Lecture Notes in Computer Science 7963

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Mei Chen Bastian Leibe
Bernd Neumann (Eds.)

Computer Vision Systems

9th International Conference, ICVS 2013
St. Petersburg, Russia, July 16-18, 2013
Proceedings

Springer

Volume Editors

Mei Chen
Carnegie Mellon University, Pittsburgh, PA, USA
E-mail: mei.chen@intel.com

Bastian Leibe
RWTH Aachen University, Germany
E-mail: leibe@vision.rwth-aachen.de

Bernd Neumann
University of Hamburg, Germany
E-mail: neumann@informatik.uni-hamburg.de

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-39401-0 e-ISBN 978-3-642-39402-7
DOI 10.1007/978-3-642-39402-7
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2013942398

CR Subject Classification (1998): I.5, I.4, I.3, I.2.10, I.5.4, C.3

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Understanding the computational processes of visual perception and constructing vision systems has been a prominent goal of computer scientists since the advent of computers. In 1955, Oliver Selfridge proclaimed to build "eyes and ears for the computer," marking the beginning of a long development, guided by changing paradigms and leading to exciting insights. While the enormous complexity of human vision is still being explored, technological developments have reached a state where computational vision processes can be integrated into operational systems, realizing a spectrum from ambitious applications to advanced experimental cognitive systems.

Whereas research on methods and theories has been well supported by scientific journals and major international conferences such as CVPR, ICCV, and ECCV from early on, the challenges of building complex operational vision systems received less attention. Therefore, a new conference series, titled "International Conference on Computer Vision Systems" (ICVS), was created, with the first conference taking place at Las Palmas on the Canary Islands in 1999. The idea of this conference series is to primarily address issues arising in the design and deployment of comprehensive computer vision systems for a broad spectrum of applications, such as robotics, monitoring, video analysis, and scene interpretation.

This volume contains the contributions to the 9th ICVS hosted in St. Petersburg, Russia. The focus is on diverse aspects of computer vision systems, ranging from video capture to high-level image interpretation, typically treated in the context of a realistic application or system building task. It is worth noting that a major fraction of the contributions deal with robot vision, mirroring the increasing support and interest for artificial cognitive systems and robotic applications. For example, several papers address the challenging task of analyzing a mobile world by moving sensors. While there is impressive progress, it is also apparent that much more work will be needed in this field and hopefully be presented at future ICVS conferences.

There were 94 submissions to ICVS 2013, out of which the Program Chairs selected 16 papers for oral presentation at the conference and 20 papers for presentation as posters, based on the careful reviews of the Program Committee and additional reviewers. All accepted papers have been revised by their authors to address reviewers' comments and are presented in this volume. We would like to thank all the authors for their submissions and all the reviewers for their valuable comments.

May 2013

Mei Chen
Bastian Leibe
Bernd Neumann

Organization

General Chair

Markus Vincze	Vienna Institute of Technology, Austria

Local Organization

Alexandr Redkin	Intel, St. Petersburg, Russia

Program Co-chairs

Mei Chen	Intel, Carnegie Mellon University, Pittsburgh, USA
Bastian Leibe	RWTH Aachen University, Germany
Bernd Neumann	University of Hamburg, Germany

Program Committee

Michael Arens	Fraunhofer IOSB, Germany
Antonios Argyros	University of Crete, Greece
Christian Bauckhage	Fraunhofer IAIS, Germany
Eduardo Bayro-Corrochano	CINVESTAV, Guadalajara, Mexico
Alexandre Bernardino	Instituto Superior Técnico, Portugal
Alain Boucher	IFI, Vietnam
Richard Bowden	University of Surrey, UK
François Bremond	INRIA, France
Jorge Cabrera	University of Las Palmas de Gran Canaria, Spain
Régis Clouard	GREYC Laboratory, France
Daniel Cremers	TU Munich, Germany
James Crowley	I.N.P. Grenoble, France
Kostas Daniilidis	University of Pennsylvania, USA
Larry Davis	University of Maryland, USA
Bruce Draper	USA
Gian Luca Foresti	Colorado State University, Italy
Jannik Fritsch	Honda Research Institutes, Germany
Antonios Gasteratos	Democritus University of Thrace, Greece
Rolf-Rainer Grigat	Technical University Hamburg-Harburg, Germany
Guodong Guo	West Virginia University, USA

