

Studies in Computational Intelligence 804

Mahmoud Hassaballah
Khalid M. Hosny *Editors*

Recent Advances in Computer Vision

Theories and Applications

 Springer

Studies in Computational Intelligence

Volume 804

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series “Studies in Computational Intelligence” (SCI) publishes new developments and advances in the various areas of computational intelligence—quickly and with a high quality. The intent is to cover the theory, applications, and design methods of computational intelligence, as embedded in the fields of engineering, computer science, physics and life sciences, as well as the methodologies behind them. The series contains monographs, lecture notes and edited volumes in computational intelligence spanning the areas of neural networks, connectionist systems, genetic algorithms, evolutionary computation, artificial intelligence, cellular automata, self-organizing systems, soft computing, fuzzy systems, and hybrid intelligent systems. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution, which enable both wide and rapid dissemination of research output.

More information about this series at <http://www.springer.com/series/7092>

Mahmoud Hassaballah · Khalid M. Hosny
Editors

Recent Advances in Computer Vision

Theories and Applications

Editors

Mahmoud Hassaballah
Department of Computer Science
South Valley University
Luxor, Egypt

Khalid M. Hosny
Department of Information Technology
Zagazig University
Zagazig, Egypt

ISSN 1860-949X

ISSN 1860-9503 (electronic)

Studies in Computational Intelligence

ISBN 978-3-030-02999-9

ISBN 978-3-030-03000-1 (eBook)

<https://doi.org/10.1007/978-3-030-03000-1>

Library of Congress Control Number: 2018958934

© Springer Nature Switzerland AG 2019

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.

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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword

In the last two decades, computer vision has established itself as a separate field inside computer science, next to (and partially overlapped with) artificial intelligence. Computer vision is an interdisciplinary area which comprises all aspects of image and video processing that can be used in artificial visual systems for automatic scene analysis, interpretation, and understanding. Unlike traditional image processing, it's history can be traced back to the 1960s and that has been focused, namely, on low-level 2D picture processing. The development of computer vision is closely connected with robotics. Intelligent autonomous robots must be equipped by fast and accurate vision system, which is able to analyze the visual field of the robot in real time and which transfers the visual information into a symbolic and/or quantitative one that serves as the input for robots control system. Recently, computer vision came into a public awareness, thanks to the appearance of autonomous vehicles, which has been widely presented and discussed.

No doubt that computer vision is one of the hottest topics of current computer science and engineering. Thanks to this, we have been witnesses to enormous growth of the number of computer vision methods, algorithms, and related papers and books. Recent progress in deep learning and convolution networks, which both can be applied to resolve many computer vision problems, illustrates the effort which has been invested in this area. If one puts the phrase “computer vision” into Google search, he/she will get more than 500000000 search results. Under these circumstances, the book entitled “Recent Advances in Computer Vision” is extremely useful, even if no single book can capture all “recent advances” on this field. The book you currently hold in your hands is a multi-authored book which presents selected topics according to the author's choice. In 16 chapters, the book presents both theoretical- and application-oriented methods and algorithms.

The chapters, Content-Based Image Retrieval using Multiresolution Feature Descriptors, Video Similarity Measurement and Search, Landmark Recognition: From Small-Scale to Large-Scale Retrieval, Analysis and Evaluation of Keypoint Descriptors for Image Matching, Motion Estimation Made Easy: Evolution and Trends in Visual Odometry, Scene Classification Using Transfer Learning, Feature Extraction of Color images using Quaternion Moments, and Hyperspectral

Image: Fundamentals and Advances, are mainly focused on theory and present state-of-the-art surveys of the respective fields. While the other chapters are oriented to some particular application area, such as face recognition (Facial Features Detection and Localization, Advances and Trends in Video Face Alignment, Face Recognition Using Exact Gaussian-Hermit Moments and Nonnegative Matrix Factorization, Face Recognition with Discrete Orthogonal Moments), biometrics (Deep Ear Recognition), medicine (Boundary Detection of Echocardiographic Images during Mitral Regurgitation), ocean biology (Ocean Ecosystems Plankton Classification), and fashion (Computer Vision for Supporting Fashion Creative Processes).

The target readership of this book includes academic researchers, advanced students, and R&D engineers from all computer vision application areas, who want to stay in touch with recent developments of this dynamic and fascinating field.

Prague, Czech Republic

Prof. Jan Flusser
Institute of Information
Theory and Automation
Czech Academy of Sciences

Preface

Indeed, scarcely a month passes where one does not hear from active research groups and industry community an announcement of some new technological breakthrough in the areas of intelligent systems and computational photography related to computer vision field. One aspect of computer vision that makes it such an interesting topic of study and active research field is the amazing diversity of our daily life applications that make use of (or depend on) computer vision or its research finds. Thus, computer vision is a key technology in these applications. For instance, computer vision is routinely used in industrial automation for the quality and process control. There are significant contributions for computer vision algorithms in pedestrian protection systems, biometric systems, driver assistance systems, and video surveillance and even in the movie industry via using avatars and creation of virtual worlds based on recorded images. This is just mentioning a few application areas, which all come with particular digital image data, where exceptional needs to analyze and process these data are a must.

