

Preface

Multimedia refers to the representation of information in varied forms. Two broad categories of multimedia information are prevalent. These are (1) discrete media and (2) continuous media. Information becomes multimedia information if and only if it contains at least a continuous media. The most common and enormous continuous media type is the video. A lot of time-complex computing facility is required for processing video contents given its variety and enormity of underlying data.

Video content classification is an important task in the computer vision community as far as intelligent analysis of video content is concerned. Typical applications include video content mining, video indexing (for media industry), video surveillance and defense applications to name a few. A score of literatures exists in this regard providing a detailed comparative study of the different techniques in vogue. Video content classification primarily entails the detection of the cuts in the video sequences through classical as well as non-classical techniques. The non-classical techniques resort to computational intelligence perspectives in ascertaining the video shot boundaries by using the inherent information distribution of the video content. Hence, larger video frames/sequences obviously require a lot of computational overhead in the shot detection task. Since, video data are essentially continuous in nature, dynamic estimation of the data distribution is a prerequisite in the commonly used shot boundary detection and video analysis techniques.

This edited volume aims to open up a new realm of research in media industry helping practitioners to securely and faithfully process their multimedia data as is evident from the table of contents. Almost no book is existent in the market as of now in this related field. Moreover, the use of intelligent techniques for processing of multimedia data would give rise to robust and time efficient systems especially when multimedia data are very bulky in nature. Add to it, the book would also serve as a treatise for effective transmission of multimedia data across reliable networks.

This edited volume is intended to bridge the gap in the media-researchers' community by introducing newer avenues in the form of recent trends in research areas in continuous media processing and manipulation. In fact, the objectives of the proposed publication are many fold. These may be summarized according to their priorities.

1. Till date, the existing techniques for video segmentation are able to detect with high accuracy the hard cuts present in a video. This book seeks avenues for advancements in these techniques to be applied to live streaming videos for on-the-fly video segmentation.
2. The problem of setting an automatic threshold for video segmentation has to be addressed. The threshold is to be set dynamically without any a priori knowledge about the type, content or length

- of the video. Newer research trends would surely target this untouched area and thereby open a new era of video summarization of real time events, such as producing highlights of sports videos.
3. Detection of representative frames (key-frames) which denote the semantic content of the shots in a video which are used for video indexing is also a challenging proposition. Redundancy reduction is an important step as it removes uninformative or redundant video frames/segments and retains the most informative parts which are concatenated to form a video summary. This book strives to seek newer research initiatives in this direction.

The intended readers of this book are the relevant research community and the media practitioners. To be precise, the book is aimed to establish the missing link between the research standing in the relevant field and that is upcoming. Infusing intelligent paradigms for multimedia information processing would surely and certainly help the readers grasp the essence and utility of the different intelligent techniques in vogue for faithful understanding of multimedia information.

This book would come to the benefits of several categories of students and researchers. At the students' level, this book can serve as a treatise/reference book for the special papers at the masters level aimed at inspiring possibly future researchers. Newly inducted PhD aspirants would also find the contents of this book useful as far as their compulsory coursework is concerned.

At the researchers' level, those interested in interdisciplinary research would also benefit from the book. The enriched interdisciplinary contents of the book would always be a subject of interest to the faculties, existing research communities and new research aspirants from diverse disciplines of the concerned departments of premier institutes across the globe. This is expected to bring different research backgrounds (due to its cross platform characteristics) close to one another to form effective research groups all over the world.

The edited volume comprises 16 well versed and self-contained chapters from diverse research domains ranging from fundamentals of multimedia to image, audio and video processing to data compression.

Chapter 1 discusses about how to analyze a video or an image by using several methods like scene analysis, and shot boundary detection and analysis, frame analysis, hypermedia representation, segmentation of media. The main focus of the chapter is to discuss as to how to use the above mentioned methods in any of the media video or image and extract the information required to represent the whole media under process.

Chapter 2 discusses the basic aspects of image segmentation and an application and presents a detailed assessment on the different methods in image segmentation and discusses a case study on it.

