

Emerging Multimedia Research and Applications

Mei-Ling Shyu
University of Miami

Shu-Ching Chen
Florida International University

Welcome to this special issue for the 2014 IEEE International Symposium on Multimedia (ISM 2014) held in Taichung, Taiwan, R.O.C. from 10–12 December 2014. This is a collaborative initiative between *IEEE MultiMedia* and ISM 2014 to facilitate the publication of the extended versions of the top symposium papers through a fast-track review and publication process.

ISM 2014

For over a decade, ISM has been an internationally renowned forum for researchers and practitioners to develop solutions and exchange ideas in emerging multimedia research and applications. In January 2015, the authors of the top ISM 2014 papers were invited to submit extended versions of their papers (with at least 30 percent new material) to this special issue. After a rigorous peer-review process, six papers were accepted for this special issue.

In this Issue

The first article, “Optimizing the Perceptual Quality of Real-Time Multimedia Applications,” by Jingxi Xu and Benjamin W. Wah, proposes a systematic and efficient method that uses an offline measured JND (Just-Noticeable Differences) profile to capture human awareness when adjusting control inputs. The method can find good operating points that achieve good perceptual quality for a real-time multimedia system.

The authors attempt to address the fact that perceptual quality is subjective in nature and lacks a well-defined model with respect to control inputs. In addition, it is challenging to design a multimedia system with high perceptual quality, because it requires tuning control inputs using preferences from subjective tests. The proposed method can be generalized to handle multiple quality metrics and control inputs. To demonstrate this, the authors developed an online method that combines multiple independent JND profiles to achieve proper tradeoffs when it comes to the corresponding quality metrics. They also developed an online search method that finds the best control inputs. A voice-over-IP system demonstrates how their proposed method can find the proper control inputs for optimizing perceptual quality.

Next, Yin-Tzu Lin, Chuan-Lung Lee, Jyh-Shing Roger Jang, and Ja-Ling Wu present a

general overview of the challenges of audio music concatenation in “Bridging Music Using Sound-Effect Insertion.” The current advancement of digital technologies has made it possible for people to easily recompose existing media content (such as music pieces) to create new ones (such as musical medleys or mashups for background music). Existing methods find the music segments separately from different music pieces that match each part of the video, and then concatenate them one by one. However, the temporal coherence between the chosen music segments might be lost if the music segments are selected based on the best match with certain parts of the video content.

This motivated the authors to develop solutions to reduce the feelings of abruptness between the chosen music clips for concatenation. For those clips that might be assigned by the users or that just matched the visual contents independently, the authors propose a method to handle and connect those not-so-coherent music clips, which are not suitable for concatenation. To further enhance the transition between clips, they introduce “sound-effect insertion” to bridge such clips and make the transition natural and euphonious. The authors designed and conducted experiments

to evaluate their proposed method, and the results demonstrate that their proposed method is effective in creating smooth and natural transitions via tempo adjustment and sound effect insertion.

The next article features a study that recommends suitable viewpoints from multiview videos recorded from multiple cameras. Although multiview videos show considerable flexibility in enhancing the quality of the viewing experience, the increasing number of cameras makes the selection of suitable viewpoints more challenging. To address this, “Viewpoint Sequence Recommendation Based on Contextual Information for Multiview Video,” by Xueting Wang, Takatsugu Hirayama, and Kenji Mase, introduces a context-dependent automatic viewpoint sequence recommendation method to support multiview video viewing.

The proposed method focuses on context dependency using viewpoint evaluation and transition processes performed by a camera agent that evaluates the view quality and by a producer agent that selects the optimal set of viewpoints based on the scene context and production context. Moreover, their proposed method lets users change their preferences during the recommendation by changing the parameters or the main object of their interest. The authors evaluate the effectiveness of their proposed method by comparing context-dependent and independent video sequences with the selections made by humans. The results show that the method could potentially satisfy both common and personal viewing preferences.

In the next article, “A Markov-Constrained Online Clustering Algorithm,” Cheng-i Wang and Shlomo Dubnov present the Variable Markov Oracle (VMO) data structure for multivariate time series to handle queries in temporal data in terms of two major aspects—namely, query-by-content and gesture following. This can enable interaction between multimedia content and human users. VMO extends both Audio Oracle and Factor Oracle by combining strengths from both of them. Consequently, VMO can identify the repetitive fragments and find sequential similarities between observations, and can be viewed as a combination of online clustering with variable-order Markov constraints.

