Series Editor

W. Bruce Croft

Editorial Board

ChengXiang Zhai
Maarten de Rijke
Nicholas J. Belkin
Charles Clarke

Henning Müller • Paul Clough • Thomas Deselaers •
Barbara Caputo
Editors

ImageCLEF

Experimental Evaluation in Visual Information Retrieval

Springer

Editors
Henning Müller
Business Information Systems
University of Applied Sciences
Western Switzerland (HES–SO)
TechnoArk 3
3960 Sierre
Switzerland
henning.mueller@hevs.ch

Paul Clough
Information School
University of Sheffield
Regent Court
Sheffield S1 4DP
England
p.d.clough@sheffield.ac.uk

Thomas Deselaers
ETH Zürich
Computer Vision Lab/ETF-C 113.2
Zürich
Switzerland
deselaers@vision.ee.ethz.ch

Barbara Caputo
Idiap Research Institute
rue Marconi 19
1920 Martigny
Switzerland
bcaputo@idiap.ch

ISSN 1387-5264 The Information Retrieval Series
ISBN 978-3-642-15180-4 e-ISBN 978-3-642-15181-1
DOI 10.1007/978-3-642-15181-1
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2010934959

ACM Classification (1998): H.3, I.4, C.4, J.3

© Springer-Verlag Berlin Heidelberg 2010
This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.
The use of general descriptive names, registered names, trademarks, 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.

Cover design: KünkelLopka GmbH, Heidelberg

Printed on acid-free paper

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

This book is dedicated to our families and the love, support and encouragement they have given us.

Foreword

The pervasive creation and consumption of content, especially visual content, is ingrained into our modern world. We're constantly consuming visual media content, in printed form and in digital form, in work and in leisure pursuits. Like our cave–man forefathers, we use pictures to record things which are of importance to us as memory cues for the future, but nowadays we also use pictures and images to document processes; we use them in engineering, in art, in science, in medicine, in entertainment and we also use images in advertising. Moreover, when images are in digital format, either scanned from an analogue format or more often than not born digital, we can use the power of our computing and networking to exploit images to great effect.

Most of the technical problems associated with creating, compressing, storing, transmitting, rendering and protecting image data are already solved. We use accepted standards and have tremendous infrastructure and the only outstanding challenges, apart from managing the scale issues associated with growth, are to do with locating images. That involves analysing them to determine their content, classifying them into related groupings, and searching for images. To overcome these challenges we currently rely on image metadata, the description of the images, either captured automatically at creation time or manually added afterwards. Meanwhile we push for developments in the area of *content–based* analysis, indexing and searching of visual media and this is where most of the research in image management is concentrated.

Automatic analysis of the content of images, which in turn would open the door to content–based indexing, classification and retrieval, is an inherently tough problem and because of the difficulty, progress is slow. Like all good science it cannot be rushed yet there is a frustration with the pace of its development because the rollout and development of other related components of image management, components such as capture, storage, transmission, rendering, etc., has been so rapid. We seem to be stuck on the problems of how to effectively find images when we are looking for them. While this is partly caused by the sheer number of images available to us, it is mostly caused by the scientific difficulty of the challenge and so it requires a

basic scientific approach to exploring the problem and finding solutions. As in all science, a fundamental aspect is measurement and benchmarking.

In any science, each new development, each approach, algorithm, model, idea or theory has to be measured in order to determine its worth and validity. That is how progress is made, and how fields advance. A theory is put forward and experiments to assess and measure the theory are carried out which may or may not support the theory and we advance the field, either by learning more about what works, or equally important we learn about what does not work. In the technology sector and in the information management area in particular, measuring the validity and worth of new ideas, approaches, etc., now takes place as part of organised benchmarking activities and there are many established examples. The Pascal Visual Object Class recognition challenge addresses recognising objects in images, TRECVid addresses content analysis, retrieval and summarization from video, the KDD competition addresses data mining, and there have been others in machine learning for stock market prediction, shape retrieval, coin classification, text detection and reading, face verification, fingerprint verification, signature verification and of course the well–known NetFlix data mining and recommender competition. All these and many others take place against a backdrop of exploring new ideas, new approaches, and measuring their efficacy in a controlled environment. Which takes us to the present volume which covers ImageCLEF.

Cross–language image retrieval is a niche application domain within the broader area of managing image/visual media. Its importance is huge, though given the directions in which Internet growth is heading with multi–linguality and cross–language resources and processes growing ever more important. Henning Müller and Paul Clough have put together an impressive collection of contributions describing the formation, the growth, the resources, the various tasks and achievements of the ImageCLEF benchmarking activity, covering seven years of development in an annual cycle and involving contributions from hundreds of researchers from across the globe. This book could be described as a capstone volume which brings together all the contributions into one place, but a capstone is a finishing stone or a final achievement, and ImageCLEF continues today, as active as ever. With four parts which address the settings and logistics of ImageCLEF, the various track reports, some reports from participants and finally some external views, the volume is balanced and presents a comprehensive view of the importance and achievements of ImageCLEF towards advancing the field of cross–lingual image retrieval. It will remain an essential reference for anybody interested in how to start up and run a sizeable benchmarking activity, as well as an invaluable source of information on image retrieval in a cross–lingual setting.

