RESEARCH

Benjamin Langmann

Wide Area 2D/3D Imaging

Development, Analysis and Applications

Springer Vieweg

Wide Area 2D/3D Imaging

Benjamin Langmann

Wide Area 2D/3D Imaging

Development, Analysis and Applications

Springer Vieweg

Benjamin Langmann
Hannover, Germany

Also PhD Thesis, University of Siegen, 2013

ISBN 978-3-658-06456-3 ISBN 978-3-658-06457-0 (eBook)
DOI 10.1007/978-3-658-06457-0

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Library of Congress Control Number: 2014942840

Springer Vieweg
© Springer Fachmedien Wiesbaden 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law. The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use. While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer Vieweg is a brand of Springer DE.
Springer DE is part of Springer Science+Business Media.
www.springer-vieweg.de

Acknowledgements

Conducting this work was made possible in the helpful and fruitful environment of the Center for Sensor System (ZESS) at the University of Siegen. I would like to thank the laboratory staff of the ZESS for the practical help and especially my supervisors Dr.-Ing. Klaus Hartmann and Prof. Dr.-Ing. Otmar Loffeld as well as Dr. Wolfgang Weihs for the valuable discussions and their support. Special thanks go to my second examiner Prof. Dr.-Ing. Andreas Kolb for his encouragement as well as the scientific guidance. Moreover, I would like to thank my sister Mirjam Spies for proof reading and my parents for their continued support during my studies.

Benjamin Langmann

Abstract

Imaging technology is widely utilized in a growing number of disciplines ranging from gaming, robotics and automation to medicine. In the last decade also 3D imaging found increasing acceptance and application, which were largely driven by the development of novel 3D cameras and measuring devices. These cameras are usually limited to indoor scenes with relatively low distances. In this thesis the development and the evaluation of medium and long-range 3D cameras are described in order to overcome these limitations. The MultiCam, a monocular 2D/3D camera which incorporates a color and a depth imaging chip, forms the basis for this research. The camera operates on the Time-of-Flight (ToF) principle by emitting modulated infrared light and measuring the round-trip time. In order to apply this kind of camera to larger scenes, novel lighting devices are required and will be presented in the course of this work. On the software side methods for scene observation working with 2D and 3D data are introduced and adapted to large scenes. An extended method for foreground segmentation illustrating the advantages of additional 3D data is presented, but also its limitations due to the lower resolution of the depth maps are addressed.

Long-range depth measurements with large focal lengths and 3D imaging on mobile platforms are easily impaired by involuntary camera motions. Therefore, an approach for motion compensation with joint super-resolution is introduced to facilitate ToF imaging in these areas. The camera motion is estimated based on the high resolution color images of the MultiCam and can be interpolated for each phase image, i.e. raw image of the 3D imaging chip. This method was applied successfully under different circumstances.

A framework for multi-modal segmentation and joint super-resolution also addresses the lower resolution of the 3D imaging chip. It resolves the resolution mismatch by estimating high resolution depth maps while performing the segmentation. Subsequently, a global multi-modal and multi-view tracking approach is described, which is able to take advantage of any type and number of cameras. Objects are modeled with ellipsoids and their appearance is modeled with color histograms as well as density estimates. The thesis concludes with remarks on future developments and the application of depth cameras in new environments.

Contents

List of Figures

1 Introduction

In a world with an increasing amount and availability of computers or micro-processors, the interaction between the real world and computers becomes more significant. Optical sensors play an important role in this interaction, but cameras, i.e. optical sensor matrices, have become more prominent since they have dropped in price in recent years. Despite posing substantially higher demands on data transmission and processing, many consumer devices from cell-phones to cars are today equipped with at least one camera. In the field of gaming cameras have become a standard input device and in medical engineering and science cameras also find more and more application. Moreover, thanks to the advancement of processing power cameras are standard sensors in industrial automation and robotics. Since many image processing algorithms can be implemented on DSPs or FPGAs, high performance CPUs are not required for all applications.

Nevertheless, the limitation of the available processing power was a key obstacle, which drove the development of depth imaging sensors in the past decade. Restrictions of stereo camera setups in respect of reliability and precision proved also to be problematic. In general, two different approaches were followed, namely the structured light and the time-of-flight approach. The structured or coded light principle is based on the observation of disparities similar to the stereo camera method. However, the disparities are observed between a camera and a light projector instead of two cameras. On the one hand this approach aims at overcoming the ambiguities encountered in stereo setups, e.g. untextured objects, and on the other hand it aims at reducing the computational demand. Structured light approaches were dominated for a long time by line patterns projected in a sequence with increasing spatial resolution. However, recently advances were made with highly resolved dot patterns, which significantly impacted the gaming market in addition to research and development.

