very careful and complete work, and strongly recommend this book to you the reader. I think that you will find reading this book an excellent use of your time.

September 1998

Alex (Sandy) Pentland
Academic Head, The Media Laboratory
Toshiba Professor of Media Arts and Sciences
Massachusetts Institute of Technology

Preface

The recovery of depth from defocused images involves calculating the depth of various points in a scene by modeling the effect that the focal parameters of the camera have on images acquired with a small depth of field. Given two images of a scene recorded with different but known camera parameter settings, the relative blur between the defocused images is measured to obtain an estimate of the depth of the scene. Research on depth from defocus (DFD) was first initiated by Alex Pentland in 1982. Since then, new algorithms have been continuously emerging in the literature for tackling the DFD problem.

The literature on vision abounds with methodologies for extracting depth (or structure) information from various cues in the scene. Among these, motion and stereo disparity are most common and extensively used. Other popular cues include shading and defocus blur. Methods based on these different cues have their own merits and limitations in terms of accuracy, computational complexity, calibration procedures, and associated assumptions. However, one aspect that is very clear is that except for the methods based on measuring the defocus blur, all methods uniformly assume a pin-hole model for the camera, or in other words, the images are not captured with a real aperture camera. This necessitates a longer shutter time for frame grabbing and the captured image may suffer from additional motion blur. The DFD methods use the properties of a real aperture camera to their advantage and provide us with yet another cue to estimate the structure of a scene. This book looks at various ways in which the above cue can be used to recover the depth accurately but refrains from comparing the performance of the DFD technique with methods based on other visual cues.

The area of DFD has witnessed a steady growth of researchers over time. Unfortunately, no single compendium of literature currently exists to facilitate a comprehensive discussion on DFD techniques. This book is an outgrowth of the Ph.D. dissertation work of the author A.N. Rajagopalan under the guidance of the author Subhasis Chaudhuri at the Indian Institute of Technology, Bombay. This book is a revised and extended version of the thesis. A substantial part of this book has appeared (or is due to appear) in many conference proceedings and international journals. Additional theorems and more results have been added in this book for the completeness of the topics discussed. What we emphasize in this book are the concepts that underlie the DFD principle and the recent trends in this area. The book presents several new approaches for computing the depth of a scene when multiple, defocused images of the scene are given. The aim of this book is to show that a rich and interesting theory exists with results that are practically applicable. Since the DFD method is already in use in several industrial applications, we believe that the material herein is both timely and relevant.

Even though the principle of DFD has its origin in physical optics, our analysis is largely from a computer vision/image-processing perspective. The manner of presentation has been designed to make the material suitable to a wide audience ranging from engineering and computer science to mathematics and physics. The text is essentially self-contained and is, for the most part, theoretically complete. The mathematical complexity of the book remains at a level well within the grasp of graduate students and requires only introductory preparation in mathematical analysis, matrix algebra, and probability theory.

In particular, this book will be of great value to computer vision/image processing researchers and practitioners who are interested in the use of the depth data and the restored image of a scene for 3D structure recovery and image interpretation, object recognition, and navigation. The intended audience includes both the novice and the expert: to the novice we provide a lucid introduction, and to the expert a fresh and stimulating perspective to breed new ideas. It is our hope that this book will arouse more curiosity and interest in this exciting and vibrant area, and that it will serve as a stepping stone towards an improvement in our collective understanding of this subject.

A very brief synopsis of the contents of the book is as follows. In Chapter 1, we cover some of the well known passive range finders in brief. The basic theory of real aperture imaging is introduced in Chapter 2. The depth from defocus technique is described, and related available literature on DFD is reviewed. Chapter 3 provides a necessary mathematical background on the time-frequency representations, the calculus of variations, and the Markov random fields (MRF) for use in subsequent chapters. A block shift-variant blur model is described in Chapter 4 to incorporate the interaction of blur among neighboring subimages to obtain improved estimates of the blur, and hence, the depth. In Chapter 5, space-variant filtering models derived from space-frequency representations are shown to successfully recover the depth. In Chapter 6, a maximum likelihood estimator is discussed for recovering the depth from blurred images. Based on the Cramér-Rao bound of the variance of the error in the estimate of blur, an optimality criterion is suggested for selecting the camera parameters. Depth recovery using the maximum likelihood principle and multiple blurred versions of the original image is studied in Chapter 7. A computationally efficient, recursive algorithm is given for computing the likelihood function as more observations are added. The problem of recovering depth given both the focused and the defocused images of a scene is addressed in Chapter 8. The space-variant blur parameter is modeled as an MRF and its maximum *a posteriori* (MAP) estimates are obtained. The problem of simultaneous depth recovery and image restoration is solved in Chapter 9. The space-variant blur parameter and the scene intensity are both modeled as separate MRFs. The book concludes with Chapter 10, wherein the discussions are summarized and suggestions for future work are indicated. A large number of experimental results have been provided in each chapter for a proper appreciation of the methods discussed.