Alan F. Smeaton
CLARITY: Centre for Sensor Web Technologies
Dublin City University
Dublin, Ireland, May 2010

Preface

This book contains a collection of texts centred on the evaluation of image retrieval systems. Evaluation, whether it be system–oriented or user–oriented, is an important part of developing effective retrieval systems that meet the actual needs of their end users. To enable reproducible evaluation requires creating standardised benchmarks and evaluation methodologies. This book highlights some of the issues and challenges in evaluating image retrieval systems and describes various initiatives that have sought to provide researchers with the necessary evaluation resources.

In particular the book summarises activities within ImageCLEF, an initiative to evaluate cross–language image retrieval systems that has been running as part of the Cross Language Evaluation Forum (CLEF) since 2003. ImageCLEF has provided resources, such as benchmarks, for evaluating image retrieval systems and complements a number of initiatives within the image retrieval research community, such as TRECvid for video retrieval, PASCAL for object recognition and detection and the many other smaller benchmarks, databases and tools available to researchers.

In addition to providing evaluation resources, ImageCLEF has also run within an annual evaluation cycle culminating in a workshop where participants have been able to present and discuss their ideas and techniques, forming a community with common interests and goals. Over the years ImageCLEF has seen participation from researchers within academic and commercial research groups worldwide, including those from Cross–Language Information Retrieval (CLIR), medical informatics, Content–Based Image Retrieval (CBIR), computer vision and user interaction.

This book comprises contributions from a range of people: those involved directly with ImageCLEF, such as the organisers of specific image retrieval or annotation tasks; participants who have developed techniques to tackle the challenges set forth by the organisers; and people from industry and academia involved with image retrieval and evaluation in general and beyond ImageCLEF. The book is structured into four parts:

- **Part I.** This section describes the context of ImageCLEF and the issues involved with developing evaluation resources, such as test collections and selecting evaluation measures. A focal point throughout ImageCLEF and across many of the

tasks has been to investigate how best to combine textual and visualisation information to improve information retrieval. Within the first section we summarise approaches explored within ImageCLEF over the years for this critical step in the retrieval process.

- **Part II.** This section includes seven chapters summarising the activities of each of the main tasks that have run within within ImageCLEF over the years. The track reports are written by those involved in co–ordinating ImageCLEF tasks and provide summaries of individual tasks, describe the participants and their approaches, and discuss some of the findings.
- **Part III.** This section is a selection of chapters by groups participating in various tasks within ImageCLEF 2009. Summaries of the techniques used for various domains such as retrieving diverse sets of photos from a collection of news photographs, multi–modal retrieval from online resources, such as Wikipedia, and retrieval and automatic annotation of medical images are presented. The chapters in this section show the variety and novelty of state–of–the–art techniques used to tackle various ImageCLEF tasks.
- **Part IV.** The final section provides an external perspective on the activities of ImageCLEF. These help to offer insights into the current and emerging needs for image retrieval and evaluation from both a commercial and research perspective. The final chapter helps to put ImageCLEF into the context of existing activities on evaluating multimedia retrieval techniques, providing thoughts on the future directions for evaluation over the coming years.

Sierre, Zürich, Martigny, Switzerland
Sheffield, UK
July 2010

Henning Müller
Paul Clough
Thomas Deselaers
Barbara Caputo

Acknowledgements

It is hard to know where to begin when considering who to thank because ImageCLEF has been very much a collaborative effort and involved a large number of people over the past seven years. We thank the UK Engineering and Physical Sciences Research Council (EPSRC) for funding the Eurovision project (grant number GR/R56778/01): a proposal written by Mark Sanderson that included running a large–scale benchmarking event for cross–language image retrieval that was to become ImageCLEF. In addition to the EPSRC, we thank all of the national and international funding bodies who have supported ImageCLEF in one way or another over the years. In particular we thank the European Union for support through the funding of various EU projects that have supported task organisers. We would like to acknowledge the following projects for supporting ImageCLEF activities:

- UK projects: Eurovision (EPSRC, grant GR/R56778/01)
- EU projects: MUSCLE, Multimatch, SemanticMining, DebutIT, Promise, TrebleCLEF, DELOS;
- Swiss projects: FNS (205321–109304/1), HES–SO for the BeMeVIS project;
- German projects: Theseus (grant 01MQ07017);
- companies such as Google (medical task) and LTU (for the object recognition task);
- American projects: NSF (grant ITR–0325160)

From the start ImageCLEF has been a collaborative activity and involved many people, many bringing new ideas, which is necessary in evaluation as it needs to advance as technology advances. People have given up their time and put tremendous effort into helping co–ordinate and run tasks. This has enabled us to produce re–useable resources for various image retrieval communities. The following is hopefully a complete list of those involved in the organisation of ImageCLEF: Thomas Arni, Peter Dunker, Thomas M. Deserno, Julio Gonzalo, Michael Grubinger, Allan Hanbury, William Hersh, Mark Huiskes, Patric Hensfelt, Charles Kahn, Jayashree Kalpathy–Cramer, Jussi Karlgren, Jana Kludas, Monica Lestari–Paramita, Stefanie Nowak, Adrian Popescu, Andrzej Pronobis, Mark Sanderson, Tatiana Tommasi and Theodora Tsikrika.

