Editorial:
Special Issue on Soft Computing in Multimedia Processing

I. Introduction

Soft Computing (SC) is an emerging field that consists of complementary elements of fuzzy logic, neural computing, evolutionary computation, machine learning and probabilistic reasoning. Due to their strong learning and cognitive ability and good tolerance of uncertainty and imprecision, soft computing techniques have found wide applications. Needless to say, multimedia data (video, image, audio, text, color, etc.) is one of these applications.

Multimedia processing is a very important scientific research domain with a broadening range of applications. The development of new insights and applications results from both fundamental scientific research and the development of new technologies. One of these emerging technologies is soft computing, which is a generic term for a specific collection of tools to model uncertainty, imprecision, evolutionary behavior and complex models.

This special issue is devoted to the recent developments in the applications of soft computing (SC) techniques to multimedia processing. We received 16 papers, of which 6 were accepted for publication. The topics covered in this issue cover a wide range of research areas of soft computing in multimedia processing including video sequence, color quantization, image retrieval, meeting video, document image analysis, image segmentation and biometric application.

II. Scanning through the issue

Effective and efficient representation of video sequences is an important multimedia analysis challenging task for video retrieval and browsing applications. The paper by Lang Gongyan, Xu De and Yang Xu introduces a new approach for the prominent region detection from the viewpoint of the human perception intending to construct a good pattern for content representation of the video sequences. It starts by partitioning each frame into homogeneous regions using a technique based on a non-parameter clustering algorithm, then extracts a number of different mise-en-scene-based perceptual features which influence human visual attention in order to automatically determine the prominent importance of the different homogenous regions in a frame. Finally, a modified Fuzzy Inference Neural Networks is used to detect prominent regions in video sequence due to its simple structure and superior performance for automatic fuzzy rules extraction. The extracted prominent regions could be used as a good pattern to bridge semantic gap between low-level features and semantic understanding. Experimental results show the excellent performance of the approach.

The popularity of the world wide web has emerged as the largest repository of multimedia data in the world. One form of information that is very popular on the web today is the digital color image. This includes both single- and multi-frame (video) images. While many forms of information and data can be transferred quickly using the web, the transfer of digital images can be very time consuming due to their inherent size. To speed up this process, images are commonly compressed before being stored at the local site or transmitted across the internet. But the compression of digital images is not a straight forward process. Color image quantization or simply color quantization, is a form of image compression which reduces the number of colors used in an image while maintaining, as much as possible, the appearance of the original. This type of compression does not allow the original image to be reproduced, however, from the compressed image. The optimal goal in the color quantization process is to produce an image which can not be distinguished from the original. Thus, a color quantization algorithm attempts to approximate the optimal solution. The paper by Mahamed Omran, Andries Engelbrecht and Ayed Salem deals with the color image quantization problem. It is based on Particle Swarm Optimization algorithm (PSO). The proposed algorithm randomly initializes each particle in the swarm to contain K centroids. The K-means clustering algorithm is then applied to each particle at a user-specified probability to refine the chosen centroids. Each pixel is then assigned to the cluster with the closest centroid. The PSO is then applied to refine the centroids obtained from the K-means algorithm. The proposed algorithm is then applied to commonly used images. It is shown from the conducted experiments that the proposed algorithm generally results in a significant improvement of image quality compared to other well-known approaches

With the development of the Internet and database techniques, information retrieval (IR) becomes very popular. As a powerful form of delivering information, multimedia data is frequently used in many domain applications. Techniques for effectively dealing with multimedia databases management are useful and in demand. Dianhui Wang and Xiaohang Ma developed a hybrid scheme for intelligent image retrieval using neural nets. Each item in an image database is indexed by a visual feature vector, which is extracted using color moments and discrete cosine transform coefficients. Query is characterized by a set of semantic labels, which are predefined by system designers and associated with domain concerns. The system utilizes the image content features as the system input, and the semantic labels as its output. To compensate the deficiency of semantics modeling, an on-line user's relevance feedback is applied to improve the retrieval performance of the hybrid

intelligent retrieval system. The neural net acts like a pattern association memory bank that maps the low-level feature vectors into their corresponding semantic labels. During retrieval process, the weights of the neural net are updated by an interactive user's relevance feedback technique, where the feedback signal comprise the neural net actual output, semantic labels provided by users and the given query. A prototype hybrid intelligent retrieval system and evaluated using an artificial image database

