Lecture Notes in Computer Science 1464

Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Horace H.S. Ip Arnold W.M. Smeulders (Eds.)

Multimedia Information Analysis and Retrieval

IAPR International Workshop, MINAR'98
Hong Kong, China, August 13-14, 1998
Proceedings

Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Horace H.S. Ip
City University of Hong Kong, Department of Computer Science
Tat Chee Avenue, Kowloon, Hong Kong, China
E-mail: cship@cityu.edu.hk

Arnold W.M. Smeulders
University of Amsterdam, Intelligent Sensory Information Systems
Kruislaan 403, 1098 SJ Amsterdam, The Netherlands
E-mail: smeulders@wins.uva.nl

Cataloging-in-Publication data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Multimedia information analysis and retrieval : IAPR international workshop ; proceedings / MINAR '98, Hong Kong, China, August 13 - 14, 1998. Horace H. S. Ip ; Arnold W. M. Smeulders (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Singapore ; Tokyo : Springer, 1998
(Lecture notes in computer science ; Vol. 1464)
ISBN 3-540-64826-7

CR Subject Classification (1991): I.4, I.5, H.5.1, H.4.3, H.5.4, H.3

ISSN 0302-9743
ISBN 3-540-64826-7 Springer-Verlag Berlin Heidelberg New York

Typesetting: Camera-ready by author
SPIN 10638431 06/3142 – 5 4 3 2 1 0 Printed on acid-free paper

Preface

The advent of Internet coinciding with the general use of digital sensors have given the field of computer vision and pattern recognition a boost in its application. Whilst the digital camera captures images for a variety of professional and consumer markets, the Internet spreads the use of pictures in digital form over the world. In short, the Internet has turned visual rather than textual. With the new tools come new usage, new audiences and, hence, new challenges for computer vision and recognition. Such challenges come in three different forms: the much enlarged dataset, new application domains, and the needs to cater to naive users.

Pictorial databases of 500 000 to over 1 000 000 images are quite common in Internet search. This imposes great demands on computational efficiency and issues relating to indexing compressed pictures and videos become relevant. Moreover, the requirement of robustness against varying circumstances in recording and interpretation in image browsing via the Internet is also larger than previously seen.

Before the Internet age, typical application domains of picture analysis were usually complicated tasks in narrow domains (e.g. medical image analysis, bank cheque verification, pharmaceutical product classification) or quick tasks in simplified environments (e.g. industrial inspection, printed character recognition). In the Internet age, the domain shifts to more general pictures from consumer video and large archives, where we have very little control over the imaging conditions, and where prior domain knowledge is not easily applied to simplify the search and retrieval process. This poses interesting new problems each time when conditions of illumination and query purpose alter. A general framework which addresses all these problems is not yet in sight.

Apart from new domains, the Internet and digital sensor revolution also generates new questions for vision research. Searching a database or the Internet requires similarity measures over very large numbers of pictures. Also, interactive browsing of an unknown query where the emphasis is on perceptual similarity is a relative unknown topic in computer vision.

Sensing the importance and the needs of new techniques for browsing large image and multimedia database, the International Association of Pattern Recognition (IAPR) established a new technical committee on Multimedia Systems (TC12) in 1995. This technical committee directs its attention to the vast number of scientific questions related to multimedia content analysis and retrieval.

In addition, TC12 also initiated the IAPR International Workshop on Multimedia Information Analysis and Retrieval (MINAR'98) which aims to bring together researchers who are working in the field of interests of the TC.

This volume is devoted to major research issues in content-based image and video search, and contains papers presented at MINAR'98 held in Hong Kong. Among these papers are topics on the exploitation of invariant properties, colour or geometric, of images for robust image and video retrieval, fusion of pictorial with other media such as text, image indexing and retrieval in compressed domain, pictorial query languages,

video segmentation by content, as well as efficient storage organisation for multimedia data.

We would like to thank the following who served on the programme committee of MINAR'98: S.-K. Chang, A. del-Bimbo, W. Grosky, R. Kasturi, T. Kato, C. Leung, S.D. Ma, D. Petkovic, H. Samet, S. Smoliar, R. Srihara, H. Tagare and J.K. Wu. We also thank Ramesh Jain who gave the keynote talk on presence technology.

We are also grateful to Charmaine Yeung for taking care of the logistics of paper collection and the workshop organisation; T.C. Pong, Irwin King and Ken Law for publication, publicity and the local arrangements.

MINAR'98 was organised by IAPR TC12 and co-sponsored by the IEEE (Hong Kong Section) Computer Chapter.

Horace H.S. Ip
Department of Computer Science
City University of Hong Kong
Hong Kong

Arnold W.M. Smeulders
Intelligent Sensory Information Systems
University of Amsterdam
The Netherlands

June 1998

Contents

Video Segmentation and Spatial Query

Indexing and Storage

Posters