Riad Hammoud	DynaVox Technologies, USA
Vasek Hlavac	Czech Technical University in Prague, Czech Republic
Jesse Hoey	University of Dundee, UK
David Hogg	University of Leeds, UK
Costas Kotropoulos	Aristotle University of Thessaloniki, Greece
Ivan Laptev	INRIA Paris-Rocquencourt, France
Dan Lee	University of Pennsylvania, USA
Ales Leonardis	University of Ljubljana, Slovenia
Bärbel Mertsching	University of Paderborn, Germany
Giorgio Metta	University of Genova/Instituto Italiano di Tecnologia, Italy
Lazaros Nalpantidis	Aalborg University, Denmark
Ramakant Nevatia	University of Southern California, USA
Lucas Paletta	Joanneum Research, Austria
Theodore Papadopoulo	INRIA, France
Nikolaos Papamarkos	Democritus University of Thrace, Greece
Claudio Pinhanez	IBM T.J. Watson Research Center, USA
Axel Pinz	Graz University of Technology, Austria
Fiora Pirri	University of Rome “La Sapienza”, Italy
Ioannis Pratikakis	Democritus University of Thrace, Greece
Carlo Regazzoni	University of Genova, Italy
Paolo Remagnino	Kingston University, UK
Yoichi Sato	University of Tokyo, Japan
Silvio Savarese	University of Michigan, USA
Gerald Schaefer	Loughborough University, UK
Bernt Schiele	Max-Planck-Institut für Informatik, Germany
Konrad Schindler	ETH Zürich, Switzerland
Yaser Sheikh	Carnegie Mellon University, USA
Alexander Stoytchev	Iowa State University, USA
Agnes Swadzba	University of Bielefeld, Germany
Anastasios Tefas	Aristotle University of Thessaloniki, Greece
Monique Thonnat	INRIA Sophia Antipolis, France
Qi Tian	University of Texas, San Antonio, USA
Panos Trahanias	University of Crete, Greece
John Tsotsos	York University, Canada
Emre Ugur	ATR, Kyoto, Japan
Marc Van Droogenbroeck	University of Liege, Belgium
Sergio Velastin	Kingston University, UK
Sven Wachsmuth	University of Bielefeld, Germany
Christian Woehler	Technical University Dortmund, Germany
Sebastian Wrede	University of Bielefeld, Germany
Zhaozheng Yin	Missouri University, USA
Hongbin Zha	Peking University, China
Li Zhang	University of Wisconsin Madison, USA

Additional Reviewers

Wongun Choi	University of Michigan, USA
Srikanth Kallakuri	ATR, Japan
Dennis Mitzel	RWTH Aachen University, Germany
Zhen Zeng	University of Michigan, USA

Visual Processing and Understanding of Human Faces and Bodies

Takeo Kanade

Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213, USA

Abstract. A human face and body convey important information to understand a person: her identity, emotion, action, and intention. Technologies to process video of human faces and bodies have many applications, ranging from biometrics to medical diagnosis and from surveillance to cognitive human-robot interaction. This talk will give highlights of the recent work at the CMU Vision Group, in particular, to robust face (and object) alignment, real-time face tracking, facial Action Unit (AU) recognition for emotion analysis, 2D and 3D body tracking, and facial video cloning for understanding human dyadic communication.

Active Pedestrian Safety: From Research to Reality

Dariu M. Gavrila

Dept. of Environment Perception, Daimler R&D, Ulm, Germany
dariu.gavrila@daimler.com

Abstract. One of the most significant large-scale deployments of intelligent systems in our daily life nowadays involves driver assistance in smart cars. The past decade has witnessed a steady increase of interest in the plight of the vulnerable road users, i.e. pedestrians and bicyclists. Accident statistics show that roughly one quarter of all traffic fatalities world-wide involve vulnerable road users; most accidents occur in an urban setting. Devising an effective driver assistance system for vulnerable road users has long been impeded, however, by the "perception bottleneck", i.e. not being able to detect and localize vulnerable road users sufficiently accurate. The problem is challenging due to the large variation in object appearance, the dynamic and cluttered urban backgrounds, and the potentially irregular object motion. Topping these off are stringent performance criteria and real-time constraints. I give an overview of the remarkable research progress that has been achieved in this area and discuss its main enablers: the algorithms, the data, the hardware and the tests. Our long-standing research on vulnerable road users has recently paid off. Our company, Daimler, deploys an advanced set of driver assistance functions for its Mercedes-Benz 2013 E- and S-Class models, termed "Intelligent Drive", using stereo vision sensing. It includes a pedestrian safety component which facilitates fully automatic emergency braking - the system works day and night. I conclude by discussing future research directions, on the road towards accident-free driving.