The rival distance measuring principle is called Time-of-Flight (ToF) and here the distance of an object to the camera is determined by emitting spatially uniform light, which is varied over time, thus allowing to measure the time until the light is received. Cameras following this operating principle have been researched for more than a decade and models from several

B. Langmann, *Wide Area 2D/3D Imaging,*
DOI 10.1007/978-3-658-06457-0_1, © Springer Fachmedien Wiesbaden 2014

manufacturers are on the market. These cameras find application in many different disciplines and it is expected that they become a standard sensor device in image processing.

In the past decades a range of alternative depth estimation approaches have been proposed. The approaches typically operate with standard color or grayscale cameras without any active lighting. Image features are used to derive the distance to objects in the scene. Popular methods are Shape-from-Shading, Shade-from-Focus and Shape-from-Silhouette. Even more advanced methods learn common depth distributions in scenes and transfer this knowledge onto new scenes.

The use of color images is extremely widespread in computer vision, since color cameras are cheap, ubiquitous and provide valuable information. Thus, depth cameras are often used in conjunction with color cameras. Methods exist to calibrate and register a color and a depth camera and to map the measurements afterwards. This mapping is performed by relying on the measured distance to an object. Measurement noise and inaccuracies therefore affect the mapping and lead to errors. Additionally, the cameras have a different point of view and hence do not share the same view, which causes holes in the mapping for example at edges of objects. In order to overcome these problems, a monocular combination of a color and a depth imaging chip was developed at the ZESS. A beam splitter allows both imaging chips to share the same view onto the scene, which renders the mapping of both images unnecessary. This 2D/3D camera is named MultiCam and allows a depth independent registration of both modalities.

1.1 Limitations of 2D/3D Imaging and Contributions

Much progress has been made in increasing the lateral resolution, the sensitivity and the applicability of ToF-based depth imaging chips, but several limitations still persist. High reliability of depth imaging can only be achieved indoors under controlled lighting conditions. Most depth cameras cannot operate in highly illuminated surroundings especially in sunlight. Mechanisms were developed allowing depth cameras to work outdoors, but this reduces the measurement quality achievable.

A related limitation of depth cameras is their measurement range. Many indoor scenes contain only short distances of a few meters. However, higher distances occur in common scenes in professional environments or outdoors.

The distance limit of available depth cameras lies between 1 and 10 meters depending on the device. This reduces the applicability of depth cameras in many situations.

The resolution of the depth imaging chips is a widely discussed other limitation of depth cameras. The first evolution of depth imaging chips consisted of only 64 × 48 pixels, which did not allow a reconstruction of typical scene geometries. At the time of writing depth imaging chips with 160 × 120 or 200 × 200 pixels are available and chips with higher resolution are in the experimental stage. In general, higher resolutions lead to smaller pixels, which require a higher sensitivity to achieve the same measurement quality. This is the main limiting factor of the resolution of depth imaging chips.

The mentioned limitations are addressed in this thesis as follows: On the technical side two concepts for novel active lighting devices are presented to facilitate depth imaging for medium-range outdoor scenes up to 75 meters, depending on the opening angle, and for long-range scenes up to 150 meters. Processing steps for these imaging devices aiming at scene observation are introduced. In particular, different approaches on how to fuse high resolution color images with low resolution depth maps without introducing false information are discussed.

1.2 Thesis Outline

The operating principles of depth cameras available as commercial products as well as research prototypes are reviewed in Chapter 2. Their capabilities and limitations are compared by means of a number of evaluation setups. The focus lies in comparing Time-of-Flight cameras manufactured by the company PMD Technologies (PMDTec) and its competitor SoftKinetic as state-of-the-art depth imaging technology to the Microsoft Kinect, which is based on a rivaling structured light technology.

On the basis of this characterization of current depth cameras, the theory of a common branch of depth imaging chips named Photonic Mixer Device (PMD) is explained in detail in Chapter 3 and the behavior of PMD chips is analyzed. The ZESS MultiCam, a monocular 2D/3D camera, with which most experiments for this thesis were conducted, is introduced in conjunction with novel lighting devices to facilitate depth measurement for distances up to 150 meters. Additionally, an approach to gain absolute depth values for long-range depth measurements with PMD chips is demonstrated. Methods to improve the depth images in order to obtain accurate depth measurements