Preface

Multimedia uses multiple forms of information content and processing mainly text, audio, graphics, animation and video for communication to cater the various user demands. Multimedia presentation etc. are used today in movies, education, entertainment, marketing, advertising, information services, teleconferencing, publishing, interactive television, product demonstration, and alike. With a rapid transfer of information and a growing need to present this information in a powerful way, individuals who have appropriate skills and knowledge to effectively communicate will succeed in the multimedia industry.

In the last few years, multimedia processing has emerged as an important technology to generate contents based on images, audio, graphics, animation, full-motion video, and text, and it have opened a wide range of applications by combining these varieties of information sources thus giving insights in the interpretation of the multimedia content. Furthermore, the recent new development represented by the high definition multimedia content and interactive television will surely generate a huge volume of data and imply serious computing problems connected with the creation, processing and management of multimedia content. Multimedia processing is a challenging domain for several reasons as, for example: it requires both high computational processing requirements and memory bandwidth, it is a multi-rate computing problem, and it requires low-cost implementations for high-volume markets.

Computational intelligence is one of the most exciting and rapidly expanding fields which attract a large number of scholars, researchers, engineers and practitioners working in such areas as: rough sets, neural networks, fuzzy logic, evolutionary computing, artificial immune systems, and swarm intelligence. Computational intelligence has been a tremendously active area of research for the past decade or so. There are many successful applications of computational intelligence in many subfields of multimedia, including image processing, audio processing, and text processing or image retrieval. However, there still are numerous open problems in multimedia processing exemplified by multimedia communication, multimedia computing and computer animation that are in a desperate need for advanced and efficient computational methodologies to deal with huge amounts of data generated by those problems.

This volume provides an up-to-date and state-of-the-art coverage of diverse aspects related to computational intelligences in multimedia processing. It address the use of different computational intelligence based approaches to various problems in multimedia computing, networking and communications such as video processing, virtual reality, movies, audio processing, information graphics in multimodal documents, multimedia tasks scheduling, modeling interactive non-linear stories video authentication, text localization in images, organizing multimedia information, and visual sensor networks. This volume comprises of 18 chapters including an overview chapter providing an up-to-date and state-of-the review of the current literature on computational intelligence based approaches to various problems in multimedia computing and communication, and some important research challenges.

The book is divided into 5 parts devoted to: foundation of computational intelligence in multimedia processing, computational intelligence in 3D multimedia virtual environment and video games, computational intelligence in image/audio processing, computational intelligence in multimedia networks task scheduling; and computational intelligence in video processing.

The part on **foundation of computational intelligence in multimedia processing** contains two introductory chapters. It presents a broad overview of computational intelligence (CI) techniques including Neural Network (NN), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Fuzzy Set (FS), Reinforcement Learning (RL) and Rough Sets (RS). In addition, a very brief introduction to near sets and near images which offer a generalization of traditional rough set theory and a new approach to classifying perceptual objects by means of features in solving multimedia problems is presented. A review of the current literature on CI based approaches to various problems in multimedia computing, networking and communications. Challenges to be addressed and future directions of research are also presented.

Chapter 1, by *Aboul ella Hassanien, Ajith Abraham, and Janusz Kacprzyk*, , presents a review of the current literature on computational intelligence based approaches to various problems in multimedia computing such as speech, audio and image processing, video watermarking, content-based multimedia indexing and retrieval. The article also discuss some representative methods to provide inspiring examples to illustrate how CI could be applied to resolve multimedia computing problems and how multimedia could be analyzed, processed, and characterized by computational intelligence.

Chapter 2, by *Parthasarathy Guturu*, presents a review of the current literature on computational intelligence based approaches to various problems in multimedia networking and communications such as call admission control, management of resources and traffic, routing, multicasting, media composition, encoding, media streaming and synchronization, and on-demand servers and services

The part on **Computational intelligence in 3D multimedia virtual environment and video game** contains four chapters. It discusses the application of computational intelligence techniques in the area of virtual environment in which humans can interact with a virtual 3D scene and navigate through a virtual environment as well as music information retrieval approaches. Dynamic models are also employed to obtain a more formal design process for (story-driven) games and on improving the current approaches to interactive storytelling.

In Chapter 3, *Ronald Genswaider, Helmut Berger, Michae Dittenbach, Andreas Pesenhofer, Dieter Merkl, Andreas Rauber, and Thomas Lidy,* introduce the MediaSquare, a synthetic 3D multimedia environment that allows multiple users to collectively explore multimedia data and interact with each other. The data is organized within the 3D virtual world either based on content similarity, or by mapping a given structure (e.g. a branch of a file system hierarchy) into a room structure. With this system it is possible to take advantage of spatial metaphors such as relations between items in space, proximity and action, common reference and orientation, as well as reciprocity

In Chapter 4, *Tauseef Gulrez, Manolya Kavakli* and *Alessandro Tognetti*, developed a testbed for robot mediated neurorehabilitation therapy, that combines the use of robotics, computationally intelligent virtual reality and haptic interfaces. They employed the theories of neuroscience and rehabilitation to develop methods for the treatment of neurological injuries such as stroke, spinal cord injury, and traumatic brain injury. As a sensor input they have used two state of the art technologies, depicting the two different approaches to solve the mobility loss problem. In their experiment, a 52 piezo-resistive sensor laden shirt were used as an input device to capture the residual signals arising from the patient's body.

