
**PERSPECTIVES ON
CONTENT-BASED
MULTIMEDIA SYSTEMS**

THE KLUWER INTERNATIONAL SERIES ON INFORMATION RETRIEVAL

Series Editor

W. Bruce Croft

University of Massachusetts, Amherst

Also in the Series:

MULTIMEDIA INFORMATION RETRIEVAL: Content-Based Information Retrieval from Large Text and Audio Databases, by *Peter Schüuble*; ISBN:0-7923-9899-8

INFORMATION RETRIEVAL SYSTEMS: Theory and Implementation, by *Gerald Kowalski*; ISBN: 0-7923-9926-9

CROSS-LANGUAGE INFORMATION RETRIEVAL, edited by *Gregory Grefenstette*; ISBN:0-7923-8122-X

TEXT RETRIEVAL AND FILTERING: Analytic Models of Performance, by *Robert M. Losee*; ISBN: 0-7923-8177-7

INFORMATION RETRIEVAL: UNCERTAINTY AND LOGICS: Advanced Models for the Representation and Retrieval of Information, by *Fabio Crestani, Mounia Lalmas, and Cornelis Joost van Rijsbergen*; ISBN: 0-7923-8302-8

DOCUMENT COMPUTING: Technologies for Managing Electronic Document Collections, by *Ross Wilkinson, Timothy Arnold-Moore, Michael Fuller, Ron Sacks-Davis, James Thom, and Justin Zobel*; ISBN: 0-7923-8357-5

AUTOMATIC INDEXING AND ABSTRACTING OF DOCUMENT TEXTS, by *Marie-Francine Moens*; ISBN 0-7923-7793-1

ADVANCES IN INFORMATIONAL RETRIEVAL: Recent Research from the Center for Intelligent Information Retrieval, by *W. Bruce Croft*; ISBN 0-7923-7812-1

INFORMATION RETRIEVAL SYSTEMS: Theory and Implementation, Second Edition, by *Gerald J. Kowalski and Mark T. Maybury*; ISBN: 0-7923-7924-1

PERSPECTIVES ON CONTENT-BASED MULTIMEDIA SYSTEMS

by

Jian Kang Wu

Kent Ridge Digital Labs, Singapore

Mohan S. Kankanhalli

National University of Singapore

Joo-Hwee Lim

Kent Ridge Digital Labs, Singapore

Dezhong Hong

Kent Ridge Digital Labs, Singapore

KLUWER ACADEMIC PUBLISHERS

New York / Boston / Dordrecht / London / Moscow

eBook ISBN: 0-306-47033-0
Print ISBN: 0-792-37944-6

©2002 Kluwer Academic Publishers
New York, Boston, Dordrecht, London, Moscow

Print ©2000 Kluwer Academic Publishers
Massachusetts

All rights reserved

No part of this eBook may be reproduced or transmitted in any form or by any means, electronic, mechanical, recording, or otherwise, without written consent from the Publisher

Created in the United States of America

Visit Kluwer Online at: <http://kluweronline.com>
and Kluwer's eBookstore at: <http://ebooks.kluweronline.com>

Contents

Preface	xiii
1 Introduction	1
1.1 Technical Issues	2
1.1.1 Inter-operability	2
1.1.2 Automatic or semi-automatic indexing	3
1.1.3 Fool-proof retrieval methods	3
1.2 Content-Based Retrieval	4
1.2.1 Basic idea	4
1.2.2 Existing work	4
1.3 Challenges and solutions	10
1.4 Layout of the Book	12
References	12
2 Formalism of Content-based Multimedia Systems	15
2.1 The System Must be User-centered	15
2.1.1 Man-machine interaction must be extremely friendly	15
2.1.2 Reduce the work and information overload of users	15
2.2 Content-based Multimedia Information System	16
2.2.1 Definitions	17
2.2.2 Challenges in content-based retrieval	21
2.2.3 Representation of multimedia objects	23
2.2.4 Multimedia data analysis	24
2.3 Object Recall - A New Formalism for Content-Based Retrieval	26
2.3.1 Need for formalism of content-based retrieval	26
2.3.2 Contrast to pattern classification	28
2.3.3 Formalism for content-based navigation	30
2.3.4 Formalism for content-based retrieval by adaptive fuzzy concept	31