Besides those involved in the organisation and actually performing the work, we also need to thank all the data providers who have enabled us to distribute and use their content. In particular we acknowledge the help of Norman Reid from St. Andrews University Library in Scotland for providing us access to the historic set of photographs for the first ImageCLEF evaluation campaign. In addition we thank the following institutions: University of Basel (Pathopic), University of Geneva (Casimage), Mallinckrodt Institute of Radiology (MIR), RSNA, Flickr, viventura, Belga, Wikipedia, LTUtech, PASCAL, UCLA (HEAL, PEIR), OHSU (CORI), MyPACS, KTH, and the IRMA group.

Thanks go to all those involved in carrying out annotations and relevance assessments across the tasks without whom we would have no gold standard to benchmark against. Many people have given up their own time to generate relevance assessments and not been paid for their contributions. We also thank all the participants to ImageCLEF tasks. It is the participants who make an event such as ImageCLEF possible and without whose support and comments there would be no results to report. Over the seven years of ImageCLEF, approximately 200 groups have signed up for at least one for the tasks and over 100 groups have submitted results to one of the tasks in multi–lingual image retrieval.

We thank those who were involved from the start of ImageCLEF. In particular we thank Carol Peters who has provided CLEF with ten years of dedicated service and created a stable environment in which to run not just ImageCLEF, but many important evaluation tasks for comparing and improving multi–lingual information retrieval systems. We thank Carol for allowing us to include ImageCLEF within the CLEF activities and for her continual support and encouragement of our work. We also gratefully acknowledge the support of Donna Harman for her insightful comments and the discussions we had on evaluation methodologies and directions to follow.

Finally, in terms of this book we would like to thank Ralf Gerstner from Springer Verlag for his help and support, David Clough for proofreading the chapters and helping produce the final version of the book and all the contributing authors. We were perhaps a little naïve when fixing the publication deadlines and had to relieve some of the stress on our fellow writers. We apologise for this and thank everyone for staying with us despite the difficulties. We hope that you will enjoy reading your chapters now they are published.

Contents

Part I Introduction

1 Seven Years of Image Retrieval Evaluation 3
Paul Clough, Henning Müller, and Mark Sanderson
1.1 Introduction 3
1.2 Evaluation of IR Systems 5
1.2.1 IR Test Collections 6
1.2.2 Cross–Language Evaluation Forum (CLEF) 9
1.3 ImageCLEF 9
1.3.1 Aim and Objectives 9
1.3.2 Tasks and Participants 11
1.3.3 Data sets 12
1.3.4 Contributions 12
1.3.5 Organisational Challenges 14
1.4 Conclusions 15
References 16

2 Data Sets Created in ImageCLEF 19
Michael Grubinger, Stefanie Nowak, and Paul Clough
2.1 Introduction 19
2.1.1 Collection Creation 20
2.1.2 Requirements and Specification 21
2.1.3 Collection Overview 23
2.2 Image Collections for Photographic Retrieval 24
2.2.1 The St. Andrews Collection of Historic Photographs 24
2.2.2 The IAPR TC–12 Database 26
2.2.3 The Belga News Agency Photographic Collection 28
2.3 Image Collections for Medical Retrieval 29
2.3.1 The ImageCLEFmed Teaching Files 30
2.3.2 The RSNA Database 34
2.4 Automatic Image Annotation and Object Recognition 35

2.4.1 The IRMA Database 35
2.4.2 The LookThatUp (LTU) Data set 36
2.4.3 The PASCAL Object Recognition Database 37
2.4.4 The MIR Flickr Image Data Set 38
2.5 Image Collections in Other Tasks 38
2.5.1 The INEX MM Wikipedia Collection 39
2.5.2 The KTH–IDOL2 Database 40
2.6 Conclusions 41
References 42

3 Creating Realistic Topics for Image Retrieval Evaluation 45
Henning Müller
3.1 Introduction 45
3.2 User Models and Information Sources 48
3.2.1 Machine–Oriented Evaluation 48
3.2.2 User Models 49
3.2.3 Information Sources for Topic Creation 50
3.3 Concrete Examples for Generated Visual Topics in Several Domains 53
3.3.1 Photographic Retrieval 53
3.3.2 Medical Retrieval 54
3.4 The Influence of Topics on the Results of Evaluation 55
3.4.1 Classifying Topics Into Categories 56
3.4.2 Links Between Topics and the Relevance Judgments 57
3.4.3 What Can Be Evaluated and What Can Not? 57
3.5 Conclusions 58
References 59