In Chapter 5, *Fabio Zambetta* builds the case for a story-driven approach to the design of a computer role playing game using a mathematical model of political balance and conflict and scripting based on fuzzy logic. The model introduced differs from a standard HCP (hybrid control process) by the use of fuzzy logic (or fuzzy state machines) to handle events, while an ordinary differential equation is needed to generate continuous level of conflict over time. By using this approach, not only can game designers express game play properties formally using a quasi-natural language, but they can also propose a diverse role-playing experience to their players. The interactive game stories designed with this methodology can change under the pressure of a variable political balance, and propose a different and innovative game play style.

Time flow is the distinctive structure of various kinds of data, such as multimedia movie, electrocardiogram, and stock price quote. To make good use of these data, locating desired instant or interval along the time is indispensable. In addition to domain specific methods like automatic TV program segmentation, there should be a common means to search these data according to the changes along the time flow. Chapter 6, by Ken Nakayama et. al presents *I*-string and *I*-regular expression framework with some examples and a matching algorithm. *I*-string is a symbolic string-like annotation model for continuous media which has a virtual continuous branchless time flow. *I*-regular expression is a pattern language over *I*-string, which is an extension of conventional regular expression for text search. Although continuous media are often treated as a sequence of time-sliced data in practice, the framework adopts continuous time flow. This abstraction allows the annotation and search query be independent from low-level implementation such as frame rate.

Computational intelligence in image/audio processing is the third part of the book. It contains six chapters discussing the application of computational intelligence techniques in image and audio processing.

In Chapter 7, *Barca J.C., Rumantir G and Li R.* present a set of illuminated contour-based markers for optical motion capture has been presented along with a modified K-means algorithm that can be used for removing inter-frame noise. The new markers appear to have features that solve and/or reduce several of the drawbacks associated with other marker systems currently available for optical motion capture. The new markers provide solutions to central problems with current standard spherical flashing LED based markers. The modified K-means algorithm that can be used for removing noise in optical motion capture data is guided by constraints on the compactness and

number of data points per cluster. Experiments on the presented algorithm and findings in literature indicate that this noise removing algorithm outperforms standard filtering algorithms such as the mean and median because it is capable of completely removing noise with both the spike and Gaussian characteristics.

In Chapter 8, *Sandra Carberry* and *Stephanie Elzer* present a corpus study that shows the importance of taking information graphics into account when processing a multimodal document. It then presents a Bayesian network approach to identifying the message conveyed by one kind of information graphic, simple bar charts, along with an evaluation of the graph understanding system.

In Chapter 9, *Klaas Bosteels* and *Etienne E. Kerre* present a recently introduced triparametric family of fuzzy similarity measures, together with several constraints on its parameters that warrant certain potentially desirable or useful properties. In particular, they present constraints for several forms of restrictability, which allow reducing the computation time in practical applications. They use some members of this family to construct various audio similarity measures based on spectrum histograms and fluctuation patterns.

Chapter 10, by *Przemyslaw Gorecki*, *Laura Caponetti* and *Ciro Castiello*, deals with the particular problem of text localization, which aims at determining the exact location where the text is situated inside a document image. The strict connection between text localization and image segmentation is highlighted in the chapter and a review of methods for image segmentation is proposed. Particularly, the benefits of this chapter and the employment of fuzzy and neuro-fuzzy techniques in this field are assessed, thus indicating a way to combine computational intelligence methods and document image analysis. Three peculiar methods based on image segmentation are presented to show different applications of fuzzy and neuro-fuzzy techniques in the context of text localization.

In Chapter 11, *Kui Wu and Kim-Hui Yap*, present a soft-labeling framework that addresses the small sample problem in interactive CBIR systems. The technique incorporates soft-labeled images into the fuzzy support vector machine (FSVM) for effective learning along with labeled images for effective retrieval. By exploiting the characteristics of the labeled images, soft-labeled images are selected through an unsupervised clustering algorithm. Further, the relevance of the soft-labeled images is estimated using the fuzzy membership function. The FSVM-based active learning is then performed based on the hybrid of soft-labeled and explicitly labeled images. Experimental results based on a database of 10,000 images demonstrate the effectiveness of the proposed method.

