SCIENCES

IMAGE

Sensors and Image Processing

Omnidirectional Vision

From Theory to Applications

Coordinated by

Pascal Vasseur

Fabio Morbidi

iSTE

WILEY

Table of Contents

List of Tables

List of Illustrations

Chapter 4

The editors dedicate this book to their families.
They are grateful for their support throughout the course of this long project.

SCIENCES

Image, Field Director – Laure Blanc-Féraud

Sensors and Image Processing, Subject Head – Cédric Demonceaux

Omnidirectional Vision

From Theory to Applications

Coordinated by

Pascal Vasseur
Fabio Morbidi

iSTE

WILEY

First published 2023 in Great Britain and the United States by ISTE Ltd and John Wiley & Sons, Inc.

ISTE Ltd
27-37 St George's Road
London SW19 4EU
UK

www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA

www.wiley.com

Library of Congress Control Number: 2023938467

British Library Cataloguing-in-Publication Data
A CIP record for this book is available from the British Library
ISBN 978-1-78945-143-6

ERC code:
PE6 Computer Science and Informatics

PE6_2 Computer systems, parallel/distributed systems, sensor networks, embedded systems, cyber-physical systems

PE6_11 Machine learning, statistical data processing and applications using signal processing (e.g. speech, image, video)

Acknowledgments

This book would not have been possible without the support provided by our host institution, the University of Picardie Jules Verne, Amiens, France. Fabio Morbidi is indebted to his colleagues, E. Mouaddib and G. Caron, for the many refreshing discussions about omnidirectional vision over a beer.

The editors express their sincere thanks to the authors who have contributed to the six main chapters of this book, for the professionalism, reactivity, valuable comments and patience, which finally led to the publication of this collective work, after multiple delays due to the Covid-19 pandemic and our busy schedules. We learnt a lot from all of them, and this common journey strengthened our bonds of mutual esteem and friendship.

Finally, the editors wish to extend their heartfelt thanks to the team at ISTE Ltd for the smooth and efficient book production process, and to L. Blanc-Féraud (director of the “Image” field) and C. Demonceaux (subject head for “Sensors and Image Processing”) for their guidance and encouragement.

Pascal VASSEUR and Fabio MORBIDI
Amiens, August 2023

List of Acronyms

AR: Augmented reality

BEV: Bird's eye view

CCD: Charge coupled device

CNN: Convolutional neural network

DNN: Deep neural network

DoF: Degrees of freedom

DoG: Difference of Gaussians

EKF: Extended Kalman filter

FoV: Field of view

GCM: General camera model

GEM: Generalized essential matrix

GPS: Global positioning system

GPU: Graphical processing unit

IMU: Inertial measurement unit

LM: Levenberg–Marquardt

MAP: Maximum a posteriori

MSCKF: Multi-state constraint Kalman filter

PnP: Perspective-*n*-point

RANSAC: RANdom SAmple Consensus

RGB-D: Red, green, blue – depth

RGBA: Red, green, blue, alpha

SfM: Structure-from-motion

SGM: Semi-global matching

SLAM: Simultaneous localization and mapping

SSD: Sum of squared differences

SURF: Speeded up robust features

SVD: Singular value decomposition

UAV: Unmanned aerial vehicle

VR: Virtual reality

WTA: Winner takes all

Preface

Fabio MORBIDI and Pascal VASSEUR
MIS Laboratory, University of Picardie Jules Verne, Amiens, France

P.1. Omnidirectional vision: a historical perspective

“Charge-coupled devices” (CCDs) were invented by W. Boyle and G.E. Smith at Bell Labs in 1969. They consist of a sensor that converts an incoming 2D light pattern into an electrical signal that, in turn, is transformed into an image. Although the CCDs could capture an image, they could not store it. Digital cameras, invented in 1975 at Eastman Kodak in Rochester, NewYork, by S.J. Sasson, fixed this problem. The first digital camera, equipped with a Fairchild Semiconductor’s 100-by-100-pixel CCD, was able to display photos on a TV screen (Goodrich 2022).