Meeting videos are important multimedia documents consisting of captured meetings in specialized smart room environments. Research activities cover for instance recording, representing, and browsing of meeting videos. Speech can be very useful cue in indexing videos, but precise speech recognition in meting rooms remains a challenging task because of extensive vocabulary topics, speech styles and so on. The sound cue can also be used in teleconferencing scenarios to identify the speaker and to improve the tracking performance. Indexing videos using visual content is also a challenging task. On the basis of visual cues it is possible to recognize what single participants are doing throughout the meeting. The paper by Bogdan Kwolek deals with the action recognition meeting videos using the head trajectory and fuzzy color histogram where the knowledge was extracted from such video. The tracking of the head is done using a particle filter built on cues such as color, gradient and shape. The head is represented by an ellipse with fuzzy color histogram in its interior and an intensity gradient along the ellipse boundary. By comparing pixels in entry zones to a model of the background we can detect the entry of the person quickly and reliable. The fuzzy color is constructed in the interior of an ellipse fitting best the oval shape of the head. When a new person appears in the scene a creation of new trajectory is initialized. The recognition of actions is performed using kernel histograms built on head positions as well as segmented trajectories that are related to the layout of the room.

Document analysis or more precisely, document image analysis, is the process that performs the overall interpretation of document images. Document image processing is now an established field within the electronic imaging world. It is becoming even more prevalent in an area where paper documents need to be transformed into electronic format for long term storage, backup, multiple access and retrieval. The process of extracting information from often poor quality images of documents is a topic of active research. In a multimedia environment where sound, moving images and graphics could be part of a compound document, the role of image processing becomes even more important. The paper by Andras Barta and Istvan Vajk presents a hierarchical object recognition system for document image processing. It is based on a spatial tree structure representation and Bayesian network framework. The image components are built up from lower level image components stored in a library. The tree representations of the objects are assembled from these components. A probabilistic framework is used in order to get robust behaviour. The method is able to convert general circuit diagrams to their components and store them in a hierarchical data-structure. The paper presents simulation for extracting the components of sample circuit diagrams.

The utilization of digital techniques in the creation, editing and distribution of multimedia data offers a number of opportunities to a pirate user, such as high fidelity copying. Furthermore, the widespread usage of Internet is providing additional channels for a pirate to quickly and easily distribute the copyrighted digital content without the fear of being tracked. As a result, the protection of multimedia content (image, video, audio, etc.) is now receiving a substantial amount of attention. Digital fingerprinting is an emerging technology to protect multimedia from unauthorized redistribution. The paper by Mohamed Mostafa deals with the problem of authentication. It presents a novel and fast fingerprint identification technique, which uses a novel clustering algorithm to detect similar feature groups from multiple template images generated from the same finger and create the cluster core set. It is based on a new supervised recurrent neural-network. A quick response was achieved by manipulating the search order inside the experimental databases. The experiments results demonstrate that the similarity search approach with neural networks proves suitable one-to many matching of fingerprints on large databases.

Acknowledgments.

The team of Guest Editors would like to take this opportunity to thank all those authors who submitted papers, and all of the reviewers who took such care in reviewing these papers, as well as the Editor-In-Chief and professor Anton P. Zeleznikar and professor Matjaz Gams managing Editor who give us the permission to edit this issue and their support and guidance in the publication. As Guest editors, we hope that the papers in this issue will stimulate further progress in this direction. We believe that the best is yet to come.

Karen Egiazarian and Aboul Ella Hassanien