

Computer Communications and Networks

For further volumes:

<http://www.springer.com/series/4198>

The **Computer Communications and Networks** series is a range of textbooks, monographs and handbooks. It sets out to provide students, researchers and non-specialists alike with a sure grounding in current knowledge, together with comprehensible access to the latest developments in computer communications and networking.

Emphasis is placed on clear and explanatory styles that support a tutorial approach, so that even the most complex of topics is presented in a lucid and intelligible manner.

Naeem Ramzan • Roelof van Zwol • Jong-Seok Lee
Kai Clüver • Xian-Sheng Hua

Editors

Social Media Retrieval

 Springer

Editors

Naeem Ramzan
School of Electronic Engineering
and Computer Science
Queen Mary University of London
London
United Kingdom

Roelof van Zwol
Director of Product Innovation, Search
Netflix
Los Gatos, California
USA

Jong-Seok Lee
School of Integrated Technology
Yonsei University
Incheon
Korea, Republic of (South Korea)

Kai Clüver
Institut für Telekommunikationssysteme
Technische Universität Berlin
Berlin
Germany

Xian-Sheng Hua
Media Computing Group
Microsoft Research
Bellevue, Washington
USA

Series Editor

A.J. Sammes
Centre for Forensic Computing
Cranfield University
Shrivenham campus
Swindon, UK

ISSN 1617-7975

ISBN 978-1-4471-4554-7

ISBN 978-1-4471-4555-4 (eBook)

DOI 10.1007/978-1-4471-4555-4

Springer London Heidelberg New York Dordrecht

Library of Congress Control Number: 2012953932

© Springer-Verlag London 2013

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

Multimedia content has become ubiquitous on the web, creating new challenges for indexing, access, search, and retrieval. At the same time, much of this content is made available on content-sharing websites such as YouTube or Flickr or shared on social networks like Facebook. In such environments, the content is usually accompanied with metadata, tags, ratings, comments, information about the uploader and their social network, etc. Analysis of these “social media” shows a great potential in improving the performance of traditional multimedia information analysis/retrieval approaches by bridging the semantic gap between the “objective” multimedia content analysis and “subjective” users’ needs and impressions. The integration of these aspects, however, is non-trivial and has created a vibrant, interdisciplinary field of research.

This book presents in-depth knowledge that explicitly exploits the synergy between multimedia content analysis, personalisation, and next-generation networking and community aspects of social networks. We believe that this integration could result in robust, personalised multimedia services, providing users with an improved multimedia centric quality of experience (QoE) awareness. In response to the booming developments in social networks, this book offers a practical step-by-step walk through the various challenges, concepts, components, and technologies involved in the development of applications and services. Researchers and students interested in the field of social media retrieval will find this book a valuable resource, covering a broad overview of current state of the art, current research, and development trends in this area. Practical engineers from industry will find this book useful in envisioning and building innovative social media applications and services.

This book is divided in four major parts. The first part introduces the fundamentals of social media retrieval by presenting the most important area of research in this domain. The second part covers an essential area of multimedia tagging in social environment. Personalisation and privacy in social media domain is discussed in the third part. In the last part, applications related to social media is presented.

The aim of the Part I “Fundamentals of Social Media” is to give a comprehensive overview on the state of the art, the challenges, and the potential of social

media retrieval. This part opens with chapter “Social Video Retrieval: Research Methods in Controlling, Sharing, and Editing of Web Video” by K. Choriantopoulos, D.A. Shamma, and L. Kennedy. It presents a survey of the state of the art in social video retrieval areas. Three case studies are presented, and the findings through these case studies are discussed in detail with the perspective of few future research directions. Chapter “Social Media Recommendation” by Z. Wang, W. Zhu, P. Cui, L. Sun, and S. Yang presents the framework of social media recommendation, with a focus on two important types of recommendations: interest-oriented social media recommendation and influence-oriented social media recommendation. The following chapter “Multimedia Indexing, Search and Retrieval in Large Databases of Social Networks” by T. Semertzidis, D. Rafailidis, E. Tiakas, M.G. Strintzis, and P. Daras addresses multimedia indexing methods and a useful case study that demonstrates the challenges faced by the content-based multimedia retrieval research community. To address such challenges, the chapter focuses on state-of-the-art strategies that target the huge volume of heterogeneous data offered by social networks. Chapter “Survey on Social Community Detection” by M. Plantie and M. Crampes presents a review/survey of the many different methods in the literature of community detection based on semantics, output type, and the evaluation of communities. The last chapter of the Part I “Detecting Multimedia Contents of Social Events in Social Networks” by M. Rabbath and S. Boll summarises the state of the art of event detection and clustering approaches. It deals with several types of social media platforms and how do they deal with events. It also shows the features used in the multimedia content analysis to support event detection including metadata, visual, people, and structural-based features.