The authors also proposed a probabilistic interpretation for the query-matching algorithm with VMO. With this probabilistic inter-

pretation, VMO is extended from a data structure to a model for time series. The authors conduct query-by-content experiments using the leave-one-out principle on a gesture database that was recorded using a Kinect 3D camera. The experimental results demonstrate that the VMO approach outperforms the hidden Markov model and dynamic time warping (both considered the baseline approaches for time-series query-retrieval experiments) and is comparable to the state-of-the-art support vector machine with the Covariance of 3D Joints (Cov3DJ) method.

In “Interleaved Time Bases in Hypermedia Synchronization,” Marcio Ferreira Moreno, Romualdo Monteiro de Resende Costa, and Luiz Fernando Gomes Soares discuss the challenges of handling time bases associated with continuous interleaved content during hypermedia presentations to preserve intermedia synchronization defined by causal/constraint relationships. The authors focus on intermedia synchronization defined by the event-based relationships. It is very common to have unsolicited interleaved media content in digital TV systems. For example, in broadcast TV, advertisements are inserted into the programs transmitted sequentially. The authors provide high-level abstractions that hide or minimize the complexity of dealing with interleaved media content as part of the hypermedia applications.

The proposed method has been incorporated in the Nested Context Language (NCL) and in its player, the main component of the Ginga digital TV middleware. It takes advantage of the declarative flexibility and expressiveness from the NCL language and the capability of identifying time bases and their associations with the corresponding media content from Ginga. As a result, it has shown that the users can define the intermedia synchronization without worrying about the artifacts or standard details. In addition, the authors provide a broader spectrum so that the method can be adapted for use in other languages and respective engines.

The last article in this special issue investigates the use of computer-based technologies in dance training (especially for ballet). Paisarn Muneesawang, Naimul Mefraz Khan, Matthew Kyan, Bruce Elder, N. Dong, Guoyu Sun, Haiyan Li, and Ling Zhong propose a novel framework for the real-time capture, assessment, and visualization of balletic movements performed by a student in the 3D cave automatic virtual

environment (CAVE) in their article, “A Machine Intelligence Approach to Virtual Ballet Training.” In this article, the authors train a spherical self-organizing map to quantize movements over the space of typical balletic postures and use the projection of the posture sequences onto this space to form the gesture trajectories. Then, they use these trajectories to generate a collection of movement templates, which form an instructional set.

To evaluate the performance of their proposed framework, the authors use two different histogram models and their sparse codes to see their potential in identifying and describing gesture trajectories. Their proposed framework shows promise, because it is remarkably accurate in restructuring a student’s performance into constituent gesture units that can be analyzed. The authors also discuss future work to further enhance their proposed framework.

We would like to thank all the authors and the reviewers for their efforts on this special issue under a very tight schedule. In addition, we thank Yong Rui (the Editor In Chief) and Wenjun Zeng (the Associate Editor In Chief) of *IEEE MultiMedia* for giving us the

opportunity to publish these high-quality articles in multimedia research, thereby helping us foster interdisciplinary and collaborative research. We hope that you will enjoy reading the extended versions of these top papers from ISM 2014.

MM

Mei-Ling Shyu is a professor in the Department of Electrical and Computer Engineering, University of Miami. She was also one of the ISM 2014 technical program committee co-chairs. Her research interests include multimedia data mining, management and retrieval, big data, and security. Shyu has a PhD in electrical and computer engineering from Purdue University. Contact her at shyu@miami.edu.

Shu-Ching Chen is an Eminent Scholar Chaired Professor in the School of Computing and Information Sciences, Florida International University. He is also an associate editor for *IEEE MultiMedia*. His research interests include content-based image/video retrieval, distributed multimedia database management systems, multimedia data mining, multimedia systems, and disaster information management. Chen has a PhD in electrical and computer engineering from Purdue University. Contact him at chens@cs.fiu.edu.



**Take the
CS Library
wherever
you go!**

 IEEE Computer Society magazines and Transactions are now available to subscribers in the portable ePub format.

Just download the articles from the IEEE Computer Society Digital Library, and you can read them on any device that supports ePub. For more information, including a list of compatible devices, visit

www.computer.org/epub

 **IEEE**  **IEEE Computer Society**