4 Relevance Judgments for Image Retrieval Evaluation 63
Jayashree Kalpathy–Cramer, Steven Bedrick, and William Hersh
4.1 Introduction 63
4.2 Overview of Relevance Judgments in Information Retrieval 64
4.2.1 Test Collections 64
4.2.2 Relevance Judgments 65
4.3 Relevance Judging for the ImageCLEF Medical Retrieval Task . . . 72
4.3.1 Topics and Collection 72
4.3.2 Judges 73
4.3.3 Relevance Judgment Systems and the Process of Judging 74
4.4 Conclusions and Future Work 78
References 79

5 Performance Measures Used in Image Information Retrieval 81
Mark Sanderson
5.1 Evaluation Measures Used in ImageCLEF 81
5.2 Measures for Retrieval 82
5.2.1 Measuring at Fixed Recall 83

5.2.2 Measuring at Fixed Rank 85
5.2.3 Measures for Diversity 87
5.2.4 Collating Two Measures Into One 88
5.2.5 Miscellaneous Measures 88
5.2.6 Considering Multiple Measures 89
5.2.7 Measures for Image Annotation and Concept Detection . . 90
5.3 Use of Measures in ImageCLEF 91
5.4 Conclusions 92
References 92

6 Fusion Techniques for Combining Textual and Visual Information Retrieval 95
Adrien Depeursinge and Henning Müller
6.1 Introduction 95
6.1.1 Information Fusion and Orthogonality 97
6.2 Methods 98
6.3 Results 98
6.3.1 Early Fusion Approaches 98
6.3.2 Late Fusion Approaches 99
6.3.3 Inter–media Feedback with Query Expansion 104
6.3.4 Other Approaches 105
6.3.5 Overview of the Methods from 2004–2009 105
6.4 Justification for the Approaches and Generally Known Problems . . 105
6.5 Conclusions 108
References 108

Part II Track Reports

7 Interactive Image Retrieval 117
Jussi Karlgren and Julio Gonzalo
7.1 Interactive Studies in Information Retrieval 117
7.2 iCLEF Experiments on Interactive Image Retrieval 119
7.2.1 iCLEF Image Retrieval Experiments: The Latin Square Phase 120
7.2.2 iCLEF Experiments with Flickr 123
7.2.3 The Target Collection: Flickr 124
7.2.4 Annotations 124
7.2.5 The Task 125
7.2.6 Experiments 127
7.3 Task Space, Technology and Research Questions 134
7.3.1 Use Cases for Interactive Image Retrieval 134
7.3.2 Challenges: Technology and Interaction 135
References 137

8 Photographic Image Retrieval 141
Monica Lestari Paramita and Michael Grubinger
8.1 Introduction 141
8.2 Ad hoc Retrieval of Historic Photographs: ImageCLEF 2003–2005 142
8.2.1 Test Collection and Distribution 143
8.2.2 Query Topics 144
8.2.3 Relevance Judgments and Performance Measures 147
8.2.4 Results and Analysis 147
8.3 Ad hoc Retrieval of Generic Photographs: ImageCLEFphoto 2006-2007 149
8.3.1 Test Collection and Distribution 150
8.3.2 Query Topics 151
8.3.3 Relevance Judgments and Performance Measures 152
8.3.4 Results and Analysis 153
8.3.5 Visual Sub–task 154
8.4 Ad hoc Retrieval and Result Diversity: ImageCLEFphoto 2008–2009 155
8.4.1 Test Collection and Distribution 155
8.4.2 Query Topics 156
8.4.3 Relevance Judgments and Performance Measures 158
8.4.4 Results and Analysis 158
8.5 Conclusion and Future Prospects 160
References 161

9 The Wikipedia Image Retrieval Task 163
Theodora Tsikrika and Jana Kludas
9.1 Introduction 163
9.2 Task Overview 164
9.2.1 Evaluation Objectives 164
9.2.2 Wikipedia Image Collection 165
9.2.3 Additional Resources 165
9.2.4 Topics 166
9.2.5 Relevance Assessments 167
9.3 Evaluation 169
9.3.1 Participants 169
9.3.2 Approaches 170
9.3.3 Results 175
9.4 Discussion 179
9.4.1 Best Practices 179
9.4.2 Open Issues 180
9.5 Conclusions and the Future of the Task 181
References 181

