

Guest Editorial

Introduction to the Special Issue on Image and Video Processing for Digital Libraries

Welcome to the Special Issue on Digital Libraries. It is a nice coincidence that the first issue of the TRANSACTIONS ON IMAGE PROCESSING for the 21st Century be on a topic that has seen a remarkable growth in the last few years, and one that is very likely to have a significant impact on various aspects of technology and human life in the new millennium!

Digital media and, in particular, images and video data, present many challenging problems in storage, access, and distribution. Large collections of multimedia documents can be found in diverse application domains such as the entertainment industry, education, medical imaging, and geographic information systems. The next generation multimedia standards (MPEG4 and MPEG7) have placed emphasis not only on coding efficiency but also on representations that facilitate content-based access and manipulation of coded data. While the individual needs for each of these domains could be quite diverse, they signify the problems and issues facing collections of large heterogeneous documents that include text, images, video, and audio. This special issue is devoted to the latest research efforts and results in addressing image and video processing issues, such as automatic visual content analysis, content similarity measures, and intellectual property identification and management, related to retrieval for digital libraries.

A total of 43 papers were submitted in response to the Call for Papers for this Special Issue. The guest editors (or their students/close collaborators) agreed not to submit any papers to this issue, to keep the reviewing process objective. Each paper was reviewed by at least three experts. At the end, we selected nine regular papers and six correspondence articles for inclusion in this special issue.

Among the nine regular papers, five are on similarity-based retrieval in image databases, one on analyzing DCT domain digital watermarking, and three on video indexing and retrieval.

Mojsilovic *et al.* develop methods for feature extraction and matching based on analyzing perceptual similarity. They used the results from subjective experiments and multidimensional scaling to compute what they refer to as the *perceptual dimensions*. They provide experimental results on a wide range of color image databases.

Cox *et al.* present a Bayesian framework for content based image retrieval. Their system, the PicHunter, maintains a probability distribution characterizing the probability of each of the images in the database as a likely target. This probability distribution is updated based upon user interactions with the system. They provide both quantitative and psychophysical performance

evaluation results illustrating that the Bayesian formulation improves the retrieval effectiveness.

Adjeroh and Lee present an occupancy model for selecting a subset of image features to be used for similarity based retrieval. The model guarantees that the decisions computed based on the subset of features is close (within a certain bound) to the performance that would have been achieved considering the entire feature set. Results on synthetic and real data are provided.

Gevers and Smeulders present PicToSeek, a retrieval system that combines color and shape invariants. A new set of color models invariant to viewpoint, object geometry, and illumination conditions is proposed. They conclude that combining the color and shape invariant features result in significant improvement in the retrieval effectiveness.

Jia and Kitchen present a method for comparing segments of the object contour, and encode object class information in computing the similarity measure. Segments of object contours are used to represent the object and their method is invariant to scale, rotation and translation. They provide experimental results on different shape image databases.

Hernandez *et al.* provide a detailed analysis of the DCT domain watermarking techniques based on spread spectrum embedding. This analysis helps in predicting the expected performance (such as probability of error in decoding or the probability of false alarm) for a given image. This helps in designing proper watermarks for a given data and in setting the criteria for watermark detection at the receiver.

Vasconcelos and Lippman use statistical models for characterizing structure and semantics in video sequences. They use the Bayesian formulation for the familiar shot segmentation problem and demonstrate that their model leads to improved segmentation. Extending this to the video feature space, they show that the Bayesian model captures well the semantic properties of the content and is useful for content based access to video libraries.

Stringa and Regazzoni describe a video surveillance system with automatic shot detection and indexing capability. Their system is aimed at detecting the presence of abandoned objects in a secure environment. The system can detect in real time any suspicious objects, and one can search a video library for such events as a person leaving an object behind.

Dagtas *et al.* discuss video indexing techniques capable of representing semantics of the data. They propose models for motion-based characterization of video events that are effective in content-based searches.