An *omnidirectional camera* (also known as 360° camera) is a camera with a field of view (FoV) that covers approximately the entire sphere or at least a full circle in the horizontal plane (the adjective “omnidirectional” combines two words: “omni” and “directional”. “Omni” comes from the Latin word “Omnis”, meaning “all”). A conventional camera has an FoV that ranges from a few degrees to, at most, 180°: this means that it can capture, at most, light falling onto the camera focal point through a hemisphere. On the contrary, an ideal omnidirectional camera captures light from all directions falling onto the focal point, covering a full sphere. However, in practice, most omnidirectional cameras span only part of the full sphere and many cameras, which are dubbed omnidirectional, cover only approximately a hemisphere, or the full 360° along the equator of the sphere, the top and bottom hemispheres excluded (in this case, the term *panoramic camera* is preferred). If the full sphere is covered, the captured light beams do not exactly intersect in a single focal point, i.e. the system is non-central. Human vision is an example of a system with a wide FoV. In fact, humans have slightly over a 210° horizontal FoV (without eye movements) (Strasburger 2020), while some birds and insects have a complete or nearly complete 360° visual field. The vertical range of the visual field in humans is around 150°. A large FoV has proven to be a very important asset in the preservation of certain species, and it likely plays a crucial role in the evolution of animal vision (Burkhardt 2005).

With the same ground being plowed many times by different researchers in the last decades, various camera designs have been proposed to capture 360° images: cameras with a single lens (fisheye), cameras with two lenses (dual or twin fisheye), cameras with more than two lenses (polydioptric), camera rigs, pan-tilt-zoom and cameras with rotating mechanisms, and catadioptric systems combining mirrors (cata-) and lenses (-dioptric). Some of these cameras capture 360° images in a single shot, while the others build an omnidirectional image by stitching together different regions of the FoV acquired over a prolonged period of time. Fisheye cameras (which use lens systems with very short focal lengths and strong refractive power), and catadioptric systems (which were first patented in 1970 (Rees 1970)), are in the first group. On the other hand, pan-tilt-zoom cameras and cameras with rotating mechanisms, belong to the second group.

Although the fundamental concepts have been around since at least the 1970s (Cao et al. 1986; Yagi and Kawato 1990; Ishiguro et al. 1992), *modern omnidirectional vision* dates back to the late 1990s: in fact, the seminal works on image formation and geometry by Nayar (1997); Baker and Nayar (1999); and Svoboda et al. (1998) marked the beginning of an independent field of investigation. Another milestone in the history of omnidirectional vision is the unifying theory for central panoramic systems developed by Geyer and Daniilidis (2000), and subsequently extended by Barreto (2006) and other researchers (Khomutenko et al. 2016; Usenko et al. 2018). This pioneering work has been the harbinger of a burgeoning array of papers in computer vision (image processing and descriptors adapted to spherical signals, calibration, epipolar and multi-view geometry, structure-from-motion, 3D reconstruction, etc.) and robotics (image-based localization, simultaneous localization and mapping (SLAM), visual servoing, etc.). Finally, the series of OMNIVIS workshops ("Omnidirectional Vision, Camera Networks and Non-Classical Cameras") held annually between 2000 and 2011, in conjunction with major computer vision conferences, contributed to shaping the community and bringing together researchers interested in non-conventional vision. This tradition continued with the OmniCV workshops ("Omnidirectional Computer Vision"), organized every year since 2020 (CVPR'23 marked the fourth edition).

Today, with the miniaturization of image sensors and optical components (lenses, prisms and mirrors), omnidirectional cameras have risen to prominence in consumer electronics (smartphone attachments, surveillance systems, perception systems in autonomous vehicles, etc.), and they have transformed our everyday lives and made them easier. Applications include mobile robotics, videoconferencing, art (panoramic photography), real estate (remote tours), vehicle parking assistance, virtual and augmented reality, tele-operated systems (for enhanced situational awareness), forensics, astronomy and entertainment. Several multinational electronics companies

(Samsung, Ricoh, GoPro) have invested in the field of omnidirectional vision which has experienced a renaissance over the past ten years, and they are actively producing and supporting hardware. This fostered academic research and has contributed to the growth of the community.

P.2. Why this book?

While innumerable computer vision books have made their appearance in the last two decades, for example, books by Forsyth and Ponce (2011); Hartley and Zisserman (2004) and Ma et al. (2004) just to mention the most popular ones, relatively few books or monographs have been dedicated to omnidirectional vision. In fact, we are aware of only three research-oriented books (Benosman and Kang 2000; Sturm et al. 2011; Puig and Guerrero 2013), a survey paper (Ishiguro 2005), and two dedicated chapters in robotics textbooks, (Chapter 11.3 in Corke 2011; Chapter 4.2 in Siegwart et al. 2011). Several indicators suggest that the field of omnidirectional vision is now mature: it is then time to review the core principles (image formation, mathematical modeling, camera calibration, etc.), critically assess the key achievements and present some of the main applications, with an eye on the most recent trends and research directions.