10 The Robot Vision Task 185
Andrzej Pronobis and Barbara Caputo
10.1 Introduction 185
10.2 The Robot Vision Task at ImageCLEF 2009: Objectives and Overview 187
10.2.1 The Robot Vision Task 2009 188
10.2.2 Robot Vision 2009: The Database 188
10.2.3 Robot Vision 2009: Performance Evaluation 189
10.2.4 Robot Vision 2009: Approaches and Results 192
10.3 Moving Forward: Robot Vision in 2010 194
10.3.1 The Robot Vision Task at ICPR2010 194
10.3.2 The Robot Vision Task at ImageCLEF2010 196
10.4 Conclusions 197
References 197

11 Object and Concept Recognition for Image Retrieval 199
Stefanie Nowak, Allan Hanbury, and Thomas Deselaers
11.1 Introduction 199
11.2 History of the ImageCLEF Object and Concept Recognition Tasks 200
11.2.1 2006: Object Annotation Task 201
11.2.2 2007: Object Retrieval Task 202
11.2.3 2008: Visual Concept Detection Task 203
11.2.4 2009: Visual Concept Detection Task 204
11.3 Approaches to Object Recognition 204
11.3.1 Descriptors 206
11.3.2 Feature Post–processing and Codebook Generation 207
11.3.3 Classifier 207
11.3.4 Post–Processing 208
11.4 Results 208
11.4.1 2006: Object Annotation Task 209
11.4.2 2007: Object Retrieval Task 209
11.4.3 2008: Visual Concept Detection Task 210
11.4.4 2009: Visual Concept Detection Task 211
11.4.5 Evolution of Concept Detection Performance 213
11.4.6 Discussion 214
11.5 Combinations with the Photo Retrieval Task 215
11.6 Conclusion 215
References 216

12 The Medical Image Classification Task 221
Tatiana Tommasi and Thomas Deselaers
12.1 Introduction 221
12.2 History of ImageCLEF Medical Annotation 222
12.2.1 The Aim of the Challenge 222
12.2.2 The Database 223
12.2.3 Error Evaluation 227

12.3 Approaches to Medical Image Annotation 229
12.3.1 Image Representation 230
12.3.2 Classification Methods 230
12.3.3 Hierarchy 231
12.3.4 Unbalanced Class Distribution 231
12.4 Results 231
12.5 Conclusion 235
References 237

13 The Medical Image Retrieval Task 239
Henning Müller and Jayashree Kalpathy–Cramer
13.1 Introduction 239
13.2 Participation in the Medical Retrieval Task 240
13.3 Development of Databases and Tasks over the Years 242
13.3.1 2004 242
13.3.2 2005–2007 243
13.3.3 2008–2009 247
13.4 Evolution of Techniques Used by the Participants 249
13.4.1 Visual Retrieval 250
13.4.2 Textual Retrieval 250
13.4.3 Combining Visual and Textual Retrieval 251
13.4.4 Case–Based Retrieval Topics 251
13.5 Results 251
13.5.1 Visual Retrieval 252
13.5.2 Textual Retrieval 252
13.5.3 Mixed Retrieval 253
13.5.4 Relevance Feedback and Manual Query Reformulation . . 253
13.6 Main Lessons Learned 253
13.7 Conclusions 255
References 255

Part III Participant reports

14 Expansion and Re–ranking Approaches for Multimodal Image Retrieval using Text–based Methods 261
Adil Alpkocak, Deniz Kilinc, and Tolga Berber
14.1 Introduction 262
14.2 Integrated Retrieval Model 263
14.2.1 Handling Multi–modality in the Vector Space Model 264
14.3 Document and Query Expansion 265
14.4 Re–ranking 267
14.4.1 Level 1: Narrowing-down and Re-indexing 269
14.4.2 Level 2: Cover Coefficient Based Re–ranking 269
14.5 Results 271
14.6 Conclusions 273
References 274

15 Revisiting Sub–topic Retrieval in the ImageCLEF 2009 Photo Retrieval Task 277
Teerapong Leelanupab, Guido Zuccon, and Joemon M. Jose
15.1 Introduction 278
15.2 Background and Related Work 280
15.2.1 Sub–topic Retrieval 280
15.2.2 The Probability Ranking Principle 282
15.2.3 Beyond Independent Relevance 282
15.3 Document Clustering and Inter–Cluster Document Selection 284
15.3.1 Re–examining Document Clustering Techniques 284
15.3.2 Clustering for Sub–topic Retrieval 285
15.4 Empirical Study 287
15.5 Results 290
15.6 Conclusions 291
References 293

16 Knowledge Integration using Textual Information for Improving ImageCLEF Collections 295
Manuel Carlos Díaz–Galiano, Miguel Ángel García–Cumbreras, María Teresa Martín–Valdivia, and Arturo Montejo-Ráez
16.1 Introduction 295
16.2 System Description 297
16.2.1 Photo Retrieval System 297
16.2.2 Medical Retrieval System 298
16.3 Photo Task 298
16.3.1 Using Several IR and a Voting System 301
16.3.2 Filtering 302
16.3.3 Clustering 305
16.4 The Medical Task 306
16.4.1 Metadata Selection using Information Gain 306
16.4.2 Expanding with Ontologies 308
16.4.3 Fusion of Visual and Textual Lists 311
16.5 Conclusion and Further Work 311
References 313

