

Amir R. Zamir · Asaad Hakeem
Luc Van Gool · Mubarak Shah
Richard Szeliski
Editors

Large-Scale Visual Geo-Localization

 Springer

Editors

Amir R. Zamir
Department of Computer Science
Stanford University
Stanford, CA
USA

Mubarak Shah
Department of Computer Science
University of Central Florida
Orlando, FL
USA

Asaad Hakeem
Machine Learning Division
Decisive Analytics Corporation
Arlington, VA
USA

Richard Szeliski
Facebook
Seattle, WA
USA

Luc Van Gool
Computer Vision Laboratory
ETH Zürich
Zürich
Switzerland

ISSN 2191-6586

ISSN 2191-6594 (electronic)

Advances in Computer Vision and Pattern Recognition

ISBN 978-3-319-25779-2

ISBN 978-3-319-25781-5 (eBook)

DOI 10.1007/978-3-319-25781-5

Library of Congress Control Number: 2016932324

© Springer International Publishing Switzerland 2016

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.

Printed on acid-free paper

This Springer imprint is published by SpringerNature

The registered company is Springer International Publishing AG Switzerland

Preface

One of the greatest inventions of all times is *camera*, a device for capturing images of the world. Computer vision emerged as the scientific field of understanding the images that cameras produce. Another influential invention of modern technology is *positioning systems*, instruments (namely GPS) for identifying one's location. Even though these two devices are often viewed as belonging to completely different realms, their products (images and locations) have a close relationship. For instance, by just looking at an image, one can sometimes guess its location, e.g., when a rain forest is in view or looking at Eiffel Tower. The opposite is also often true: knowing the location of an image can provide a rich context and significantly assist the process of understanding the visual content. The goal of this book is to explore this bidirectional relationship between images and locations.

In this book, we provide a comprehensive map of the state of the art in large scale visual geo-localization and discuss the emerging trends in this area. Visual geo-localization is defined as the problem of identifying the location of the camera that captured an image and/or the content of the image. Geo-localization finds numerous applications in organizing photo collections, law enforcement, or statistical analysis of commercial market trends.

The book is divided into four main parts: Data-Driven Geo-localization, Semantic Geo-localization, Geometric Matching based Geo-localization, and Real-World Applications. The first part (Data-Driven Geo-localization) discusses recent methods that exploit internet-scale image databases for devising geographically rich features and geo-localizing query images at a city, country, or global scale. The second part (Semantic Geo-localization) presents geo-localization techniques that are built upon high-level and semantic cues. The third part (Geometric Matching based Geo-localization) focuses on the methods that perform localization by geometrically aligning the query image against a 3D model (namely digital elevation model or structure-from-motion reconstructed cities).

In the fourth part of the book (Real-World Applications), methods that use the geo-location of the image, whether extracted automatically or acquired from a GPS-chip, for understanding the image content are presented. Such frameworks are

of great importance as cell phones and cameras are now being equipped with built-in localization devices. Therefore, it is of particular interest to develop techniques that accomplish image analysis *assisted by geo-location*. We also discuss several approaches for geo-localization under more practical settings, e.g., when the supervision of an expert in the form of a user-in-the-loop system is available.

The editors sincerely appreciate everyone who assisted in the process of preparing this book. In particular, we thank the authors, who are among the leading and most active researchers in the field of computer vision. Without their contribution in writing and peer reviewing the chapters, this book would not have been possible. We are also thankful to Springer for providing us with excellent support. Lastly, the editors are truly grateful to their families and loved ones for their continual support and encouragement throughout the process of writing this book.