The past and on-going research on computer vision and its related image processing cover a wide domain of topics and tasks, from basic research to a large number of real-world industrial applications. In recent years, there have been enormous progress and major results achieved in the field of computer vision, where several new methods and robust algorithms have been published in this active research field. For instance, deep learning algorithms have brought a revolution to the computer vision community, introducing nontraditional and efficient solutions to several image-related problems that had long remained unsolved. This rapid and tremendous progress is the inspiration for this book. In spite of the market has many good books on computer vision, there is always a need for another new one that covers recent progress and achievements in the field as no single book can do that. Our book is influenced by the editing books of “Emerging topics in computer vision” edited by Medioni and Kang in 2005 and “Emerging topics in computer vision and its applications” edited by Chen C. in 2012. The book gives a comprehensive overview of the most advanced methodologies and modern applications in the field of computer vision. Particularly, these did not cover by the previous books and the most recent research and development. Thus, the topics

covered by the book present a collection of high-quality research works written by renowned leaders in the field.

Because of the rapid developments in computer vision, the book is a contributed volume where well-known experts from several countries are dealing with specific research and application problems. It presents the state-of-the-art as well as the most recent trends in fashion creative processes, plankton classification, facial features detection, face recognition, feature description and matching, visual odometry, scene classification, transfer learning, video face alignment, video searching, hyperspectral image, object segmentation, ear recognition, and image retrieval. As the book covers this wide range of topics and cutting-edge methods with applications as well as recent advancements and challenges in computer vision, we would expect that the content is of considerable interest to readers worldwide, even they are not experts (or professionals) in this field. Thus, the primary target audience for the book is from both computer science and electronics engineers including researchers, scholars, postgraduate students, and developers who are interested in computer vision, pattern recognition, image processing, and connected research disciplines. The book is a valuable companion and comprehensive reference for both postgraduate and possibly some senior undergraduate students who are taking an advanced course in related topics (e.g., computer vision and image processing). However, it is also likely to provide a reference source of interest and value to practitioners from industry.

The book consists of 16 chapters, where each chapter provides the principles and fundamentals of a specific topic, introduces reviews up-to-date techniques, presents outcomes, and points out challenges and future directions. In each chapter, figures, tables, and examples are used to improve the presentation and the analysis of the covered topic. Furthermore, bibliographic references are included in each chapter providing a good starting point for deeper research and further exploration of the topics considered in this book. Further, this book is structured such that each chapter can be read independently from the others. A brief overview of the contents of the book is as follows:

Chapter “[Computer Vision for Supporting Fashion Creative Processes](#)” considers using computer vision techniques as powerful tools to support and enhance creative workflows in fashion industries, where the starting point of the fashion workflow is designer sketches and drawings—made with pen or pencil on raw paper. Then, such hand-drawn sketches must be imported into a software, to convert the prototype into a real-world product. In this chapter, some state-of-the-art techniques and a novel proposal for line extraction and thinning, applied to fashion sketches, are described. Additionally, many recently developed methods are presented and their effectiveness in the recognition of different types of feature is discussed.

Chapter “[Facial Features Detection and Localization](#)” discusses basic concepts related to the problem of facial landmarks detection and overviews the successes and failures of existing solutions. Besides, it explores the difficulties that hinder the path of progress in the topic and the challenges involved in the adaptation of existing approaches to build successful systems for real-world facial image-related

applications. Also, it discusses the performance evaluation metrics and the available benchmarking datasets. Finally, it suggests some possible future directions for research in the topic.

Chapter “[Advances and Trends in Video Face Alignment](#)” addresses the problem of face alignment in a video, which is an important research area in computer vision providing strong support for video face recognition, face animation, etc. In the chapter, definition of face alignment in a video and its significance are described. Methods for face alignment in a video are divided into three kinds: face alignment using image alignment algorithms, joint alignment of face images, and face alignment using temporal and spatial continuities. The chapter mainly focuses on joint face alignment and face alignment using temporal and spatial continuities.

Chapter “[Video Similarity Measurement and Search](#)” presents and discusses the core field concepts, problems, and recent trends in content-based video retrieval systems (CBVR), where video matching for the retrieval purpose is the core of these CBVR systems. Besides, this chapter provides readers with the required amount of knowledge to select suitable features’ set and adequate techniques to develop robust research in this field.