Addressing Some of the Challenges in Building Vision Systems

Martial Hebert

Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213, USA
hebert@ri.cmu.edu

Abstract. Despite considerable progress in the past two decades in all aspects of computer vision, building vision systems for practical applications remains a considerable challenge. This is especially true in applications such as robotics and autonomous systems, in which even a small error rate in the perception system can have catastrophic effects on the overall system. This talk will review a few ideas that can be used to start formalizing the issues revolving around the integration of vision systems. They include a systematic approach to the problem of self-assessment of vision algorithm performance and of predictive quality metrics on the inputs to vision algorithms, ideas on how to manage multiple hypotheses generated from a vision algorithm rather than relying on a single "hard" decision, and methods for using external (non-visual) domain- and task-dependent information and constraints to boost the performance of the vision system. These ideas will be illustrated with examples of recent vision systems for scene understanding, depth cueing, and object recognition for robotics and HRI applications.

Deep Hierarchies in Human and Computer Vision

Norbert Krüger

University of Southern Denmark
norbert@mmmi.sdu.dk

Abstract. In the last decade, computer vision - although being still a rather young scientific discipline - was able to provide some impressive examples of artificial vision systems that outperform humans (for example in terms of speed in well defined industrial inspection tasks or in terms of precision as for example in the latest vision based technology for goal decisions in football). However, the human visual system is still superior to any artificial vision system in visual tasks requiring generalization and reasoning (often also called 'cognitive vision') such as extraction of visual based affordances or visual tasks in the context of tool use and dexterous manipulation. Two decades ago, there has been a strong connection between the two communities dealing with human vision research and computer vision. This link however has been somehow lost recently and computer vision has been more and more developed into a sub-field of machine learning. In this talk, I argue that the reason for the superiority of human vision for 'cognitive vision tasks' is connected to the deep hierarchical architecture of the primate's visual system which contradicts the 'flat design' of many of the most successful artificial vision systems available today. I will ponder on this somehow disturbing fact and will give arguments in favor of establishing deep structures as the only way to be able to solve cognitive vision tasks. The talk is divided into three parts: First, I will give an overview about today's knowledge about the primate's (and by that the human's) visual system primarily based on neurophysiological research. In the second part of the talk, an 'Early cognitive Vision' (ECV) system is described. The ECV system is a deep hierarchical visual representation designed as a visual front-end facilitating the realization of higher level processes, in particular on intelligent robot systems. Basic design choices of the system have been motivated by analogies to the human visual system. The current status of the ECV system is presented as well as a number of applications, in particular in the context of certain learning tasks connected to robotics. In the last part, I will summarize current research attempts on deep hierarchies in computer vision made by other groups and will formulate a number of open research questions related to the design of artificial cognitive vision systems. By that I also hope to help to facilitate a renewal of a fruitful interaction between biological and computer vision.

Table of Contents

Image and Video Capture

Automatic Unconstrained Online Configuration of a Master-Slave Camera System 1
David Münch, Ann-Kristin Grosselfinger, Wolfgang Hübner, and Michael Arens

A Five-Camera Vision System for UAV Visual Attitude Calculation and Collision Warning 11
Akos Zarandy, Zoltan Nagy, Balint Vanek, Tamas Zsedrovits, Andras Kiss, and Mate Nemeth

An Integrated 4D Vision and Visualisation System 21
Csaba Benedek, Zsolt Jankó, Csaba Horváth, Dömötör Molnár, Dmitry Chetverikov, and Tamás Szirányi

Compensation for Multipath in ToF Camera Measurements Supported by Photometric Calibration and Environment Integration 31
Stefan Fuchs, Michael Suppa, and Olaf Hellwich

Coloured Video Code for In-Flight Data Transmission 42
Falk Schubert, Thilo Fath, and Harald Haas

Vision-Based Magnification of Corneal Endothelium Frames 52
Dario Comanducci and Carlo Colombo

Visual Attention and Object Detection

FACTS - A Computer Vision System for 3D Recovery and Semantic Mapping of Human Factors 62
Lucas Paletta, Katrin Santner, Gerald Fritz, Albert Hofmann, Gerald Lodron, Georg Thallinger, and Heinz Mayer

Anytime Perceptual Grouping of 2D Features into 3D Basic Shapes 73
Andreas Richtsfeld, Michael Zillich, and Markus Vincze