Among the correspondence items, Liu and Wechsler present two new methods for indexing and search of face image databases. Nguyen, Worring, and Dev describe a motion based video segmentation algorithm. Milios and Petrakis propose a

shape matching algorithm based on dynamic programming. Li, Doerman, and Kia describe methods for text detection and tracking in video. Fuh *et al.* propose an image retrieval method based on color segmentation and encoding region relationships. The work by Zhu *et al.* describes a content-based image retrieval system in the context of a geographic image library.

The research focus in recent years has shifted from low level feature descriptors to building more complex models for search and retrieval. Most of the recent developments are aimed at an object-based representation, and image/video segmentation still remains as one of the key issues. Some of these works attempt to solve the longstanding object segmentation and recognition problems in computer vision in connection with digital libraries for fully automated indexing and retrieval. However, in certain digital libraries applications, one can assume that the segmentation information is either available or can be obtained interactively. Developing schemas that can characterize spatial and temporal aspects of objects in images and video is an active research area. We expect to see more work on issues such as indexing in high-dimensional feature spaces for efficient access to multimedia objects, and the use of adaptive mechanisms to tailor the digital libraries for individual users. This requires integrating information from different modalities and different media, and deriving high level semantics—a significant leap from the current state of the art. We hope the papers presented in this Special Issue provides motivation for further research on this important subject in the new millennium.

M. Abdel-Motteleb
W. Al-Khatib
A. Alatan
A. N. Akansu
Y. Altunbasak
C. A. Bouman
F. Brill
S.-F. Chang
R. Chellappa
C. Chen
H. Chen
R. Chin
I. Cox
E. Delp
N. Dimitrova
M. S. Drew
C. Dyer
W. Effelsberg
E. Eren
M. Ferman
S. Fischer
Y. Fu
A. Ghafoor
W. Grosky
C. Gu
J. Huang
L. Huang
Q. Huang
M. Irani
H. V. Jagadish

A. K. Jain
J. Jin
M. Kankanhalli
J. Konrad
F. Kossentini
D. Krigman
R. Krishnapuram
C.-C. Kuo
Z. Lei
C.-S. Li
R. Linhart
S. Liou
A. Lippman
L. Lucchese
Q. Liu
W. Y. Ma
F. Marquis
C. Matthews
R. Mehrotra
E. Milios
M. Misra
D. Mukherjee
F. Nack
M. Naphade
W. Niblack
M. Orchard
S. Panchanathan
D. Petkovic
C. Podilchuk
D. Ponceleon

ACKNOWLEDGMENT

Thanks go to thank the Editor-in-Chief of the TRANSACTIONS ON IMAGE PROCESSING, A. Bovik, for his encouragement and support for this special issue. Special thanks to G. Kolakowski at UCSB for her help in patiently managing the administrative details, including collecting the papers and distributing them to the guest editors. Finally, this Special Issue would not have been possible without the help of the reviewers who managed to give their feedback on time to meet the strict time constraints. We express our sincere thanks to the reviewers listed below.

B. S. MANJUNATH
University of California
Santa Barbara, CA 93106 USA

T. HUANG
University of Illinois at Urbana-Champaign
Urbana, IL 61801 USA

A. M. TEKALP
University of Rochester
Rochester, NY 14627 USA

H. J. ZHANG
Microsoft Research China
Beijing, China

P. Salembier
C. Saraceno
S. Sclaroff
B. Shahraray
A. She
J. Smith
M. Smith
M. Swain
G. E. Sotak
P. Su
T. Syeda-Mahmood
C. Taskiran
C. Tomasi
A. Traina
A. Valey
N. Vasconcelos
V. Vinod
S. Wan
Y. F. Wang
J. Weng
T. Wiegand
W. Wolf
J. Wu
M. Yeung
Y. Xu
P. Yianilos
A. Zhang
X. Zhou