

Multimodal Location Estimation of Videos and Images

Jaeyoung Choi · Gerald Friedland
Editors

Multimodal Location Estimation of Videos and Images

Editors

Jaeyoung Choi
Gerald Friedland
International Computer Science Institute
Berkeley, CA
USA

ISBN 978-3-319-09860-9 ISBN 978-3-319-09861-6 (eBook)
DOI 10.1007/978-3-319-09861-6

Library of Congress Control Number: 2014951051

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2015

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

Preface

With the widespread use of GPS-equipped handheld devices, location metadata (a.k.a geotags) has rapidly become an integral part of photos and videos shared over the Web. This trend enabled location-based multimedia organization, search, and retrieval on many Internet services such as Google, Facebook, and Flickr. The main driving force behind these services is the creation of highly personalized user experiences, allowing for better recommendations and targeted advertisements. Even with this trend, it has been estimated that only about 5 % of the existing multimedia content on the Internet is actually geotagged. A significant amount of consumer-produced media content is still obtained using devices that do not have GPS functionality. Privacy concerns have motivated users to disable automatic geotagging of media. Furthermore, even GPS-enabled devices cannot provide accurate location information when the photo or video was captured in an indoor environment.

Nevertheless, the volume of high-quality geotagged videos and photos on the Web represents a quantity of training data for machine learning on an unprecedented scale, giving rise to the idea of creating an automated task that would try to locate non-geotagged media from the Web using models obtained through the geotagged subset. Put simply: Given a video and its associated textual metadata, can we infer the location where it was taken? This idea of “multimodal video location estimation” was proposed several years ago by the authors of the book.

Since then, the “Placing Task” of the MediaEval evaluation evaluated the task on a global scale and the United States Government has sponsored the research in their Finder program with a separate set of evaluations conducted by National Institute of Standards and Technology (NIST). As a result, the problem has been approached with diverse methods and ideas in the research community and significant improvements have been made. Multimodal Location Estimation has become a powerful tool and accuracies now come close to human capabilities.

The goal of the book is to present an overview of this field to software developers, engineers, and researchers and to bring together the different communities

working in the area. Apart from research interest, forensics experts, developers, and engineers for targeted advertising tools, as well as many people working in social media retrieval have become interested in the subject.

Berkeley, CA, USA, June 2014

Jaeyoung Choi
Gerald Friedland

Acknowledgments

First and foremost, we would like to thank our families. Jaeyoung is very grateful to his parents and his sister for their tireless support and encouragement. They have always supported him in everything he did, and for this he wants to thank them most of all. Gerald would like to thank his wife Yvonne and his daughter Mona for supporting him through this project and so many others over the years.

Jaeyoung is indebted to Jiwon Kim, Choongbum Lee, Jennifer Jinju Lee, Eugene Suh, Michelle Jeung-Eun Lee, and Jun Woo Lim, who were there for him when he needed them the most while going through crisis during his years in the U.S. To my good friends/housemates, Kang Wook Lee, Chan Kim, and Eunkwang Joo; and Seunghoon Lee, former captain of KGSA soccer team, for laying all the groundwork for shaping ourselves into a better team, Euiyoung Kim, Kunwoo Lee, and Grace Jeon; It would have been a real boring life here if it weren't for dear friends that I've met at Berkeley: Sunghwan Kim, Sun Choi, Kevin Ahn, Ilhyung Lee, Insoon Yang, Shinhye Choi, and Hyojung Shin. All the good and bad moments we've shared and went through together won't be forgotten. Korean folks at CS department have given me precious advices in many aspects. I'm grateful for the time I got to spend with Gunho Lee, Minkyung Kang, Chang-seo Park, Wontae Choi, Woojong Ko, Sangjin Han, and Yunsup Lee.

We are both grateful to our friends and colleagues at ICSI who directly or indirectly helped and supported the research work put into this book, in particular; Nelson Morgan, who provided us a research home at ICSI, Adam Janin, Arlo Faria, Benjamin Elizalde, Luke Gottlieb, T.J. Tsai, Howard Lei, Oriol Vinyals, Trevor Darrell, Julia Bernd, Bryan Morgan, and Michael Ellsworth.

We were privileged to be collaborating with gifted minds of Berkeley BASICS Group, Venky Ekambaram, Giulia Fanti, and Kannan Ramchandran.

This work would not have happened without the steady encouragement of Martha Larson from TU Delft, who always kept a firm grip on making MediaEval a fun, organically organized event. The resulting highly constructive, noncompetitive climate of the evaluation resulted in some of the most insightful discussions with colleagues we ever had.

Throughout the process of writing this book, many individuals from the community have taken time out to help us out. We'd like to give special thanks to the MediaEval and GeoMM community for actively participating in the feedback and contributions for this book. The book is a collection of quality research works from many institutions and individuals. Special thanks to all the authors and contributors of the guest chapters for their hard work.