Model-Based Pose Estimation for Rigid Objects 83
Manolis Lourakis and Xenophon Zabulis

A Web-Service for Object Detection Using Hierarchical Models 93
Domen Tabernik, Luka Čehovin, Matej Kristan, Marko Boben, and Aleš Leonardis

Fast Detection of Multiple Textureless 3-D Objects 103
Hongping Cai, Tomáš Werner, and Jiří Matas

Biological Models for Active Vision: Towards a Unified Architecture 113
Kasim Terzić, David Lobato, Mário Saleiro, Jaime Martins, Miguel Farrajota, J.M.F. Rodrigues, and J.M.H. du Buf

Real-Time Image Recognition with the Parallel Directed Enumeration Method 123
Andrey V. Savchenko

Self-Localization and Pose Estimation

Simultaneous Localization and Mapping for Event-Based Vision Systems 133
David Weikersdorfer, Raoul Hoffmann, and Jörg Conradt

Dimensionality Reduction for Efficient Single Frame Hand Pose Estimation 143
Petros Douvantzis, Iason Oikonomidis, Nikolaos Kyriazis, and Antonis Argyros

A Head Pose Tracking System Using RGB-D Camera 153
Songnan Li, King Ngi Ngan, and Lu Sheng

A New Hierarchical Method for Markerless Human Pose Estimation 163
Yuan Lei, Huawei Pan, Weixia Chen, and Chunming Gao

Autonomous Robot Navigation: Path Planning on a Detail-Preserving Reduced-Complexity Representation of 3D Point Clouds 173
Rohit Sant, Ninad Kulkarni, Ainesh Bakshi, Salil Kapur, and Kratarth Goel

Motion and Tracking

High Frame Rate Egomotion Estimation 183
Natesh Srinivasan, Richard Roberts, and Frank Dellaert

Is Crowdsourcing for Optical Flow Ground Truth Generation Feasible? 193
Axel Donath and Daniel Kondermann

A Comparative Study on Multi-person Tracking Using Overlapping Cameras 203
Martijn C. Liem and Dariu M. Gavrila

Visual Tracking Using Superpixel-Based Appearance Model 213
Shahed Nejhum, Muhammad Rushdi, and Jeffrey Ho

Carried Object Detection and Tracking Using Geometric Shape Models and Spatio-temporal Consistency 223
Aryana Tavanai, Muralikrishna Sridhar, Feng Gu, Anthony G. Cohn, and David C. Hogg

Robust Multi-hypothesis 3D Object Pose Tracking 234
Georgios Chliveros, Maria Pateraki, and Panos Trahanias

Automatic Parameter Adaptation for Multi-object Tracking 244
Duc Phu Chau, Monique Thonnat, and François Brémond

Probabilistic Cue Integration for Real-Time Object Pose Tracking 254
Johann Prankl, Thomas Mörwald, Michael Zillich, and Markus Vincze

3D Reconstruction

Depth Estimation during Fixational Head Movements in a Humanoid Robot 264
Marco Antonelli, Angel P. del Pobil, and Michele Rucci

LaserGun: A Tool for Hybrid 3D Reconstruction 274
Marco Fanfani and Carlo Colombo

Accurate Dense Stereo Matching of Slanted Surfaces Using 2D Integral Images 284
Gwnag Yul Song, Seong Ik Cho, Dong Yong Kwak, and Joon Woong Lee

Integrating Multiple Viewpoints for Articulated Scene Model Aquisition 294
Leon Ziegler, Agnes Swadzba, and Sven Wachsmuth

Features, Learning and Validation

Explicit Context-Aware Kernel Map Learning for Image Annotation 304
Hichem Sahbi

Integrating Cue Descriptors in Bubble Space for Place Recognition 314
Özgür Erkent and Işıl Bozma

A Hierarchical Scheme of Multiple Feature Fusion for High-Resolution Satellite Scene Categorization 324
Wen Shao, Wen Yang, Gui-Song Xia, and Gang Liu

A Validation Benchmark for Assessment of Medial Surface Quality for Medical Applications 334
Agnés Borràs, Debora Gil, Sergio Vera, and Miguel Angel González

When Is a Confidence Measure Good Enough? 344
Patricia Márquez-Valle, Debora Gil, Aura Hernàndez-Sabaté, and Daniel Kondermann

Parallel Deep Learning with Suggestive Activation for Object Category Recognition 354
Karthik Mahesh Varadarajan and Markus Vincze

Author Index 365