Chapter “[Analysis and Evaluation of Keypoint Descriptors for Image Matching](#)” addresses feature keypoint descriptors, which are indispensable tools and essential components in many computer vision systems. In this chapter, authors have discussed the most frequently used floating-point descriptors and their binary competitors. Then, a general framework to measure the performance of keypoint descriptors is given as well as the performance metrics and matching criteria followed to obtain a set of correct matches are explained in details. The framework is used to evaluate the performance of several recently proposed keypoint descriptors and a detailed explanation of the experimental results and findings is given.

Chapter “[Feature Extraction of Color Images Using Quaternion Moments](#)” covers the topic of extracting color images features, which is the backbone of many computer vision applications. It focuses exactly on using quaternion moments, where an overview of the different quaternion moments and their rotation, scaling, and translation invariants is given. A series of numerical experiments are performed using different parameters to assess the performance of extracted features using different quaternion moments.

Chapter “[Face Recognition Using Exact Gaussian-Hermit Moments](#)” addresses the problem of face recognition using exact Gaussian-Hermit moments (EGHMs) to extract the features of face images, where the higher order EGHMs are able to capture the higher order nonlinear features of images. The rotation, scaling, and translation invariants of EGHMs are used to overcome the geometric distortions. The nonnegative matrix factorization technique is investigated to classify the extracted features.

In Chapter “[Face Recognition with Discrete Orthogonal Moments](#)”, authors investigate the applicability of discrete orthogonal Hahn and Racah moments for face recognition problem under lighting, facial expression, and face detail changes. They utilize the conventional nearest neighbor algorithm with the Euclidean and Manhattan normalized distance metrics to classify the discrete orthogonal moment

features. With some experimental results, they show that the discrete Hahn and Racah moments are able to extract global facial features as well as local characteristics, thus providing the holistic, the component-based, and the fused approaches for feature representation.

Chapter “[Content-Based Image Retrieval Using Multiresolution Feature Descriptors](#)” presents the concept of multiresolution feature descriptors for content-based image retrieval systems, where the important properties and advantages of multiresolution feature descriptors are discussed. Furthermore, this chapter proposes a simple method based on utilizing a new proposed multiresolution descriptor. The proposed method depends on capturing shape feature in a localized manner at multiple resolutions of image and required no segmentation operation.

Chapter “[Landmark Recognition: From Small-Scale to Large-Scale Retrieval](#)” explains the most used techniques for solving the problem of landmark recognition. The focus is on the classical and deep learning approaches for the creation of descriptors used in many tasks of computer vision, where a particular attention is put on the major recent breakthroughs in content-based image retrieval using transfer learning and fine-tuning techniques which improves the feature representation and therefore accuracy of the retrieval system. Finally, the chapter exposes techniques for large-scale retrieval in which datasets contain at least a million images.

Chapter “[Ocean Ecosystems Plankton Classification](#)” covers a very challenging computer vision task, which is automatic plankton classification. Several related works about plankton classification are reviewed in the chapter. Besides, it presents an ensemble of heterogeneous descriptors and automated plankton recognition system based on deep learning methods combined with so-called handcrafted features.

Chapter “[Boundary Detection of Echocardiographic Images During Mitral Regurgitation](#)” introduces a method to automatically find out the boundaries of left ventricle and left atrium, i.e., boundaries of hearts chambers, which play an important role in determining the area of left ventricle or left atrium (a measure of severity of mitral regurgitation), where two basic techniques are considered in this method, namely, the fast region active contour model and selective binary Gaussian filtering regularized level set.

In Chapter “[Motion Estimation Made Easy: Evolution and Trends in Visual Odometry](#)”, an attempt is made to introduce the visual odometry topic for beginners covering different aspects of vision-based motion estimation task. The theoretical section provides a brief on different computer vision fundamentals specific to pose estimation tasks followed by a systematic discussion on the visual odometry schemes under different categories. The evolution of these schemes over last few decades is discussed under geometric and nongeometric approaches.

Chapter “[Deep Ear Recognition Pipeline](#)” covers an ear biometric problem under image data captured in completely unconstrained settings using deep-learning-based approaches, where a novel ear recognition pipeline for ear detection and recognition based on the convolutional neural network is introduced. Detailed analysis of the proposed pipeline for detection and recognition separately, as well as

joint analysis, is also presented. Several experiments are conducted to test the performance of the proposed pipeline.

Chapter “[Scene Classification Using Transfer Learning](#)” presents a transfer learning-based approach for scene classification. A pretrained convolutional neural network is used as a feature extractor. The pretrained network along with classifiers such as support vector machines or multilayer perceptron is used to classify the images. Also, the effect of single-plane images such as RGB2Gray, SVD decolorized, and modified SVD decolorized images is analyzed based on classification accuracy, class-wise precision, recall, F1-score, and equal error rate.

