Preface

Multimedia is an interdisciplinary, application-oriented technology that capitalizes on the multi-sensory nature of humans and the ability of computers to store, manipulate, and convey non-numerical information such as video, graphics, and audio in addition to numerical and textual information. Multimedia has the intrinsic goal of improving the effectiveness of computer-to-human interaction and, ultimately, of human-to-human communication.

While the Human Computer Interaction (HCI) community looks predominantly at the application layer and the telecommunications community at the lower end of the ISO OSI stack, little work has been published in bridging the gap between these two communities. Indeed, the human element is often neglected in Quality of Service (QoS) negotiation protocols. Not only does this have a negative and undesirable impact on the user's experience of multimedia, it also discards the potential for more economical resource allocation strategies With the proliferation of ubiquitous multimedia in predominantly bandwidth-constrained environments, more research is needed towards integrating and mapping perceptual considerations across the protocol stack and building truly end-to-end multimedia communication solutions.

Objectives and Structure of This Book

The book provides a comprehensive review of state-of-the-art in the area of perception-based multimedia design, covering both theories and applications in areas such as virtual reality, multi-sensory-based computing, ubiquitous communications, personalization, and adaptation according to user perceptual needs. The book covers a wide spectrum of up-to-date issues in the field, providing insight into analytical and architectural aspects of perception-based multimedia systems, as well as applications in real-world settings.

The book is structured in four main sections, reflecting the varied themes and interdisciplinarity associated with perceptual multimedia. Thus, the first section deals with perceptual modeling in multimedia, which is an important issue if one is to truly integrate human perceptual requirements into multimedia design. While multimedia is traditionally associated with the auditory and visual senses, there is no intrinsic reason why multimedia should not cover the other three senses of humans, and research done in this respect forms the focus of the second section of the book. The following section presents work dealing with a core issue at the heart of perceptual multimedia, namely the integration of human factors in the design and development of multimedia applications. Multimedia applications need to adapt at both ends: to the different characteristics and needs of the user, and to the changing requirements of the communication medium. This is the focus of the last section of the book, which following advances in networking technologies, presents research in multimedia communication and adaptation.

Perceptual Modeling in Multimedia

Understanding and modeling perceptual requirements is a challenging but very much exciting issue, due to the fickle nature of humans, their wants, and their needs. To address this issue, this section starts with **Chapter I**, by Cavallaro and Winkler, which presents work done in respect of emulating human vision. Their idea is to prioritize the visual data in order to improve the compression performance of video coders and the prediction performance of perceptual quality metrics. They propose an encoder and quality metric which incorporate visual attention and use a semantic segmentation stage, which takes into consideration some aspects of the cognitive behavior of people when watching multimedia video. Their semantic model corresponds to a specific human abstraction, which need not necessarily be characterized by perceptual uniformity. Since the semantics are defined through human abstraction, the definition of the semantic partition depends on the task to be performed, and they present the results of evaluating their proposed model in respect of segmenting moving objects and faces in multimedia video.

Chapter II summarizes the work on Mathematics of Perception performed by Chang's research team between 2000 and 2005. To support personalization, a search engine must comprehend users' query concepts (or perceptions), which are subjective and complicated to model. Traditionally, such query-concept comprehension has been performed through a process called "relevance feedback." Their work formulates relevance feedback as a machine-learning problem when used with a small, biased training dataset. His team has pioneered and developed a method of query-concept learning as the learning of a binary classifier to separate what a user wants from what she or he does not want, sorted out in a projected space. They have developed and published several algorithms to reduce data dimensions, to maximize the usefulness of selected training instances, to conduct learning on imbalanced datasets, to accurately account for perceptual similarity, to conduct indexing and learning in a non-metric, high-dimensional space, and to integrate perceptual features with keywords and contextual information.

In **Chapter III**, May sketches a semantic taxonomy of representational forms (or "sign types" in the terminology of semiotics) relevant for the compositional analysis, design,

and indexing of multimodal multimedia, specifically in the domain of instructional design of learning objects, and contributes to current attempts to bridge the "semantic gap" between the technical and spatio-temporal description of media objects and the contextual perception and interpretation by users of their content.

Chapter IV rounds off this section. Gulliver proposes a model of distributed multimedia, to allow a more structured analysis of the current literature concerning video and audio information. The model segregates studies implementing quality variation and/or assessment into three discrete information abstractions (the network, media, and content levels) and from two perspectives (the technical and user perspectives). In so doing, Gulliver places current research in the context of a quality structure, and thus highlights the need for fuller incorporation of the user perspective in multimedia quality assessment.

Multimedia and the Five Senses

Gulliver's chapter is a good link to this section, as his model incorporates four of the human senses, olfactory (smell), tactile / haptic (touch), visual (sight), and auditory (sound), and naturally leads to work done in respect of making multimedia applications truly multi-sensual and multimodal. **Chapter V**, by Axelrod and Hone, explores affective computing, illustrating how innovative technologies are capable of emotional recognition and display. Research in this domain has emphasized solving the technical difficulties involved, through the design of ever more complex recognition algorithms. But fundamental questions about the use of such technology remain neglected. Can it really improve human-computer interaction? For which types of application is it suitable? How is it best implemented? What ethical considerations are there? This field is reviewed in this chapter, and the authors discuss the need for user-centered design in multimedia systems. They then conclude the chapter by describing and giving evidence from a study that explored some of the user issues in affective computing.

