SPRINGER BRIEFS IN COMPUTER SCIENCE

Jenny Benois-Pineau
Frédéric Precioso
Matthieu Cord

Visual Indexing and Retrieval

Springer

SpringerBriefs in Computer Science

For further volumes:
http://www.springer.com/series/10028

Jenny Benois-Pineau • Frédéric Precioso
Matthieu Cord

Visual Indexing and Retrieval

Springer

Jenny Benois-Pineau
Univ. Bordeaux
LABRI UMR 5800
Talence, France
jenny.benois@labri.fr

Frédéric Precioso
UNS
I3S UMR 7271/UNS_CNRS
Sophia Antipolis Cedex, France
precioso@polytech.unice.fr

Matthieu Cord
UPMC
LIP6 UMR 7606
matthieu.cord@lip6.fr

ISSN 2191-5768 ISSN 2191-5776 (electronic)
ISBN 978-1-4614-3587-7 ISBN 978-1-4614-3588-4 (eBook)
DOI 10.1007/978-1-4614-3588-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012935105

Printed on acid-free paper

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

Preface

Visual Indexing and Retrieval is a wide-scope research domain which unites researchers from image analysis and computer vision, to information management. For nearly twenty years, this research has been extremely active world-wide. With the advent of social networks conveying huge amount of visual information to the user, the ever increasing capacities of broadcast transmission of content via heterogeneous networks, popularization of hand-carried devices, the realm of possibilities has got wide open both for industry leaders and researchers. Indeed, the always increasing size of visual content databases has brought to the fore the need for innovative content understanding, retrieval and classification methods. Such methods are henceforth of paramount importance to let people the ability to exploit such huge databases.

The book is the result of joint efforts of the French research community joined through the GDR CNRS ISIS, the national academic network in the field. Thanks to this network, the research results gathered and explained in this book have received global recognition, and are gaining more and more success in technology transfer.

The authors hope that the most recent results and fruitful trends in Visual Indexing and Retrieval presented in this book will be helpful to young and experienced researchers willing to put the ideas forward and needing a solid understanding of the state-of-the art to do so, as well as to industry people willing to find an adequate algorithmic solution for their application area.

Acknowledgments: the authors acknowledge the governing board of GDR CNRS ISIS for supporting the project of this book.

Paris, France
January 2012

Jenny Benois-Pineau
Frédéric Precioso
Matthieu Cord

Contents

Chapter 1
Introduction

Jenny Benois-Pineau, Frédéric Precioso, Matthieu Cord

1.1 Context and motivations

The research in visual information indexing and retrieval has become one of the most popular directions in the broad area of information technologies. The reasons for that are the technological maturity of capture, storage and network infrastructures, that allow for common daily-life capturing of images and recording of video with professional equipment and personal mobile devices. According to Internet sources the British Broadcasting Corporation set up team dedicated to process user-generated content, as an experimental group in April 2005 with 3 employees in the staff. The team was then made durable and got expanded, unveiling the integration of "citizen journalist" in the (broadcast) news mainstream. The same concept has been put in place by CNN that launched CNN iReport in 2006. This project was meant to allow CNN to collect user-generated news. So did the american Fox News with their 'uReport' project and the french BFM-TV broadcast channels. YouTube, FaceBook, FileMobile, DailyMotion, host and supply facilities for accessing a tremendous amount of professional and user-generated content for educational and entertainment purposes. The areas of societal activity such as video surveillance and security also generate thousands of tera-bytes of video content with specific issues to be tackled. Finally, the digitization and storage of cultural heritage be it Byzantine frescoes, Medieval miniatures, old manuscripts or feature films, documentaries and broadcasting programs or web-sites, lead to the production of a mass of visual data which has to be accessed and searched both by professionals for re-mastering and production of new visual content and by common users for various humanities research.

Thus, visual information indexing and retrieval has attracted a lot of research efforts since the early nineties [24]. However nowadays, owing to the size of large scale databases, complexity of visual interpretation tasks, the famous "semantic gap" between the (semantic) concept(s) the user is looking for in images and the digital representation of the (semantic) visual content, the research directions are widely open. This research field has been dramatically enriched by the last achievements

J. Benois-Pineau et al., *Visual Indexing and Retrieval*, SpringerBriefs in Computer Science,
DOI 10.1007/978-1-4614-3588-4_1, © The Author(s) 2012

in machine learning methods, the standardization of content descriptors and large-scale evaluation campaigns providing researchers with access to amounts of visual data associated to benchmarks for testing and evaluation. In this book, we propose a detailed review of the cutting-edge approaches in Visual Information Indexing, along with the original contributions of the authors of chapters.

1.2 Outline of the book

The book structure has naturally emerged from the key components of the visual indexing and retrieval problems. The next chapters, – Spatial and multi-resolution descriptions in visual indexing and – Machine learning approaches for visual information retrieval, are presenting the ingredients of the whole process. The next ones are dedicated to advanced considerations for the image representations, the scalability issues, and the evaluation.

Visual indexing and retrieval tasks always start with feature extraction from the raw data. Many approaches in several contexts have been developed since more than 30 years for now. Many primitives as Points of interest, regions, lines and so on, have been studied. Additionally, efficient descriptors are required. The Chapter 2 provides a deep overview of the basic feature extraction and description methods in the literature.

Representing these features at a higher semantic level description is the second stage of the process. The main steps for deriving image representation from visual local descriptors are described in Chapter 3. Well-known Bag-of-Visual-Words including recent extensions as sparse coding and spatial pooling methods are detailed. Higher semantic level data representation and similarity design are not independent but strongly related processes. Similarity measures between histograms, but also more complex functions such as kernels are presented. In order to address classification, retrieval or detection tasks, these similarities must be integrated into machine learning frameworks. We chose to focus on two major contributions from the Machine Learning community, namely Support Vector Machines and Boosting, very successful in multimedia retrieval applications.

Spatial structure is a key point in the building of the image representation. This information is usually ignored in basic representations. Advanced approaches try to overcome this drawback by adding spatial arrangements in the process. Furthermore, the natural willing to use information according to its visual importance and compressed-stream analysis yields multi-resolution approaches. In Chapter 4, integration of spatial context into visual description is deeply studied. Two trends on how to account for such information are considered. The first is the design of structural descriptors and integration of the spatial context in the signature matching. The second relates to the multi-resolution visual indexing.

Since image and video databases are becoming ever larger, scalability is definitely a critical requirement for visual information retrieval. The chapter 5 is dedicated to these scalability issues. The nature of the problems both for content-based

retrieval and for mining are described. Main ideas and recent advances are presented, like the use of approximation or of shared-neighbour similarity. We also highlight some prospective directions, like embeddings, filtering based on simplified descriptions, optimization of content representations and distributed processing.

Finally, it is crucial to evaluate the visual indexing and retrieval methods and systems against common benchmark corpora. Having a deep understanding of the evaluation process is the best way to identify the strengths and weaknesses of the different approaches, and to trace promising ways to investigate. The chapter 6 gives an overview of the major evaluation campaigns and benchmarks for visual indexing and retrieval tasks. Data collection, relevance judgements, performance measures and experimentation protocols are discussed. The State-of-the-Art performance in recent campaigns and the lessons learnt from these campaigns are also presented.