Stanford, CA, USA
Arlington, VA, USA
Zürich, Switzerland
Orlando, FL, USA
Seattle, WA, USA
May 2015

Amir R. Zamir
Asaad Hakeem
Luc Van Gool
Mubarak Shah
Richard Szeliski

Contents

1	Introduction to Large-Scale Visual Geo-localization	1
	Amir R. Zamir, Asaad Hakeem, Luc Van Gool, Mubarak Shah and Richard Szeliski	
Part I Data-Driven Geo-localization		
2	Discovering Mid-level Visual Connections in Space and Time	21
	Yong Jae Lee, Alexei A. Efros and Martial Hebert	
3	Where the Photos Were Taken: Location Prediction by Learning from Flickr Photos	41
	Li-Jia Li, Rahul Kumar Jha, Bart Thomee, David Ayman Shamma, Liangliang Cao and Yang Wang	
4	Cross-View Image Geo-localization	59
	Tsung-Yi Lin, Serge Belongie and James Hays	
5	Ultrawide Baseline Facade Matching for Geo-localization	77
	Mayank Bansal, Kostas Daniilidis and Harpreet Sawhney	
Part II Semantic Reasoning Based Geo-localization		
6	Semantically Guided Geo-location and Modeling in Urban Environments	101
	Gautam Singh and Jana Košecká	
7	Recognizing Landmarks in Large-Scale Social Image Collections	121
	David J. Crandall, Yunpeng Li, Stefan Lee and Daniel P. Huttenlocher	

Part III Geometric Matching Based Geo-localization

8	Worldwide Pose Estimation Using 3D Point Clouds	147
	Yunpeng Li, Noah Snavely, Daniel P. Huttenlocher and Pascal Fua	
9	Exploiting Spatial and Co-visibility Relations for Image-Based Localization	165
	Torsten Sattler, Bastian Leibe and Leif Kobbelt	
10	3D Point Cloud Reduction Using Mixed-Integer Quadratic Programming	189
	Hyun Soo Park, Yu Wang, Eriko Nurvitadhi, James C. Hoe, Yaser Sheikh and Mei Chen	
11	Image-Based Large-Scale Geo-localization in Mountainous Regions	205
	Olivier Saurer, Georges Baatz, Kevin Köser, L'ubor Ladický and Marc Pollefeys	
12	Adaptive Rendering for Large-Scale Skyline Characterization and Matching	225
	Jiejie Zhu, Mayank Bansal, Nick Vander Valk and Hui Cheng	
13	User-Aided Geo-location of Untagged Desert Imagery	239
	Eric Tzeng, Andrew Zhai, Matthew Clements, Raphael Townshend and Avidesh Zakhor	
14	Visual Geo-localization of Non-photographic Depictions via 2D–3D Alignment	255
	Mathieu Aubry, Bryan Russell and Josef Sivic	

Part IV Real-World Applications

15	A Memory Efficient Discriminative Approach for Location-Aided Recognition	279
	Sudipta N. Sinha, Varsha Hedau, C. Lawrence Zitnick and Richard Szeliski	
16	A Real-World System for Image/Video Geo-localization	299
	Himaanshu Gupta, Yi Chen, Minwoo Park, Kiran Gunda, Gang Qian, Dave Conger and Khurram Shafique	
17	Photo Recall: Using the Internet to Label Your Photos	321
	Neeraj Kumar and Steven Seitz	
	Index	349

Contributors

Mathieu Aubry LIGM (UMR CNRS 8049), ENPC/Université Paris-Est, Marne-la-Vallée, France

Georges Baatz Google Inc., Zürich, Switzerland

Mayank Bansal Vision Technologies Lab, SRI International, Princeton, NJ, USA; Center for Vision Technologies, SRI International, Princeton, NJ, USA

Serge Belongie Cornell University, Ithaca, NY, USA

Liangliang Cao IBM Watson Research, New York, USA

Mei Chen Intel Corporation, Hillsboro, OR, USA

Yi Chen Object Video, Inc., Reston, VA, USA

Hui Cheng Vision Technologies Lab, SRI International, Princeton, NJ, USA

Matthew Clements UC Berkeley, Berkeley, CA, USA

Dave Conger Object Video, Inc., Reston, VA, USA

David J. Crandall Indiana University, Bloomington, IN, USA

Kostas Daniilidis Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA, USA

Alexei A. Efros Department of Electrical Engineering and Computer Science, UC Berkeley, Berkeley, CA, USA