17 Leveraging Image, Text and Cross–media Similarities for Diversity–focused Multimedia Retrieval 315
Julien Ah-Pine, Stephane Clinchant, Gabriela Csurka, Florent Perronnin, and Jean-Michel Renders
17.1 Introduction 315
17.2 Content–Based Image Retrieval 317
17.2.1 Fisher Vector Representation of Images 318
17.2.2 Image Retrieval at ImageCLEF Photo 320
17.3 Text Representation and Retrieval 321
17.3.1 Language Models 321
17.3.2 Text Enrichment at ImageCLEF Photo 322

17.4 Text–Image Information Fusion 326
17.4.1 Cross–Media Similarities 327
17.4.2 Cross–Media Retrieval at ImageCLEF Photo 329
17.5 Diversity–focused Multimedia Retrieval 332
17.5.1 Re–ranking Top–Listed Documents to Promote Diversity 333
17.5.2 Diversity–focused Retrieval at ImageCLEF Photo 336
17.6 Conclusion 339
References 340

18 University of Amsterdam at the Visual Concept Detection and Annotation Tasks 343
Koen E. A. van de Sande and Theo Gevers
18.1 Introduction 343
18.2 Concept Detection Pipeline 344
18.2.1 Point Sampling Strategy 345
18.2.2 Color Descriptor Extraction 346
18.2.3 Bag–of–Words model 347
18.2.4 Machine Learning 348
18.3 Experiments 349
18.3.1 Spatial Pyramid Levels 349
18.3.2 Point Sampling Strategies and Color Descriptors 350
18.3.3 Combinations of Sampling Strategies and Descriptors . . . 351
18.3.4 Discussion 353
18.4 ImageCLEF 2009 353
18.4.1 Evaluation Per Image 355
18.4.2 Conclusion 355
18.5 ImageCLEF@ICPR 2010 356
18.6 Conclusion 356
References 357

19 Intermedia Conceptual Indexing 359
Jean–Pierre Chevallet and Joo Hwee Lim
19.1 Introduction 359
19.2 Conceptual Indexing 361
19.2.1 Concept Usage and Definition in IR 361
19.2.2 Concept Mapping to Text 362
19.2.3 Mapping Steps 363
19.2.4 IR Models Using Concepts 366
19.2.5 Experiments using the ImageCLEF Collection 367
19.3 Image Indexing using a Visual Ontology 369
19.3.1 Image Indexing Based on VisMed Terms 370
19.3.2 FlexiTile Matching 373
19.3.3 Medical Image Retrieval Using VisMed Terms 374
19.3.4 Spatial Visual Queries 375
19.4 Multimedia and Intermedia Indexing 376

19.5 Conclusions 378
References 379

20 Conceptual Indexing Contribution to ImageCLEF Medical Retrieval Tasks 381
Loïc Maisonasse, Jean–Pierre Chevallet, and Eric Gaussier
20.1 Introduction 382
20.2 Semantic Indexing Using Ontologies 382
20.3 Conceptual Indexing 383
20.3.1 Language Models for Concepts 383
20.3.2 Concept Detection 384
20.3.3 Concept Evaluation Using ImageCLEFmed 2005–07 385
20.4 From Concepts to Graphs 386
20.4.1 A Language Model for Graphs 386
20.4.2 Graph Detection 387
20.4.3 Graph Results on ImageCLEFmed 2005–07 388
20.5 Mixing Concept Sources 388
20.5.1 Query Fusion 389
20.5.2 Document Model Fusion 389
20.5.3 Joint Decomposition 390
20.5.4 Results on ImageCLEFmed 2005–07 392
20.6 Adding Pseudo–Feedback 393
20.6.1 Pseudo–Relevance Feedback Model 393
20.6.2 Results 394
20.7 Conclusions 395
References 395

21 Improving Early Precision in the ImageCLEF Medical Retrieval Task 397
Steven Bedrick, Saïd Radhouani, and Jayashree Kalpathy–Cramer
21.1 Introduction 397
21.1.1 What is Early Precision? 398
21.1.2 Why Improve Early Precision? 399
21.2 ImageCLEF 399
21.3 Our System 400
21.3.1 User Interface 400
21.3.2 Image Database 401
21.3.3 Query Parsing and Indexing 402
21.4 Improving Precision 403
21.4.1 Modality Filtration 403
21.4.2 Using Modality Information for Retrieval 406
21.4.3 Using Interactive Retrieval 408
21.5 Conclusions 411
References 412

