



PERGAMON

Pattern Recognition 35 (2002) 2671–2672

**PATTERN
RECOGNITION**

THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Foreword

Pattern recognition in information systems

Advances in information systems and information technology have dramatically transformed the way we live, work, communicate and learn. For instance, easy and rapid access to information (text, audio, images and video) via the Internet, the WWW, TV programming, etc. in digital format is having an enormous impact on research in education, science and engineering. To enhance the positive impact of this technology, multidisciplinary investigations must be carried out to address new system design and integration issues in various emerging applications. Huge amounts of information in digital format are already available; for example, the World Wide Web and digital libraries are immense sources of text, images, video and sound. Structuring and understanding this vast amount of multimedia information are difficult tasks.

Accessibility issues arise in this context, with personal identification techniques playing a key role in designing human-machine interfaces for security reasons and restricting access to privileged information. Pattern recognition (PR) and machine learning (ML) techniques provide a formal framework in which this class of problems can be adequately addressed. On the other hand, exploratory data analysis (data mining) and content-based retrieval of multimedia data serve both as important tools for PR and ML techniques, and as a source of challenging research problems. The efficient use of information, therefore, involves new ways of organizing, searching, processing, understanding and securing data as well as interacting with large data repositories.

This special issue contains a collection of papers addressing the multidisciplinary topics listed above. Both searching and organizing information involve the efficient and effective retrieval, feature extraction and automated analysis of multimedia data. The analysis and retrieval techniques used are closely related to the type of data being processed: images, text, sound, video or multimedia documents. Therefore, we organize and illustrate these topics in the special issue according to the type of media involved, namely image, text and general documents (text and figures) and video.

Efficient image indexing and access tools are of paramount importance in order to fully utilize the increasing amounts of digital image data, available on the Internet and in digital libraries. Content-based image retrieval (CBIR) approaches, based on properties automatically extracted

from images, provide a means to analyze, interpret, catalog and index image databases. Iqbal and Aggarwal present an approach to CBIR, integrating isotropic (represented by structure extraction and color histograms) and anisotropic mappings (based on texture analysis). A recent trend in CBIR focuses on applying learning techniques, especially active learning, in CBIR. The paper by Li, Chen and Zhang follows this direction, proposing a statistical correlation model that can accumulate semantic knowledge learned from relevance feedback information based on previous queries.

Accurate layout extraction is an important step in document analysis. The special spatial arrangement of Chinese documents, as compared to English documents, provides a significant challenge. The approach proposed by Xi, Hu and Wu consists of techniques for both text block segmentation and identification of Chinese documents. Juan and Vidal address text classification based on the application of mixtures of multivariate Bernoulli distributions to binary features.

Processing of video information is another source of challenging problems. The work by Nascimento and Marques concerns the tracking of objects with complex motion dynamics or shape changes in video sequences. Sadlier, Marlow, O'Connor and Murphy present a technique for automatic detection of TV advertisement, aiming at the development of efficient video browsing/viewing tools.

Information privacy and intellectual property protection aim at restricting access to value-sensitive information and to help owners track copying and distribution of digital media. Encryption and watermarking techniques are commonly used for this purpose. Encryption allows digital documents to be scrambled so that they can be unscrambled only by legitimate users. Watermarking techniques embed information (e.g. about ownership) into the digital documents such as an image. Biometric techniques are also utilized for user-authentication. Bolle, Connell and Ratha present a discussion on secure-biometrics-based user authentication and introduce a new technique of intentional, non-invertible distortions called cancellable biometrics to handle the problem of privacy protection. The paper by Gunsel, Uludag and Tekalp addresses watermarking techniques in digital security, presenting two spatial watermarking methods and their application to fingerprint images.

The foundations and principles of data mining and decision support systems include, for instance, algorithms and

methods for pre-processing, compression, feature selection and feature transformation, visualization, interpretation and classification of large-scale heterogeneous data sets. Two papers on this topic are included dealing with general PR techniques. Feature selection techniques are important for reducing data dimensionality and computational complexity and avoiding the curse of dimensionality. The paper by Somol and Pudil describes a software package for feature selection. A number of well-known feature selection algorithms are illustrated and compared on several data sets. Another paper by Rodriguez, Soraluze, Muguerza, Martin and Alvarez focuses on neighborhood-based classifiers, proposing the use of hierarchical classifiers to reduce the computational cost.

This special issue contains a subset of the papers presented at the First International Workshop on Pattern Recognition in Information Systems (PRIS 2001) held at the Escola Superior de Tecnologia, Setubal, Portugal, on July

6–7, 2001, under the sponsorship of the International Association of Pattern Recognition. We would like to thank all the authors and reviewers and express our gratitude to the Editor-in-Chief Robert S. Ledley and Managing Editor Blaire V. Mossman for their support.

Ana Fred
*Telecommunications Institute,
Instituto Superior Tecnico,
1049-001 Lisbon,
Portugal*
E-mail address: afred@lx.it.pt

Anil K. Jain
*Michigan State University,
Department of Computer Science and Engineering,
East Lansing, MI 48824, USA*
E-mail address: jain@cse.msu.edu