2.4	A Content-Based Similarity Retrieval Formalism	33
2.4.1	Feature measures and similarity functions	33
2.4.2	Training data set and loss function	34
2.5	Learning of Similarity Function	35
2.5.1	Problem statement	35
2.5.2	Related works	36
2.5.3	Learning of simple similarity functions	37
2.5.4	Learning of multi-level similarity functions	38
2.6	Experimental Results	41
2.6.1	The data set	41
2.6.2	Feature extraction	41
2.6.3	Within-sequence learning	43
2.6.4	Across-sequence learning	44
2.7	Conclusions	47
	References	47
3	Color Feature Extraction	49
3.1	Color Spaces Selection	49
3.1.1	<i>RGB</i> color space	49
3.1.2	Munsell color system	50
3.1.3	CIE color systems	50
3.1.4	HSV color space	51
3.2	Color Measures	52
3.2.1	A brief review of color features	53
3.2.2	The distance method	55
3.3	Reference Color Table Method	56
3.3.1	The color clustering-based method	58
3.3.2	Experimental results	64
3.3.3	Remarks	65
	References	66
4	Texture Feature Extraction	69
4.1	Discrete Cosine Transformed Texture	70
4.1.1	Discrete cosine transform	70
4.1.2	Texture feature based on discrete cosine transform	71
4.1.3	Feature vector formation	72
4.1.4	Test results	72
4.2	Wavelet Transformed Texture Feature	74
4.2.1	Wavelet Transform	74
4.2.2	Feature vector formation	75
4.2.3	Performance evaluation and test results	79

4.3	Texture Features Based on Second Moment Matrix	82
4.3.1	Performance evaluation and test results	83
4.4	Comparative Study	84
4.4.1	Comparison based on classification	84
4.4.2	Multiple features - combination of color and texture	84
4.4.3	Retrieval test	87
	References	90
5	Video Processing	93
5.1	Review of Video Processing Techniques	94
5.1.1	Video features	94
5.1.2	Video applications	95
5.1.3	Research areas	96
5.1.4	State of the art review	97
5.2	Content-Based Representative Frame Extraction and Video Summary	100
5.2.1	Definition	100
5.2.2	Related work	101
5.2.3	Extraction of representative frames	102
5.2.4	Application of representative frame extraction technique	114
	References	117
6	Object Segmentation	121
6.1	Edge-preserved smoothing of features	123
6.1.1	Principle of edge-preserved smoothing	123
6.1.2	EPSM for 2D signal	127
6.1.3	Application in color feature	128
6.1.4	Application in texture feature	129
6.2	Segmentation algorithm: clustering and region merging	134
6.2.1	Clustering in the feature space	135
6.2.2	Cluster validation for unsupervised segmentation	138
6.2.3	Markov random field model and Gibbs distribution	145
6.2.4	Region analysis and region merging	152
6.3	Experiment results	156
6.3.1	Experiment setup	156
6.3.2	Experiment results	157
6.3.3	Summary	169
	References	170
7	Human Face Detection	173