Part II of the book addresses converting social media into implicitly or explicitly tagging (textual description). It has long been a dream of a number of multimedia, computer vision, and machine-learning researchers that gain popularity exponentially due to the explosion of social media on the Internet, bringing us both challenges and opportunities. Online services based on users’ geographical location are becoming very popular. Chapter “Georeferencing in Social Networks” by P. Kelm, V. Murdock, S. Schmiedeke, S. Schockaert, P. Serdyukov, and O. Van Laere provides an overview of issues in extracting and exploiting the geographical information from contents in social networks. By broadening the topic towards general content annotation, chapter “Predicting User Tags in Social Media Repositories Using Semantic Expansion and Visual Analysis” by T. Piatrik, Q. Zhang, X. Sevillano, and E. Izquierdo discusses how tags for multimedia contents can be automatically predicted based on the associated textual metadata and visual features and complementary information in external resources. Chapter “A Rule-Based Flickr Tag Recommendation System” by L. Cagliero, A. Fiori, and L. Grimaudo focuses on a rule-based tag recommendation approach that is robust to keyword data sparsity and able to capture different viewpoints of tags. In chapter “Sentic Computing for Social Media Analysis, Representation and Retrieval” by E. Cambria, M. Grassi, S. Poria, and A. Hussain, the concept of sentic computing, which is based on simultaneous usage of common sense knowledge and emotional information, is applied to bridge the gap between word-level natural language

data and concept-level opinions conveyed by these. Chapter “Highlight Detection in Movie Scenes Through Inter-users, Physiological Linkage” by C. Chenes, G. Chanel, M. Soleymani, and T. Pun shows how the problem of automatic video summarisation can be solved by aggregating users’ physiological signals, based on the assumption that synchronised emotional excitation of users provides the information about the importance of the video segment corresponding to the time instance. Chapter “Toward Emotional Annotation of Multimedia Contents” by A. Yazdani, J.-S. Lee, and T. Ebrahimi also deals with the problem of multimedia annotation via emotion recognition using physiological signals, which is called implicit tagging. The potential of the electroencephalogram (EEG), peripheral physiological signals, and content features for extracting emotional aspects of contents is examined.

Part III covers different aspects of privacy and personalisation of social media. It is introduced by a chapter “Privacy in Recommender Systems” by A. Jeckmans, M. Beye, Z. Erkin, P. Hartel, R. Lagendijk, and Q. Tang. It discusses existing recommender systems, data that is used in recommender systems, risks to data privacy, privacy-protection techniques in the literature, and how they apply to recommender systems. Chapter “Geotag Propagation with User Trust Modeling” by I. Ivanov, P. Vajda, J.-S. Lee, P. Korshunov, and T. Ebrahimi presents an overview of methods for user trust modelling in automatic geotags propagation and then presents and evaluates their own approach for automatic geotagging. Chapter “Context-Aware Content Adaptation for Personalised Social Media Access” by H.K. Arachchi and S. Dogan elaborates issues related to content adaptation for personalised access to social media. The importance of using context-aware content adaptation is stressed, and several signal processing-based techniques for actual content adaptation are presented. Both technological and non-technological challenges for performing content adaptation in social media are also addressed.

Part IV of the book focuses on the different applications that exploit social media to enhance the overall acceptability of the system. Exchange of information in social media would not be possible without the enormous progress of communication techniques. First, efficient methods for flexible encoding and compression, especially of bitrate-intensive video data, are a prerequisite for acceptable distribution and retrieval speeds is explained. Further, intelligent network architectures ensure a minimum quality of experience (QoE) and help establish new applications within social networks. This part starts with the L. Anania from European Commission (EC) statement is elaborated how the advancement of social media is supported by EC. The following chapter “Video Technology for Storage and Distribution of Personalised Media” by G.V. Wallendael, J.D. Cock, D.V. Deursen, M. Mrak, and R.V. de Walle deals with scalable coding of video. The authors give an overview of current video compression techniques with the focus on strategies for scaling and adaptation of content delivery. The use of metadata for personalised adaptation of media is discussed, and standardisation efforts as well as test results are presented. Dealing with networking aspects, chapter “Social Aware TV Content Delivery Over Intelligent Networks” by F. Fraile, P. Arce, R. Belda, I. de Fez, J.C. Guerri, and A. Pajares discusses IP network architectures, quality criteria,