Feature extraction in a multimedia environment deals with two complicated tasks first deciding and then extracting. There are certain properties expected from good features viz., Repeatability, Distinctiveness, Locality, Quantity, Accuracy, Efficiency, and Invariance. Chapter 3 describes the different feature extraction techniques. It also concentrates on taking a survey on the topic of Feature extraction and Image formation.

Particle Swarm Optimization (PSO) is a well-known swarm optimization technique. It is very efficient to optimize the image segmentation problem. PSO algorithm has some drawbacks as the possible solutions may follow the global best solution at one stage. As a result, the probable solutions may be bound within locally optimized solutions. Chapter 4 tries to get over the drawback of the PSO algorithm and proposes a Modified Particle Swarm Optimization (MfPSO) algorithm to segment the multilevel

Preface

images. The proposed method is compared with the original PSO algorithm and the renowned k-means algorithm. Comparison of the above mentioned existing methods with the proposed method are applied on three real life multilevel gray scale images and the results are demonstrated.

Chapter 5 focuses on the concept of Content based image retrieval. Searching of an image or video database based on text based description is a manual labor intensive process. Descriptions of the file are usually typed manually for each image by human operators because the automatic generation of keywords for the images is difficult without incorporation of visual information and feature extraction. This method is impractical in today's multimedia information era. "Content-based" means that the search will analyze the actual contents of the image rather than the metadata such as keywords, tags, and descriptions associated with the image. The term 'content' in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. Several important sections are highlighted in this chapter, like architectures, query techniques, multidimensional indexing, video retrieval and different application sections of CBIR.

Among the various applications like physics, security, photonics, biomedical, astronomy, remote sensing, ecological, environmental, etc.; biomedical is one of the many important areas researchers are focusing on. So for the intelligent analysis of multimedia information like biomedical image, Chapter 6 covers this thrust area. This chapter therefore aids both biomedical engineers and non-technical people to get an overview of the basics of biomedical image processing and analysis.

In Chapter 7 different solutions are presented to develop algorithms for digital image processing focusing particularly on edge detection, which is one of the most important phases used in computer vision and image processing applications and also in human image understanding. In this chapter, implementation of classical edge detection algorithms is presented and also implementation of algorithms based on the theory of Cellular Automata (CA) is demonstrated.

Image segmentation among overlapping land cover areas in satellite images is a very crucial task. Detection of belongingness is the important problem for classifying mixed pixels. Chapter 8 proposes an approach for pixel classification using a hybrid approach of Fuzzy C-Means and Cellular automata methods. This new unsupervised method is able to detect clusters using 2-Dimensional Cellular Automata model based on fuzzy segmentations. This approach detects the overlapping regions in remote sensing images by uncertainties using fuzzy set membership values.

Fuzzy classification techniques are used for image classification for quite a long time back by allowing pixels to have membership in more than one class. However, handling information at the pixel level is time consuming and there is a high chance of biased assessment of images if class labels are assigned by a single human observer. An individual's perception in assessing quality of images is not reflected even if multiple observers' opinions are considered. In Chapter 9, the Fuzzy Relational Classifier (FRC) is used to assess quality of images distorted by information loss or noise, unlike the existing methods where images are preprocessed to remove the noise before classification.

eHealth is a set of systems and services that enable the sharing of medical diagnostic imaging data remotely. The application of eHealth solves the problem of the lack of specialized personnel, unnecessary execution of multiple diagnostic imaging and rapid exchange of information and remote diagnostics. Medical imaging generates large amounts of data. An MRI study can contain up to several Gigabytes (GB). The exchange of such large amounts of data in the local network facilities is a significant problem due to bandwidth sharing which is even more significant in mobile and wireless networks. A possible

solution to this problem is data compression with the requirement that there is no loss of data. The goal of Chapter 10 is a conceptual compression prototype that will allow faster and more efficient exchange of medical images in systems with limited bandwidth and communication speeds (cellular networks, wireless networks). To obtain this conceptual compression prototype we will use wavelets.