7.1	Color Segmentation of Faces	174
7.1.1	Chromaticity diagrams	176
7.1.2	Effects of the projection of 3-D color spaces on chromaticity diagrams	180
7.1.3	Method for face detection using chromaticity diagrams	184
7.1.4	Experiments	185
7.2	Shape Information As a Cue	188
7.2.1	Geometric characteristic of face	188
7.2.2	Shape descriptors	190
7.2.3	Experimental results	191
7.3	Face Feature Detection Using DOG Operators	191
7.3.1	The DOG (Difference of Gaussians) operator	191
7.3.2	Face feature detection by DOG operator	193
7.3.3	Experimental results	194
7.3.4	Discussion	196
7.4	Template-based Human Face Detection	198
7.4.1	Overview	198
7.4.2	Normalized “face space”	199
7.4.3	Dimensionality reduction	199
7.4.4	Clustering and face template generation	200
7.4.5	Template matching	203
	References	206
8	Visual Keywords	209
8.1	Introduction	209
8.2	Related Works	212
8.3	Methodology	214
8.3.1	Typification	215
8.3.2	Description scheme	219
8.3.3	Selection	222
8.3.4	Coding scheme	222
8.4	Image Retrieval	223
8.4.1	Unsupervised learning	224
8.4.2	Learning by instruction	228
8.5	Image Categorization	234
8.6	Conclusions and Future Directions	236
	References	237
9	Fuzzy Retrieval	239
9.1	Problem Definition	240

9.1.1	Content-based fuzzy retrieval	240
9.1.2	Fuzzy sets over multi-dimensional universes	242
9.1.3	Fuzzy queries cannot be processed in the feature space	243
9.1.4	Fuzzy querying of multimedia information calls for new technology for its processing	244
9.2	Fuzzy Database Model	246
9.2.1	Extended tuple relation calculus for image retrieval	249
9.3	Fuzzy Query Processing	251
9.3.1	Feature space and fuzzy space	251
9.3.2	Fuzzy query interface	252
9.3.3	Context model	253
9.3.4	Extraction of the confidence value	254
9.3.5	Incomplete query condition	255
9.3.6	Similarity measure	256
9.3.7	Relevance feedback for query refinement	258
9.4	Learning Fuzzy Membership Functions	258
9.4.1	Need for learning fuzzy membership functions	259
9.4.2	A Fuzzy neuro membership function	262
9.4.3	Training the neural network	263
9.5	Experimental Results of Fuzzy Retrieval	264
9.5.1	Face database indexing and retrieval	264
9.6	Conclusion and Remarks	270
	References	270
10	Face Retrieval	273
10.1	CAFIIR system	274
10.1.1	Data model	277
10.1.2	Indexing of facial images	279
10.1.3	Feature extraction	279
10.1.4	Content-based indexing	280
10.2	Content Based Indexing of Multimedia Object	282
10.2.1	Definition	282
10.2.2	Horizontal links	283
10.2.3	A practical example	285
10.2.4	Iconic images construction	286
10.2.5	Content-based retrieval using ContIndex	286
10.3	ContIndex Creation by Self-organization Neural Networks	287
10.3.1	LEP neural network architecture	289
10.3.2	Fusion of multi-modal feature measures	290
10.3.3	Spatial self-organization	291

10.3.4 Bi-directional learning on experiences	293
10.4 Experimental Results	294
10.5 Conclusion and Remarks	300
10.6 Visual Retrieval of Facial Images	300
10.6.1 Facial composition	300
10.6.2 Browsing and similarity retrieval	301
10.7 Descriptive Queries	303
10.7.1 Fuzzy retrieval of facial images	303
10.7.2 Free text retrieval	306
10.8 Further Improvement of Queries	307
10.8.1 Feedback for query refinement	307
10.8.2 Combined query	308
10.9 Implementation and Concluding Remarks	308
References	309
11 System for Trademark Archival and Retrieval	311
11.1 Representation of Trademarks	312
11.2 Segmentation of Trademarks	314
11.2.1 Color segmentation	315
11.3 Capturing Visual Features of Trademark Images	318
11.3.1 Structural description	319
11.3.2 Feature measures	319
11.3.3 Match shape interpretation using fuzzy thesaurus	322
11.4 Composite Similarity Measures	323
11.5 Evaluation and Learning of Similarity Measures	326
11.5.1 Selection of training and test data sets	326
11.5.2 Learning of similarity functions	326
11.5.3 Evaluating the shape retrieval	328
11.6 Experimental Results	330
11.7 Conclusions	332
References	332
12 Digital Home Photo Album	335
12.1 Digital photo is becoming popular	335
12.1.1 What do the home users want?	336
12.1.2 User study 1	337
12.1.3 User study 2	339
12.1.4 The gap between ideal and realistic	339
12.2 Object-based Indexing and Retrieval	341
12.2.1 Object categories	343
12.2.2 Object models	344