social aspects of TV services, and socially aware techniques for content delivery. Motivated by the increasing demand for video delivery, the authors present several approaches for ensuring QoE in video transmission by dynamic management of network resources. Video watching over a distributed social network requires synchronisation in order to ensure a shared experience for the participants. Chapter “Distributed Media Synchronisation for Shared Video Watching: Issues, Challenges and Examples” by Fernando Boronat, Rufael Mekuria, Mario Montagud, and Pablo Cesar presents an overview of the problem, results of measurements and subjective tests, a comprehensive description of possible synchronisation schemes, and an own proposal which is currently in the process of standardisation. Chapter “eGuided: Sharing Media in Academic and Social Networks Based on Peer-Assisted Learning e-Portfolios” by Paulo N.M. Sampaio, Rúben H. de Freitas Gouveia, and Pedro A.T. Gomes presents a software platform for peer-assisted learning. A student’s “e-portfolio” may be used and/or edited by its owner, teachers, peers, etc., in order to assess the student’s skills and weaknesses and derive suitable learning strategies. The chapter includes a discussion of different peer-assisted learning methods and of the authors’ approach, together with a comparison of existing e-portfolio platforms. The last chapter of the book, “Exploiting Social Media for Music Information Retrieval” by M. Schedl, deals with an application of data mining, which deals with music information retrieval in social media. The chapter gives a comprehensive overview of approaches, describing similarity and popularity estimation and auto-tagging methods, including the results of recent work in the field.

We would like to thank all the authors and reviewers for their excellent original contributions and essential help in providing expert opinions and comments on the numerous chapter submitted to this book, respectively. It is our hope that the readers of this book will get useful knowledge, ideas, and insights for their own research.

Naeem Ramzan
Roelof van Zwol
Jong-Seok Lee
Kai Clüver
Xian-Sheng Hua

Contents

Part I Fundamentals of Social Media

Social Video Retrieval: Research Methods in Controlling, Sharing, and Editing of Web Video	3
Konstantinos Chorianopoulos, David A. Shamma, and Lyndon Kennedy	
Social Media Recommendation	23
Zhi Wang, Wenwu Zhu, Peng Cui, Lifeng Sun, and Shiqiang Yang	
Multimedia Indexing, Search, and Retrieval in Large Databases of Social Networks	43
Theodoros Semertzidis, Dimitrios Rafailidis, Eleftherios Tiakas, Michael G. Strintzis, and Petros Daras	
Survey on Social Community Detection	65
Michel Plantié and Michel Crampes	
Detecting Multimedia Contents of Social Events in Social Networks	87
Mohamad Rabbath and Susanne Boll	

Part II Tagging of Social Media

Georeferencing in Social Networks	115
Pascal Kelm, Vanessa Murdock, Sebastian Schmiedeke, Steven Schockaert, Pavel Serdyukov, and Olivier Van Laere	
Predicting User Tags in Social Media Repositories Using Semantic Expansion and Visual Analysis	143
Tomas Piatrik, Qianni Zhang, Xavier Sevillano, and Ebroul Izquierdo	

A Rule-Based Flickr Tag Recommendation System	169
Luca Cagliero, Alessandro Fiori, and Luigi Grimaudo	
Sentic Computing for Social Media Analysis, Representation, and Retrieval	191
Erik Cambria, Marco Grassi, Soujanya Poria, and Amir Hussain	
Highlight Detection in Movie Scenes Through Inter-users, Physiological Linkage	217
Christophe Chênes, Guillaume Chanel, Mohammad Soleymani, and Thierry Pun	
Toward Emotional Annotation of Multimedia Contents	239
Ashkan Yazdani, Jong-Seok Lee, and Touradj Ebrahimi	
Part III Privacy and Personalisation of Social Media	
Privacy in Recommender Systems	263
Arjan J.P. Jeckmans, Michael Beye, Zekeriya Erkin, Pieter Hartel, Reginald L. Lagendijk, and Qiang Tang	
Geotag Propagation with User Trust Modeling	283
Ivan Ivanov, Peter Vajda, Jong-Seok Lee, Pavel Korshunov, and Touradj Ebrahimi	
Context-Aware Content Adaptation for Personalised Social Media Access	305
Hemantha Kodikara Arachchi and Safak Dogan	
Part IV Applications and Services	
Research in Social Media: How the EC Facilitates R&D Innovation	341
Loretta Anania	
Video Technology for Storage and Distribution of Personalised Media	347
Glenn Van Wallendael, Jan De Cock, Davy Van Deursen, Marta Mrak, and Rik Van de Walle	
Social Aware TV Content Delivery Over Intelligent Networks	373
Francisco Fraile, Pau Arce, Román Belda, Ismael de Fez, Juan Carlos Guerri, and Ana Pajares	
Distributed Media Synchronisation for Shared Video Watching: Issues, Challenges and Examples	393
Fernando Boronat, Rufael Mekuria, Mario Montagud, and Pablo Cesar	

**eGuided: Sharing Media in Academic and Social Networks
Based on Peer-Assisted Learning e-Portfolios** 433
Paulo N.M. Sampaio, Rúben H. de Freitas Gouveia,
and Pedro A.T. Gomes

Exploiting Social Media for Music Information Retrieval 449
Markus Schedl

Index 479