Temporal textures are textures with motion like real world image sequences of sea-waves, smoke, etc. that possess some stationary properties over space and time. The motion assembly by a flock of flying birds, water streams, fluttering leaves, and waving flags also serve to illustrate such a motion. The characterization of temporal textures is of a vital importance to computer vision, electronic entertainment, and content-based video coding research with a number of potential applications in areas including recognition (automated surveillance and industrial monitoring), synthesis

(animation and computer games), and segmentation (robot navigation and MPEG-4). Chapter 12 by *Ashfaqur Rahmanand* and *Manzur Murshed*, provides a comprehensive literature survey of the existing temporal texture characterization techniques.

Part-IV in Computational intelligence in multimedia networks and task scheduling contains four chapters that describe several approaches to develop video analysis and segmentation systems based on visual sensor networks using computational intelligence as well as a discussion about detecting hotspots in the cockpits in view of the Swissair 111 and ValuJet 592 flight disasters, and answer the question that how distributed sensor networks could help in near real-time event detection, disambiguate faults and events by using artificial intelligence techniques. In addition, a review chapter of the current literature on computational intelligence based approaches to various problems in multimedia networking and communications

In Chapter 13, *Mitsuo Gen* and *Myungryun Yoo*, discuss a task scheduling problem by introducing many scheduling algorithms for soft real-time tasks using a genetic algorithm (GA). They propose reasonable solutions for NP-hard scheduling problem which much less difficulties than those solved by traditional mathematical methods. In addition, a continuous task scheduling, real-time task scheduling on homogeneous system and real-time task scheduling on heterogeneous system are discussed and included in this chapter.

Chapter 14 by *Miguel A. Patricio*, *F. Castanedo, A. Berlanga, O. Pérez, J. Garćıa*, and *José M. Molina,* describe several approaches to develop video analysis and segmentation systems based on visual sensor networks using computational intelligence. They discuss how computational intelligence paradigms can help obtain competitive solutions. The knowledge about the domain is used in the form of fuzzy rules for data association and heuristic evaluation functions to optimize the design and guide the search of appropriate decisions.

In Chapter 15, *Slawomir T. Wierzchon, Krzysztof Ciesielski,* and *Mieczyslaw A. Klopotek*, focus on some problems concerning application of an immune-based algorithm to extraction and visualization of cluster structure. The chapter presents a novel approach, based on artificial immune systems, within a broad stream of map type clustering methods. Such approach leads to many interesting research issues, such as context-dependent dictionary reduction and keywords identification, topic-sensitive document summarization, subjective model visualization based on particular user's information requirements, dynamic adaptation of the document representation and local similarity measure computation

In Chapter 16, *S. Srivathsan, N. Balakrishnan* and *S.S. Iyengar*, discuss some safety issues in commercial planes particularly focusing on hazards in the cockpit area. The chapter discusses a few methodologies to detect critical features and provide unambiguous information about the possible sources of hazards to the end user in near real time. They explore the application of Bayesian probability, the Iyengar-Krishnamachari method, probabilistic reasoning, reasoning under uncertainty, and the

Dempster-Shafer theory, and analyze how these theories could help in the data analysis gathered from wireless sensor networks deployed in the cockpit area.

The final part of the book deals with the use of **computational intelligence in video processing.** It contains three chapters which discuss the use of computational intelligence techniques in video processing.

In Chapter 17, *Nicholas Vretos, Vassilios Solachidis* and *Ioannis Pitas,* provide a uniform framework by which media analysis can be rendered more useful for retrieval applications as well as for human-computer interaction based application. All the algorithms presented in this chapter are focused on humans and thus provides interesting features for an anthropocentric analysis of a movie.

In Chapter 18, *Thomas Bärecke*, *Ewa Kijak, Marcin Detyniecki,* and *Andreas N˙ürnberger*, present an innovative way of automatically organizing multimedia information to facilitate content-based browsing. It is based on self-organizing maps. The visualization capabilities of the self-organizing map provide an intuitive way of representing the distribution of data as well as the object similarities. The main idea is to visualize similar documents spatially close to each other, while the distance between different documents is bigger. They introduce a novel time bar visualization that re-projects the temporal information.

In chapter 19, *Mayank Vatsa, Richa Singh, Sanjay K. Singh and Saurabh Upadhyay,* present an efficient intelligent video authentication algorithm using support vector machine. The proposed algorithm which can detect multiple video tampering attacks. It computes the local relative correlation information and classifies the video which or non-tampered. The proposed algorithm computes the relative correlation information between all the adjacent frames of a video and projects them into a non-linear SVM hyperplane to determine if the video is tampered or not. The algorithm is validated on an extensive video database containing 795 tampered and non-tampered videos. The results show that the proposed algorithm yields a classification accuracy of 99.2%.

Acknowledgements

Aboul-Ella Hassanien, Janusz Kacprzyk and Ajith Abraham (Editors)