Obviously, the field is too vast and dynamic to be fully covered in a single book. Therefore, a precise editorial choice has been made, and some “trendy topics” have been intentionally left out. A notable omission in coverage is the growing body of research on machine learning applied to omnidirectional vision (Ai et al. 2022), which is only briefly mentioned in Chapters 3 and 6. Moreover, we pass over the recent progress made in the field of graph image processing (Cheung et al. 2018). This book brings together the contributions of 10 renowned international scientists with multidisciplinary interests in image processing, computer vision, vehicle engineering and robotics. It is intended for a general audience: young beginners interested in discovering the field, professionals, instructors and experimented scientists in academia.

P.3. Organization of the book

This book consists of a preface, six chapters and a conclusion, and it is organized as follows:

– *Preface*: providing a brief history of omnidirectional vision, it defines the position and scope of the book, and presents its general structure.

– *Chapter 1* reviews basic geometric concepts relevant to omnidirectional vision. These include the image formation process, with a special focus on catadioptric cameras. A brief discussion on how camera models approximate the image formation process is also provided.

– *Chapter 2* presents the geometric models behind the formation of an omnidirectional image and critically assesses the different existing techniques for the estimation of the intrinsic parameters of an omnidirectional camera.

– *Chapter 3* describes different techniques for the reconstruction of 3D environments from images captured by static or moving omnidirectional cameras.

– *Chapter 4* is devoted to image processing, adapted to the spherical signals provided by catadioptric cameras.

– *Chapter 5* presents a special class of omnidirectional cameras, the so-called non-central vision sensors, and provides an overview of their main geometric properties and applications.

– *Chapter 6* deals with the application of omnidirectional cameras to robot localization and navigation.

– *Chapter 7* concludes the book. The main contributions are summarized and some prospects for future research are discussed.

It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment. When I have clarified and exhausted a subject, then I turn away from it, in order to go into darkness again. The never-satisfied man is so strange; if he has completed a structure, then it is not in order to dwell in it peacefully, but in order to begin another. I imagine the world conqueror must feel thus, who, after one kingdom is scarcely conquered, stretches out his arms for others.

Extract from a letter of Carl Friedrich GAUSS to Farkas BOLYAI, dated 2 September 1808.

June 2023

P.4. References

Ai, H., Cao, Z., Zhu, J., Bai, H., Chen, Y., Wang, L. (2022). Deep learning for omnidirectional vision: A survey and new perspectives [Online]. Available at: https://arxiv.org/abs/2205.10468.

Baker, S. and Nayar, S. (1999). A theory of single-viewpoint catadioptric image formation. *Int. J. Comput. Vision*, 35(2), 175–196.

Barreto, J. (2006). A unifying geometric representation for central projection systems. *Comput. Vis. Image Und.*, 103(3), 208–217.

Benosman, R. and Kang, S. (eds) (2000). *Panoramic Vision: Sensors, Theory, and Applications*. Springer-Verlag, New York.

Burkhardt Jr., R.W. (2005). *Patterns of Behavior: Konrad Lorenz, Niko Tinbergen, and the Founding of Ethology*. The University of Chicago Press.

Cao, Z., Oh, S., Hall, E. (1986). Dynamic omnidirectional vision for mobile robots. *J. Robotic Syst.*, 3(1), 5–17.

Cheung, G., Magli, E., Tanaka, Y., Ng, M. (2018). Graph spectral image processing. *Proc. IEEE*, 106(5), 907–930.

Corke, P. (2011). *Robotics, Vision and Control: Fundamental Algorithms in MATLAB*, volume 73. Springer-Verlag, Berlin/Heidelberg.

Forsyth, D. and Ponce, J. (2011). *Computer Vision: A Modern Approach*, 2nd edition. Pearson Education, Upper Saddle River.

Geyer, C. and Daniilidis, K. (2000). A unifying theory for central panoramic systems and practical implications. In *Proc. 6th Eur. Conf. Comput. Vis.* Springer, Berlin/Heidelberg.

Goodrich, J. (2022). The first digital camera was Kodak's biggest secret: The toaster-sized device displayed photos on a TV screen. *The Institute*, 60–61.

Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*, 2nd edition. Cambridge University Press.

Ishiguro, H. (2005). Omnidirectional vision. In *Handbook of Pattern Recognition and Computer Vision*, Chen, C.H.

G, H

I, J, K

L, M

N, O

P, Q, R

S, T, V