12.2.3	System training	345
12.2.4	Image categorization	348
12.2.5	Image retrieval	352
12.2.6	Experimental results	354
References	358
13	Evaluation of Content-based Retrieval	361
13.1	Definition of the Problem	361
13.1.1	Retrieval as a function of data and query	362
13.1.2	Definition of benchmarking of multimedia databases	363
13.2	Benchmarking for Content-based Systems	364
13.2.1	Benchmarking for database retrieval methods	364
13.2.2	Benchmarking of information retrieval methods	365
13.2.3	Benchmarking links	367
13.2.4	Benchmarks for approximate retrieval methods	368
13.2.5	Complete benchmark for multimedia databases	380
13.3	Testing Multimedia Databases	380
13.3.1	Scalability with respect to time	381
13.3.2	Scalability with respect to quality	381
13.3.3	Examples of testing and evaluation	383
13.4	Conclusion	386
References	386
Index		389

Preface

Due to the rapid advances in computing and communication technologies, human beings are constantly being inundated by information in form of text, image, audio, video and spatial data. There is an overwhelming need for an integrated multimedia system to reduce the work and information overload for people. The technologies for handling multimedia data are most important and most challenging: our society is increasingly generating vast amount of multimedia data by means of cameras, satellites, etc.; the complexity and variety of multimedia data are beyond one's imagination.

It is quite natural for us human beings to recall multimedia information by content. We recognize our friends by their appearances. The story of one video may remind us of the ones similar to that. Therefore, content-based access to multimedia is of primary importance. This book deals with various aspects of content-based multimedia systems.

Formalization of the whole paradigm of content-based retrieval to bring it to a sufficient level of consistency and integrity is essential to the success of the field. After introduction, we will discuss the formalism of content-based multimedia systems, which is proposed and developed by the authors of the book. With the formalism, a reconfigurable retrieval engine can be developed.

For the content-based retrieval of a large multimedia database, a unified feature space (referred to as vector space in information retrieval community) is desirable. In such a unified feature space, data items can be represented and measures of similarity can be defined.

Unlike conventional databases, where primary data types are well handled by standard query languages such as SQL, the retrieval of multimedia databases is approximate, and is not likely able to be handled by a single language like SQL. In the fourth chapter, we will present methods for content-based indexing for visual browsing, fuzzy retrieval method, customizable similarity retrieval, and case-based reasoning.

As show cases, three application examples will be presented. A face database is a typical image database where each data item share the same structure. Many concepts can be demonstrated using this system. Because of the diversity of the trademarks, the retrieval of conflicting trademarks

relies on several properties of those trademarks. These include the word, phonetics, shape and interpretation of the trademark image. The system retrieval engine should be able to find feature spaces to accommodate those properties in order to define proper similarity functions. The medical database example is used to illustrate case-based reasoning. Digital photo album is a very practical application as digital cameras become popular.

This book covers the research results supported by Kent Ridge Digital Labs (KRDL), Singapore and funded by Real World Computing Partnership (RWCP), Japan. The authors sincerely appreciate the support and help given by the management and colleagues from these two organizations.

The following people made contributions to the book: Chian Prong Lam (Chapter 10 and 11), Fuchun Shang (Chapter 4, 7 and 12), Xinding Sun (Chapter 5), Bin Wang (Chapter 6), and Harro Stokman (Chapter 7). We would also like to thank Ms. Anuradha Srinivasan for the proof-reading help.

Jian-Kang WU