Chapter “[Hyperspectral Image: Fundamentals and Advances](#)” provides the overview of fundamentals and advances in hyperspectral images. The hyperspectral image enhancement, denoising and restoration, classical classification techniques, and the most recently popular classification algorithm are discussed with more details. Besides, the standard hyperspectral datasets used for the research purposes are covered in this chapter.

Finally, we hope that our readers find the presented chapters in the book interesting and will inspire future research both from theoretical and practical viewpoints to spur further advances in the computer vision field.

Acknowledgements

The editors would like to take this opportunity to express their sincere gratitude to the authors of the chapters for extending their wholehearted support in sharing some of their latest results and findings. Without their significant contribution, this book volume could not have fulfilled its mission. The reviewers deserve our thanks for their constructive and timely input. Special profound thanks go to Prof. Ing. Jan Flusser, Institute of Information Theory and Automation, Czech Republic for writing the foreword of the book. The editors are very grateful to Editor-in-Chief, Prof. Janusz Kacprzyk, Dr. Thomas Ditzinger, the senior editor, and Holger Schpe, the editorial assistant of the Studies in Computational Intelligence series. Finally, the editors and the authors acknowledge the efforts of the Studies in Computational Intelligence team at Springer for their support and cooperation in publishing the book as a volume in the SCI series. Really, the editorial staff at Springer has done a meticulous job and working with them was a pleasant experience.

Luxor, Egypt
Zagazig, Egypt

Mahmoud Hassaballah
Khalid M. Hosny

Contents

Computer Vision for Supporting Fashion Creative Processes	1
Luca Donati, Eleonora Iotti and Andrea Prati	
Facial Features Detection and Localization	33
M. Hassaballah, Saddam Bekhet, Amal A. M. Rashed and Gang Zhang	
Advances and Trends in Video Face Alignment	61
Gang Zhang, Yuding Ke, Weikang Zhang and M. Hassaballah	
Video Similarity Measurement and Search	85
Saddam Bekhet, M. Hassaballah, Amr Ahmed and Ali H. Ahmed	
Analysis and Evaluation of Keypoint Descriptors for Image Matching	113
M. Hassaballah, Hammam A. Alshazly and Abdelmgeid A. Ali	
Feature Extraction of Color Images Using Quaternion Moments	141
Khalid M. Hosny and Mohamed M. Darwish	
Face Recognition Using Exact Gaussian-Hermit Moments	169
Khalid M. Hosny and Mohamed Abd Elaziz	
Face Recognition with Discrete Orthogonal Moments	189
Fatima Akhmedova and Simon Liao	
Content-Based Image Retrieval Using Multiresolution Feature Descriptors	211
Prashant Srivastava and Ashish Khare	
Landmark Recognition: From Small-Scale to Large-Scale Retrieval	237
Federico Magliani, Tomaso Fontanini and Andrea Prati	
Ocean Ecosystems Plankton Classification	261
A. Lumini and L. Nanni	

Boundary Detection of Echocardiographic Images During Mitral Regurgitation 281
K. Chauhan and R. K. Chauhan

Motion Estimation Made Easy: Evolution and Trends in Visual Odometry 305
Shashi Poddar, Rahul Kottath and Vinod Karar

Deep Ear Recognition Pipeline 333
Žiga Emeršič, Janez Kržaj, Vitomir Štruc and Peter Peer

Scene Classification Using Transfer Learning 363
Nikhil Damodaran, V. Sowmya, D. Govind and K. P. Soman

Hyperspectral Image: Fundamentals and Advances 401
V. Sowmya, K. P. Soman and M. Hassaballah

Author Index 425

About the Editors



Mahmoud Hassaballah was born in 1974, Qena, Egypt. He received a B.Sc. degree in Mathematics in 1997, and then M.Sc. degree in Computer Science in 2003, all from South Valley University, Egypt. In April 2008, he joined the Lab of Intelligence Communication, Department of Electrical and Electronic Engineering and Computer Science, Ehime University, Japan as a Ph.D. student, where he received a Doctor of Engineering (D.Eng.) in Computer Science on September 2011 for his work on facial features detection. He is currently an Associate Professor of Computer Science at Faculty of Computers and Information, South Valley University, Luxor, Egypt. His research interests include feature extraction, object detection/recognition, biometrics, image processing, computer vision, machine learning, and data hiding.



Khalid M. Hosny was born in 1966, Zagazig, Egypt. He is a Professor of Information Technology at Faculty of Computers and Informatics, Zagazig University. He received the B.Sc., M.Sc., and Ph.D. from Zagazig University, Egypt in 1988, 1994, and 2000, respectively. From 1997 to 1999, he was a Visiting Scholar, University of Michigan, Ann Arbor and University of Cincinnati, Cincinnati, USA. He is a senior member of ACM and IEEE. His research interests include image processing, pattern recognition, multimedia, and computer vision. He published more than 60 papers in international journals. He is an editor and scientific reviewer for more than 30 international journals.