22 Lung Nodule Detection 415
Luca Bogoni, Jinbo Bi, Charles Florin, Anna K. Jerebko, Arun Krishnan, Sangmin Park, Vikas Raykar, and Marcos Salganicoff
22.1 Introduction 415
22.1.1 Lung Cancer — Clinical Motivation 416
22.1.2 Computer–Aided Detection of Lung Nodules 418
22.1.3 Ground Truth for Lesions 419
22.2 Review of Existing Techniques 420
22.2.1 Gray–Level Threshold 421
22.2.2 Template Matching 421
22.2.3 Spherical Enhancing Filters 422
22.3 Description of Siemens LungCAD System 423
22.3.1 Lung Segmentation 423
22.3.2 Candidate Generation 423
22.3.3 Feature Extraction 424
22.4 Classification 425
22.4.1 Multiple Instance Learning 425
22.4.2 Exploiting Domain Knowledge in Data–Driven Training–Gated Classifiers 426
22.4.3 Ground Truth Creation: Learning from Multiple Experts . 427
22.5 ImageCLEF Challenge 428
22.5.1 Materials and Methods 428
22.5.2 Results 429
22.6 Discussion and Conclusions 430
22.6.1 Clinical Impact 430
22.6.2 Future Extensions of CAD 432
References 433

23 Medical Image Classification at Tel Aviv and Bar Ilan Universities . . 435
Uri Avni, Jacob Goldberger, and Hayit Greenspan
23.1 Introduction 435
23.1.1 Visual Words in Medical Archives 436
23.2 The Proposed TAU–BIU Classification System Based on a Dictionary of Visual–Words 437
23.2.1 Patch Extraction 438
23.2.2 Feature Space Description 438
23.2.3 Quantization 439
23.2.4 From an Input Image to a Representative Histogram 440
23.2.5 Classification 441
23.3 Experiments and Results 442
23.3.1 Sensitivity Analysis 444
23.3.2 Optimizing the Classifier 446
23.3.3 Classification Results 449
23.4 Discussion 450
References 451

24 Idiap on Medical Image Classification 453
Tatiana Tommasi and Francesco Orabona
24.1 Introduction 453
24.2 Multiple Cues for Image Annotation 454
24.2.1 High–Level Integration 455
24.2.2 Mid–Level Integration 456
24.2.3 Low–Level Integration 456
24.3 Exploiting the Hierarchical Structure of Data: Confidence Based Opinion Fusion 457
24.4 Facing the Class Imbalance Problem: Virtual Examples 458
24.5 Experiments 458
24.5.1 Features 458
24.5.2 Classifier 461
24.5.3 Experimental Set–up and Results 462
24.6 Conclusions 463
References 464

Part IV External views

25 Press Association Images — Image Retrieval Challenges 469
Martin Stephens and Dhavalkumar Thakker
25.1 Press Association Images — A Brief History 469
25.1.1 The Press Association 469
25.1.2 Images at the Press Association 471
25.2 User Search Behaviour 472
25.2.1 Types of Users 472
25.2.2 Types of Search 473
25.2.3 Challenges 474
25.3 Semantic Web for Multimedia Applications 475
25.3.1 Introduction to the Semantic Web 475
25.3.2 Success Stories and Research Areas 475
25.3.3 The Semantic Web Project at Press Association Images . . 477
25.4 Utilizing Semantic Web Technologies for Improving User Experience in Image Browsing 478
25.4.1 PA Data set: Linking to the Linked Data Cloud 478
25.4.2 Information Extraction and Semantic Annotation 480
25.5 Conclusions and Future Work 481
References 481

26 Image Retrieval in a Commercial Setting 483
Vanessa Murdock, Roelof van Zwol, Lluis Garcia, and Ximena Olivares
26.1 Introduction 483
26.2 Evaluating Large Scale Image Search Systems 486
26.3 Query Logs and Click Data 487
26.4 Background Information on Image Search 490
26.5 Multilayer Perceptron 491

26.6 Click Data 493
26.7 Data Representation 495
26.7.1 Textual Features 495
26.7.2 Visual Features 497
26.8 Evaluation and Results 498
26.8.1 Analysis of Features 500
26.9 Discussion of Results 502
26.10 Looking Ahead 503
References 504

27 An Overview of Evaluation Campaigns in Multimedia Retrieval 507
Suzanne Little, Ainhoa Llorente, and Stefan Rüger
27.1 Introduction 507
27.2 ImageCLEF in Multimedia IR (MIR) 509
27.2.1 INEX XML Multimedia Track 510
27.2.2 MIREX 511
27.2.3 GeoCLEF 511
27.2.4 TRECVid 512
27.2.5 VideOlympics 514
27.2.6 PASCAL Visual Object Classes (VOC) Challenge 514
27.2.7 MediaEval and VideoCLEF 515
27.2.8 Past Benchmarking Evaluation Campaigns 516
27.2.9 Comparison with ImageCLEF 517
27.3 Utility of Evaluation Conferences 518
27.4 Impact and Evolution of Metrics 519
27.5 Conclusions 521
References 522