During video editing, the shots composing the video are coalesced together by different types of transition effects. These transition effects are classified abrupt and gradual transitions based on the inherent nature of these transitions. In abrupt transitions, there is an instantaneous change in the visual content of two consecutive frames. Gradual transitions are characterized by a slow and continuous change in the visual contents occurring between two shots. In Chapter 11, the challenges faced in this field along with an overview of the different approaches are presented. Also, a novel method for detection of dissolve transitions using a two-phased approach is enumerated.

Music listening is one of the most common things of human behaviors. Normally mobile music are downloaded to mobile phones and played by mobile phones. Today millennial people use mobile music in about all the age groups. Music recommendation system enhances personalized music classifications that create a profile with the service and build up a music library based on the choice preferences using mobile cloud services. Music recommendation through cloud is therefore an emerging field, and this can be done using various parameters like song genre similarity, human behavior, human mood, song rhythmic patterns, seasons etc. Chapter 12 presents an intelligent music recommender system that identifies the raga name of one particular song music and then mapping with the raga time database and classifies the songs according to their playing time that create time slot based personalized music libraries.

Over the past decade, research in the field of Content-Based Video Retrieval Systems (CBVRS) has attracted much attention as it encompasses processing of all the other media types i.e. text, image and audio. Video summarization is one of the most important applications as it potentially enables efficient and faster browsing of large video collections. A concise version of the video is often required due to constraints in viewing time, storage, communication bandwidth as well as power. Thus, the task of video summarization is to effectively extract the most important portions of the video, without sacrificing the semantic information in it. The results of video summarization can be used in many CBVRS applications like semantic indexing, video surveillance copied video detection etc. However, the quality of the summarization task depends on two basic aspects: content coverage and redundancy removal. These two aspects are both important and contradictory to each other. Chapter 13 aims to provide an insight into the state-of-the-art approaches used for this booming field of research.

Biometric system is used by many institutions, organizations and industries for automatic recognition of person. One of the main reasons for popularity of used for biometric system is that the ability of the system to identify between an authorized person and unauthorized person. There are many challenges associated with the biometric system such as designing of human recognition algorithm, compression of biometric templates, privacy and security of biometric templates in biometric systems. Chapter 14 gives an application of Compressive Sensing (CS) theory for solutions of the above mentioned challenges in biometric systems.

Chapter 15 introduces a methodology on how to use analytical potential of multimedia contents like YouTube, Bing Videos or Vimeo for discovering behavioral consumer characteristics. The chapter also enumerates on to how to consolidate unstructured text data sources from blogs and Twitter with revealed knowledge from multimedia contents for better understanding consumer habits and needs.

Preface

Nowadays, multimedia contents such as images, videos, e-books and texts are easily available for download on the internet. Duplication of multimedia contents is done using different software. This type of operation sometimes creates problem of copyright and ownership authentication. Digital watermarking technique is one of the solutions for providing protection to multimedia contents. Chapter 16 gives various watermarking techniques in the transformation and sparse domain for protection of multimedia contents. This chapter demonstrates various watermarking techniques such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Singular Value Decomposition (SVD), Fast Discrete Curvelet Transform (FDCT), and CS theory based technique.

The primary objective of the book is to bring a broad spectrum of multimedia application domains under the purview of intelligent techniques so that it is able to trigger further inspiration among various research communities and contribute in their respective fields of applications. Thereby, these application fields may be oriented towards using intelligence techniques.

Once the purpose stated above, is achieved, a larger number of research communities may be brought under one umbrella to ventilate their ideas in a more structured manner. Thus, the present endeavor may be seen as the beginning of such an effort in bringing various research domains close to one another.

Siddhartha Bhattacharyya
RCC Institute of Information Technology, India

Hrishikesh Bhaumik
RCC Institute of Information Technology, India

Sourav De
The University of Burdwan, India

Goran Klepac
*University College for Applied Computer Engineering Algebra, Croatia & Raiffeisenbank Austria,
Croatia*

