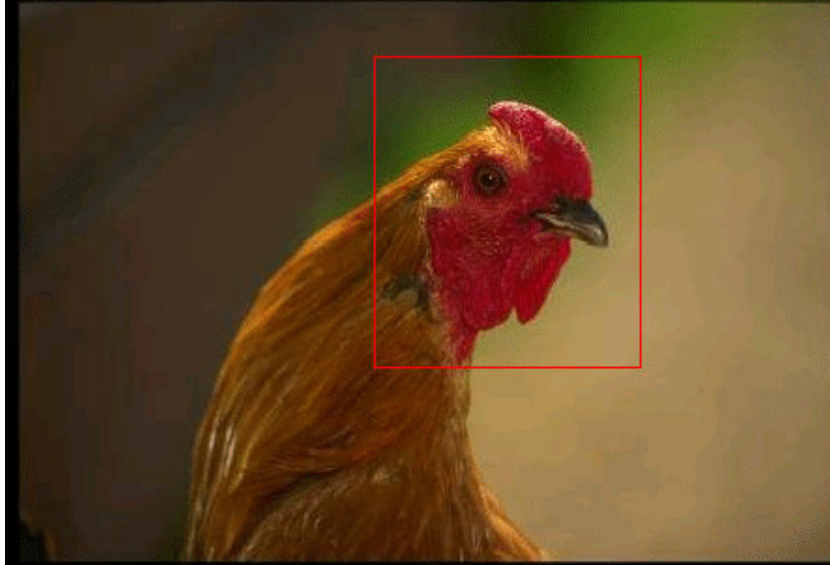


---

# STATE-OF-THE-ART IN CONTENT-BASED IMAGE AND VIDEO RETRIEVAL





# STATE-OF-THE-ART IN CONTENT-BASED IMAGE AND VIDEO RETRIEVAL

Edited by  
**REMCO C. VELTKAMP**  
Utrecht University  
**HANS BURKHARDT**  
University of Freiburg  
**HANS-PETER KRIEGEL**  
University of Munich

**Kluwer Academic Publishers**  
Boston/Dordrecht/London



# Contents

Preface	vii
1 Image Content Analysis and Description <i>Xenophon Zabulis, Stelios C. Orphanoudakis</i>	1
2 Local Features for Image Retrieval <i>Luc Van Gool, Timme Tuytelaars, Andreas Turina</i>	21
3 Fast Invariant Feature Extraction for Image Retrieval <i>Sven Siggelkow, Hans Burkhardt</i>	43
4 Shape Description and Search for Similar Objects in Image Databases <i>Longin Jan Latecki, Rolf Lakaemper</i>	69
5 Features in Content-based Image Retrieval Systems: a Survey <i>Remco C. Veltkamp, Mirela Tanase, Danielle Sent</i>	97
6 Probablistic Image Models for Object Recognition and Pose Estimation <i>Joachim Hornegger, Heinrich Niemann</i>	125
7 Distribution-based Image Similarity <i>Jan Puzicha</i>	143
8 Distribution Free Statistics for Segmentation <i>Greet Frederix, Eric J. Pauwels</i>	165
9 Information Retrieval Methods for Multimedia Objects <i>Norbert Fuhr</i>	191
10 New descriptors for image and video indexing <i>Patrick Gros, Ronan Fablet, Patrick Bouthemy</i>	213

11		
	Facial and Motion Analysis for Image and Video Retrieval	235
	<i>Massimo Tistarelli, Enrico Grosso</i>	
12		
	Asymmetric Similarity Measures for Video Summarisation	255
	<i>Sorin M. Iacob, Reginald L. Lagendijk, M. E. Iacob</i>	
13		
	Video Retrieval using Semantic Data	279
	<i>Alberto Del Bimbo</i>	
14		
	Adaptable Similarity Search in Large Image Databases	297
	<i>Thomas Seidl, Hans-Peter Kriegel</i>	
15		
	Parallel NN-search for large multimedia repositories	319
	<i>Roger Weber, Klemens Böhm, Hans-Jörg Schek</i>	

## Preface

*Content-based image and video retrieval* is concerned with retrieving images and video sequences on the basis of automatically derived features such as color, texture, and shape information that is present in the images. The need for efficient storage and retrieval of images is not new, but the increase in the number and size of digital image collections made the problems of image retrieval widely recognized. The shortcomings with traditional methods of indexing, often based on keywords, have led to the interest in retrieval on the basis of features that are automatically derived from the image content. The 1992 USA's NSF workshop on Visual Information Management Systems [1] was one of the events that gave this research area a push. The search for solutions is now an active area of research.