In **Chapter VI**, Bussell examines the use of haptic feedback in a multimedia simulation as a means of conveying information about physical science concepts. The case study presented investigates the effects of force feedback on children's conceptions of gravity, mass, and related concepts following experimentation with a force-feedback-enabled simulation. Guidelines for applying these technologies effectively for educational purposes are discussed. This chapter adds to the limited research on the application of haptic feedback for conceptual learning, and provides a basis for further research into the effects of computer-based haptic feedback on children's cognition.

In **Chapter VII**, Giannakis investigates the use of visual texture for the visualization of multi-dimensional auditory information. In his work, 20 subjects with a strong musical background performed a series of association tasks between high-level perceptual dimensions of visual texture and steady-state features of auditory timbre. The results indicated strong and intuitive mappings among (a) texture contrast and sharpness, (b) texture coarseness-granularity and compactness, and (c) texture periodicity and sensory dissonance. The findings contribute in setting the necessary groundwork for the application of empirically-derived auditory-visual mappings in multimedia environments.

Chapter VIII winds up this section. In their chapter, Crosby and Ikehara describe their research focused on deriving changing cognitive state information from the patterns of data acquired from the user with the goal of using this information to improve the presentation of multimedia computer information. Detecting individual differences via performance and psychometric tools is supplemented by using real-time physiological sensors. Described is an example computer task that demonstrates how cognitive load is manipulated. The different types of physiological and cognitive state measures are discussed, along with their advantages and disadvantages. Experimental results from eye tracking and the pressures applied to a computer mouse are described in greater detail. Finally, adaptive information filtering is discussed as a model for using the physiological information to improve computer performance. Study results provide support that we can create effective ways to adapt to a person's cognition in real time and thus facilitate real-world tasks.

Human Factors

Multimedia technology is widely applied to Web-based applications, which are faced by diverse users, in terms of their background, skills, and knowledge, and, we would argue, perceptual requirements. Therefore, human factors represent an essential issue in the design of perceptual multimedia systems, which is addressed in this section.

In **Chapter IX**, Ghinea and Chen investigate the effect of cognitive style on the perceived quality of distributed multimedia. They use two dimensions of cognitive style analysis – field dependent/independent and verbalizer/visualizer – and the quality of perception metric to characterize the human perceptual experience. This is a metric which takes into account multimedia's "infotainment" (combined informational and entertainment) nature, and comprises not only a human's subjective level of enjoyment with regards to multimedia content quality, but also his/her ability to analyze, synthesizes, and assimilate the informational content of such presentations. Their results show that multimedia content and dynamism are strong factors influencing perceptual quality.

Chapter X, by Kalyuga, provides an overview of theoretical frameworks and empirical evidence for the design of adaptive multimedia that is tailored to individual levels of user expertise to optimize cognitive resources available for learning. Recent studies indicate that multimedia design principles which benefit low-knowledge users may disadvantage more experienced ones due to the increase in the cognitive load required for the integration of presented information with the available knowledge base. The major implication for multimedia design is the need to tailor instructional formats to individual levels of expertise. The suggested adaptive procedure is based on empirically-established interactions between levels of user proficiency and formats of multimedia presentations (the expertise reversal effect), and on real-time monitoring of users' expertise using rapid cognitive diagnostic methods.

Chapter XI, by Tilinger and Sik-Lanyi, presents the differences between left- and right-handed persons in the aspect of computer-presented information and virtual realities. It introduces five test scenarios and their results addressing this question. They show

that there are moderate differences between groups preferring different hands, and that the different needs of left- and right-handed people may play an important role in user-friendly interface and virtual environment design, since about a tenth of the population is left-handed. This could help to undo the difficulties which left-handed and ambidextrous individuals routinely encounter in their daily lives.

Multimedia Communication and Adaptation

To address the diverse needs of users and the underlying communication networks, adaptation is a necessity in distributed multimedia systems. The fluctuating bandwidth and time-varying delays of best-effort networks makes providing good quality streaming a challenge; when such quality has to take into account perceptual considerations, the problem is compounded. This section kicks off with **Chapter XII**, by Cranley and Murphy. Their work describes research that proposes that an optimal adaptation trajectory through the set of possible video encodings exists, and indicates how to adapt transmission in response to changes in network conditions in order to maximize user-perceived quality.

Chapter XIII, by Wikstrand, introduces a framework for understanding, developing, and controlling networked services and applications based on a division into three layers. While previous framework models mostly strive to be content- and usage-independent, his proposed model incorporates user considerations. The model is based on three layers: usage, application, and network. By setting targets for and measuring "utility" in appropriate layers of the model, it becomes possible to consider new ways to solve problems and of integrating perceptual requirements in the design of distributed multimedia systems.

Last, but by no means least, the final contribution of our book is **Chapter XIV**, by Germanakos and Mourlas. They investigate the new multi-channel constraints and opportunities emerged by mobile and wireless technologies, as well as the new user demanding requirements that arise. It further examines the relationship between the adaptation and personalization research considerations, and proposes a three-layer architecture for adaptation and personalization of Web-based multimedia content based on the "new" user profile, with visual, emotional, and cognitive processing parameters incorporated.

Summary

Perceptual multimedia design is an exciting but notoriously difficult research area — first, foremost, and last, because it attempts to integrate human requirements and considerations in multimedia communication systems and applications. This book is, to the best of our knowledge, the first book to reunite state-of-the-art research in this chal-

lenging area. It will be useful to professionals and academic researchers working in various areas, such as multimedia telecommunications, Web-based technologies, human-computer interaction, virtual reality, content infrastructure, customization, and personalization. However, this is a first small step. We hope that readers of our book will be imbued with passion and motivation to take this fascinating research area forward, capable of meeting the challenges of the 21st century.

Gheorghita Ghinea and Sherry Y. Chen