Last but not least, we would like to acknowledge our sponsors. Initial experiments of this work were supported by an NGA NURI grant #HM11582-10-1-0008 and Korean Foundation for Advanced Studies. Later work was supported by National Science Foundation Award No. 1065240 and also National Science Foundation Award No. 1251276.

Contents

| | | |
|----------|---|------------|
| 1 | Introduction | 1 |
| | Jaeyoung Choi and Gerald Friedland | |
| 2 | The Benchmark as a Research Catalyst: Charting the Progress of Geo-prediction for Social Multimedia | 5 |
| | Martha Larson, Pascal Kelm, Adam Rae, Claudia Hauff, Bart Thomee, Michele Trevisiol, Jaeyoung Choi, Olivier Van Laere, Steven Schockaert, Gareth J.F. Jones, Pavel Serdyukov, Vanessa Murdock and Gerald Friedland | |
| 3 | Large-Scale Image Geolocalization | 41 |
| | James Hays and Alexei A. Efros | |
| 4 | Vision-Based Fine-Grained Location Estimation | 63 |
| | Heng Liu, Tao Mei, Houqiang Li and Jiebo Luo | |
| 5 | Image-Based Positioning of Mobile Devices in Indoor Environments | 85 |
| | Jason Zhi Liang, Nicholas Corso, Eric Turner and Avidesh Zakhori | |
| 6 | Application of Large-Scale Classification Techniques for Simple Location Estimation Experiments | 101 |
| | Howard Lei, Jaeyoung Choi and Gerald Friedland | |
| 7 | Collaborative Multimodal Location Estimation of Consumer Media | 115 |
| | Venkatesan Ekambaram, Kannan Ramchandran, Jaeyoung Choi and Gerald Friedland | |

| | | |
|-----------|---|------------|
| 8 | Georeferencing Flickr Resources Based on Multimodal Features | 127 |
| | Pascal Kelm, Sebastian Schmiedeke, Steven Schockaert, Thomas Sikora, Michele Trevisiol and Olivier Van Laere | |
| 9 | Human Versus Machine: Establishing a Human Baseline for Multimodal Location Estimation | 153 |
| | Jaeyoung Choi, Howard Lei, Venkatesan Ekambaram, Pascal Kelm, Luke Gottlieb, Thomas Sikora, Kannan Ramchandran and Gerald Friedland | |
| 10 | Personalized Travel Navigation and Photo-Shooting Navigation Using Large-Scale Geotags | 173 |
| | T. Yamasaki, A. Gallagher, T. Chen and K. Aizawa | |

Contributors

T. Chen Department of Information and Communication Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

Jaeyoung Choi International Computer Science Institute, Berkeley, CA, USA; California State University, East Bay, CA, USA

Nicholas Corso Department of EECS, UC Berkeley, Berkeley, CA, USA

Alexei A. Efros University of California, Berkeley, CA, USA

Venkatesan Ekambaram University of California, Berkeley, CA, USA

Gerald Friedland International Computer Science Institute, Berkeley, CA, USA; California State University, East Bay, CA, USA

A. Gallagher School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA

Luke Gottlieb International Computer Science Institute, Berkeley, CA, USA

Claudia Hauff Delft University of Technology, Delft, The Netherlands

James Hays Brown University, Providence, RI, USA

Gareth J.F. Jones Dublin City University, Dublin, Ireland

Pascal Kelm Technische Universität, Berlin, Germany

Martha Larson Delft University of Technology, Delft, The Netherlands

Howard Lei International Computer Science Institute, Berkeley, CA, USA; California State University, East Bay, CA, USA

Houqiang Li Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China

Jason Zhi Liang Department of EECS, UC Berkeley, Berkeley, CA, USA

Heng Liu Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China

Jiebo Luo Department of Computer Science, University of Rochester, Rochester, NY, USA

Tao Mei Microsoft Research, Beijing, China

Vanessa Murdock Microsoft, Bellevue, WA, USA

Adam Rae Future Cities Catapult, London, UK

Kannan Ramchandran University of California, Berkeley, CA, USA

Sebastian Schmiedeke Technische Universität, Berlin, Germany

Steven Schockaert Cardiff University, Cardiff, UK

Pavel Serdyukov Yandex, Moscow, Russia

Thomas Sikora Technische Universität, Berlin, Germany

Bart Thomee Yahoo Labs, San Francisco, CA, USA

Michele Trevisiol Yahoo Labs, Pompeu Fabra University, Barcelona, Spain

Eric Turner Department of EECS, UC Berkeley, Berkeley, CA, USA

Olivier Van Laere Yahoo Labs, Barcelona, Spain

T. Yamasaki Department of Information and Communication Engineering, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

Avideh Zakhor Department of EECS, UC Berkeley, Berkeley, CA, USA