Needless to say, we welcome comments and suggestions from the readers.

Bombay
September 1998

Subhasis Chaudhuri
A.N. Rajagopalan

Acknowledgments

A task like writing a book is never possible without the active help and support of a large number of people.

The authors are greatly indebted to Dr. U.B. Desai at IIT, Bombay, and Dr. P.G. Poonacha at SAS, Bangalore, for their insightful comments and creative ideas at various stages of research over the last five years. Thanks are due to Dr. V.M. Gadre of IIT, Bombay, for his many useful remarks, Dr. S.C. Sahasrabudhe of IIT, Bombay, for his constant encouragement, Dr. Shankar Chatterjee of Cinesite, Hollywood, CA, for his very valuable and thought-provoking suggestions and ideas, and Dr.–Ing. Bernd Girod of the Universität Erlangen-Nürnberg for his critical yet constructive comments. The authors are thankful to Dr. Alex Pentland, Media lab., MIT, for kindly consenting to write the foreword for this book.

The authors thank the members of the SPANN laboratory at IIT, Bombay where the work was conducted. In particular, thanks are due to Sunil Kumar, Mohammed Yeasin, Manivasakan, Jayashree Karlekar, Vinaydutt Kohir, Umesh Kaulgud, and Deepu Rajan for several stimulating and helpful discussions, and to Namita Maiti for proofreading an earlier version of this document.

Both the authors have spent some time at the Lehrstuhl für Nachrichtentechnik, Universität Erlangen-Nürnberg, and the research stay there was quite stimulating and helpful for the completion of this book. The authors are quite thankful to Eckehard Steinbach, Thomas Wiegand, Peter Eisert, Jonathan Su, Gerhard Runze, Klauss Stuhlmüller, Frank Hartung and Alex Stenger for their kind cooperation and valuable comments.

The first author also takes this opportunity to thank Alexander von Humboldt Stiftung, Germany, for their support and financial assistance. Partial financial aid from the Curriculum Development Programme at IIT, Bombay, is also gratefully acknowledged.

The authors are extremely grateful to the IEEE, the Academic Press, Elsevier Science, and the Kluwer Academic publishers for permitting them to reuse their works in this book. Special thanks go to Dr. William R. Sanders, Senior Editor, Computer Science, Springer-Verlag, NY, who encouraged us to publish this manuscript. Thanks are also due to Ms. Mary Ann Cottone and Mr. Frederick Bartlett of the editorial staff at Springer-Verlag, NY, for their kind cooperation and assistance.

Finally, the authors wish to express their deep sense of gratitude to their families for their patience, and their constant love and encouragement without which this book could not have been completed.

Bombay
September 1998

Subhasis Chaudhuri
A.N. Rajagopalan

Contents

List of Symbols

D	Depth of the scene
F_l, F_{l_1}, F_{l_2}	Focal length of camera
v_0, v_1, v_2	Image plane to lens distance
r_b	Blur radius
r_0, r_1, r_2	Lens aperture
$\sigma, \sigma_1, \sigma_2$	Blur parameter
ρ, ρ_1, ρ_2	Constants of camera
$f(\cdot)$	Focused image of the scene
$h(\cdot)$	Point spread function of camera
$g(\cdot), g_1(\cdot), g_2(\cdot)$	Defocused images
$C_x(\cdot, \cdot)$	Complex spectrogram of $x(\cdot)$
$\tilde{W}_x(\cdot, \cdot)$	Pseudo-Wigner distribution of $x(\cdot)$
$\check{s}(i, j)$	$\sigma_2^2(i, j) - \sigma_1^2(i, j)$
λ	Regularization parameter
$a(\cdot)$	AR coefficients
σ_w^2	Variance of observation noise
σ_v^2	Variance of noise in AR process
$\acute{s}$	σ_1^2
η	MRF neighborhood without line fields
$\underline{\eta}$	MRF neighborhood with line fields
$U(\cdot)$	Energy function in MRF model
$V_c(\cdot)$	Clique potential
l^{r^s}	Horizontal line field
v^{r^s}	Vertical line field
μ	Weightage for smoothness in MRF model
γ	Penalty associated with line fields
$s_{i,j}$	$\sigma_{i,j}$