Pascal Fua EPFL, Lausanne, Switzerland

Kiran Gunda Object Video, Inc., Reston, VA, USA

Himaanshu Gupta Object Video, Inc., Reston, VA, USA

Asaad Hakeem Decisive Analytics Corporation, Arlington, VA, USA

James Hays Brown University, Providence, RI, USA

Martial Hebert Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA

Varsha Hedau Apple, Cupertino, CA, USA

James C. Hoe Carnegie Mellon University, Pittsburgh, PA, USA

Daniel P. Huttenlocher Cornell University, Ithaca, NY, USA

Rahul Kumar Jha University of Michigan, Ann Arbor, USA

Leif Kobbelt RWTH Aachen University, Aachen, Germany

Jana Košecká George Mason University, Fairfax, VA, USA

Kevin Köser GEOMAR, Kiel, Germany

Neeraj Kumar Department of Computer Science and Engineering, University of Washington, Seattle, WA, USA

L'ubor Ladický ETH Zürich, Zürich, Switzerland

Stefan Lee Indiana University, Bloomington, IN, USA

Yong Jae Lee Department of Computer Science, UC Davis, Davis, CA, USA

Bastian Leibe RWTH Aachen University, Aachen, Germany

Li-Jia Li Yahoo! Research, Sunnyvale, USA

Yunpeng Li École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Tsung-Yi Lin Cornell University, Ithaca, NY, USA

Eriko Nurvitadhi Intel Corporation, Hillsboro, OR, USA

Hyun Soo Park Carnegie Mellon University, Pittsburgh, PA, USA

Minwoo Park Object Video, Inc., Reston, VA, USA

Marc Pollefeys ETH Zürich, Zürich, Switzerland

Gang Qian Object Video, Inc., Reston, VA, USA

Bryan Russell Adobe Research, Lexington, KY, USA

Torsten Sattler Department of Computer Science, ETH Zürich, Zürich, Switzerland

Olivier Saurer ETH Zürich, Zürich, Switzerland

Harpreet Sawhney Center for Vision Technologies, SRI International, Princeton, NJ, USA

Steven Seitz Department of Computer Science and Engineering, University of Washington, Seattle, WA, USA

- Khurram Shafique** Object Video, Inc., Reston, VA, USA
- Mubarak Shah** University of Central Florida, Orlando, FL, USA
- David Ayman Shamma** Yahoo! Research, San Francisco, USA
- Yaser Sheikh** Carnegie Mellon University, Pittsburgh, PA, USA
- Gautam Singh** George Mason University, Fairfax, VA, USA
- Sudipta N. Sinha** Microsoft Research, Redmond, WA, USA
- Josef Sivic** Inria, WILLOW Project-team, Département d'Informatique de l'Ecole Normale Supérieure, ENS/INRIA/CNRS UMR, Paris, France
- Noah Snaveley** Cornell University, Ithaca, NY, USA
- Richard Szeliski** Facebook, Seattle, WA, USA
- Bart Thomee** Yahoo! Research, San Francisco, USA
- Raphael Townshend** Stanford University, Stanford, CA, USA
- Eric Tzeng** UC Berkeley, Berkeley, CA, USA
- Luc Van Gool** ETH Zurich, Zurich, Switzerland
- Nick Vander Valk** Vision Technologies Lab, SRI International, Princeton, NJ, USA
- Yang Wang** University of Manitoba, Winnipeg, Canada
- Yu Wang** Carnegie Mellon University, Pittsburgh, PA, USA
- Avideh Zakhor** UC Berkeley, Berkeley, CA, USA
- Amir R. Zamir** Stanford University, Stanford, CA, USA
- Andrew Zhai** UC Berkeley, Berkeley, CA, USA
- Jiejie Zhu** Vision Technologies Lab, SRI International, Princeton, NJ, USA
- C. Lawrence Zitnick** Facebook AI Research, Palo Alto, CA, USA