The emergence of content-based image and video retrieval has brought together and broadened the extend of a number of disciplines, such as image understanding, database indexing, object recognition, visual data modeling, feature extraction, visual querying, perception and cognition, and information visualization.

These developments motivated us to organize the Dagstuhl Seminar on Content-Based Image and Video Retrieval, 5-10 December 1999, Schloss Dagstuhl, Wadern, Germany [2]. The purpose of this seminar was to bring together people from the various fields in order to promote information exchange and interaction among researchers who are interested in various aspects of accessing the content of image and video data.

The past decade has witnessed the development of the first few commercial systems. These are still very limited in their functionality. The user requirements on image features can vary considerably, but they can be classified into three levels of abstraction: primitive features such as color or textures, logical features such as the identity of objects, and abstract features such as meaning or significance of the image [3]. While most current systems operate on the lowest of these levels, the user demands higher levels.

Therefore, the emphasis of the following seminar [4] will lie on identifying the principal obstacles that hamper progress in content-based retrieval. Fundamental questions such as whether image ‘understanding’ is necessary for effective image ‘retrieval’ and whether ‘low’ level features are sufficient for ‘high’ level querying.

## **Structure**

This book is a result of the 1999 Dagstuhl Seminar on Content-based Image and Video Retrieval [2]. It contains a collection of works that represent the latest thinking in content-based image and video retrieval and cover a wide spectrum of areas. The chapters are a selection from the presentations at the seminar, and written for this book.

The chapters are grouped as follows:

- The first five chapters are dealing with features in image retrieval. To start off, chapter 1 discusses perceptual relevant features. Chapter 2 derives local image features from affine invariant regions. Chapter 3 is about local features integrated over rigid motions, and the evaluation with a Monte Carlo method. Chapter 4 treats 2D object retrieval by establishing correspondence between parts. Chapter 5 concludes this quintet with an overview of features used in 44 contemporary image retrieval systems.
- The next four chapters exploit, in different ways, probabilistic methods. Chapter 6 takes a probabilistic approach to classification and pose parameter estimation. Chapter 7, after an empirical evaluation, derives a framework that allows heterogeneous image models based on feature distributions. Chapter 8 is based on non-parametric density estimation of image feature clustering. Chapter 9 presents a logic-based retrieval system, where semantic information is transformed into a probabilistic object-oriented logic.
- The following four chapters are (in part) concerned with video retrieval. Chapter 10 introduces a multi-scale approach to image features, and a hierarchical motion classification for video. Chapter 11 describes, besides facial features, motion features derived from optic flow. Chapter 12 is about the hierarchical clustering of video key frames. In chapter 13, semantic information is derived from lower level features and signs.
- The last two chapters focus on indexing. Chapter 14 is treating similarity search that is adaptable to application specific requirements and user preferences. Finally, chapter 15 is about parallel nearest neighbor searching on a network of workstations.

Because this grouping is not strict and unambiguous, we have not made an explicit division of the book into parts.

## Acknowledgment

We wish to thank all authors for their contributions, and their collaborative effort in producing an up-to-date state-of-the-art book in such a dynamic field. We thank also all other participants in the seminar for their presentations and lively discussions.

Last, but not least indeed, we thank the Dagstuhl organization for offering such a wonderful facility, and the Dagstuhl office for their perfect support.

## References

- [1] Ramesh Jain (ed.). NSF Workshop on Visual Information Management Systems, Redwood, CA, 24-25 Feb. 1992. In *Storage and Retrieval for Image and Video Databases, Proceedings SPIE 1908*, pages 198–218, 1993. Also appeared in SIGMOD Record 22(3), 57-75, 1993.
- [2] Content-Based Image and Video Retrieval, Dagstuhl Seminar 99491, Schloss Dagstuhl, Wadern, Germany, 5-10 December, 1999. <http://www.dagstuhl.de/DATA/Reports/99491>.
- [3] John P. Eakins and Margaret E. Graham. Content-Based Image Retrieval, A Report to the JISC Technology Application Programme. Technical report, Institute for Image Data Research, University of Northumbria at Newcastle, UK, January 1999. <http://www.unn.ac.uk/iidr/report.html>.
- [4] Content-Based Image and Video Retrieval, Dagstuhl Seminar 02021, Schloss Dagstuhl, Wadern, Germany, 6-11 January, 2002. <http://www.dagstuhl.de/DATA/Reports/02021>.