Glossary 527

Index 533

List of Contributors

Julien Ah-Pine
Xerox Research Centre Europe, Meylan, France

Adil Alpkocak
Dokuz Eylul University, Department of Computer Engineering, Izmir, Turkey

Uri Avni
Tel Aviv University, Tel Aviv, Israel

Steven Bedrick
Oregon Health and Science University, Portland, OR, USA

Tolga Berber
Dokuz Eylul University, Department of Computer Engineering, Izmir, Turkey

Jinbo Bi
Siemens at Malvern, PA, USA

Luca Bogoni
Siemens at Malvern, PA, USA

Barbara Caputo
Idiap Research Institute, Martigny, Switzerland

Jean–Pierre Chevallet
University of Grenoble, Laboratoire d'Informatique de Grenoble, Grenoble, France

Stephane Clinchant,
Xerox Research Centre Europe, Meylan, France

Paul D. Clough
University of Sheffield, Sheffield, UK

Gabriela Csurka,
Xerox Research Centre Europe, Meylan, France

Adrien Depeursinge
University and University Hospitals of Geneva (HUG), Geneva 14, Switzerland

Thomas Deselaers
Computer Vision Laboratory, ETH Zurich, Zurich, Switzerland

Manuel Carlos Díaz Galiano
SINAI Research group, University of Jaén, Jaén, Spain

Charles Florin
Siemens at Malvern, PA, USA

Lluis Garcia
Yahoo! Research, Barcelona, Spain

Miguel Ángel García Cumbreras
SINAI Research group, University of Jaén, Jaén, Spain

Eric Gaussier
University of Grenoble, Laboratoire d'Informatique, Grenoble, France

Theo Gevers
University of Amsterdam, Amsterdam, The Netherlands

Jacob Goldberger
Bar Ilan University, Ramat–Gan, Israel

Julio Gonzalo
E.T.S.I. Informática de la UNED, Madrid, Spain

Hayit Greenspan
Tel Aviv University, Tel Aviv, Israel

Michael Grubinger
Medellín, Colombia

Allan Hanbury
Information Retrieval Facility, Vienna, Austria

William Hersh
Oregon Health and Science University, Portland, OR, USA

Joo Hwee Lim
Institute for Infocom Research, Singapore

Anna K. Jerebko
Siemens at Malvern, PA, USA

Joemon M. Jose
University of Glasgow, Glasgow, UK

Jayashree Kalpathy–Cramer
Oregon Health and Science University, Portland, OR, USA

Jussi Karlgren
SICS, Kista, Sweden

Deniz Kilinc
Dokuz Eylul University, Department of Computer Engineering, Izmir, Turkey

Jana Kludas
CUI, University of Geneva, Switzerland

Arun Krishnan
Siemens at Malvern, PA, USA

Teerapong Leelanupab
University of Glasgow, Glasgow, UK

Monica Lestari Paramita
University of Sheffield, Sheffield, UK

Suzanne Little
KMi, The Open University, UK

Ainhoa Llorente
KMi, The Open University, UK

Loïc Maisonasse
TecKnowMetrix, Voiron, France

María Teresa Martín Valdivia
SINAI Research group, University of Jaén, Jaén, Spain

Arturo Montejo Ráez
SINAI Research group, University of Jaén, Jaén, Spain

Henning Müller
University of Applied Sciences Western Switzerland (HES–SO), Sierre, Switzerland

Vanessa Murdock
Yahoo! Research, Barcelona, Spain

Stefanie Nowak
Fraunhofer IDMT, Ilmenau, Germany

Ximena Olivares
Universitat Pompeu Fabra, Barcelona, Spain

Francesco Orabona
Dipartimento di Scienze dell'Informazione, Universita' degli Studi di Milano, Milano, Italy

Sangmin Park
Siemens at Malvern, PA, USA

Florent Perronnin
Xerox Research Centre Europe, Meylan, France

Andrzej Pronobis
Department of Computer Science, Royal Institute of Technology, Stockholm, Sweden

Saïd Radhouani
Koodya sàrl, Bou Salem, Tunisia

Vikas Raykar
Siemens at Malvern, PA, USA

Jean-Michel Renders
Xerox Research Centre Europe, Meylan, France

Stefan Rüger
KMi, The Open University, Milton Keynes, UK

Marcos Salganicoff
Siemens at Malvern, PA, USA

Koen E. A. van de Sande
University of Amsterdam, Amsterdam, The Netherlands

Mark Sanderson
University of Sheffield, Sheffield, UK

Alan F. Smeaton
CLARITY: Centre for Sensor Web Technologies, Dublin City University, Dublin, Ireland

Martin Stephens
Press Association Images, Nottingham, UK

Dhavalkumar Thakker
Press Association Images, Nottingham, UK

Tatiana Tommasi
Idiap Research Institute, Martigny, Switzerland

Theodora Tsikrika
Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

Guido Zuccon
University of Glasgow, Glasgow, UK

Roelof van Zwol
Yahoo! Research, Barcelona, Spain