International Journal of Semantic Computing
Vol. 10, No. 1 (2016) 1–4

DOI: 10.1142/S1793351X16020013

World Scientific
www.worldscientific.com

Guest Editors' Introduction

Min Chen
University of Washington Bothell, USA
minchen2@uw.edu

Leonel Sousa
INESC-ID, IST, Universidade de Lisboa, Portugal
las@inesc-id.pt

Yonghong Tian
Peking University, P. R. China
yhtian@pku.edu.cn

Multimedia Computing: Challenges and Opportunities

Recent advances in storage, hardware, information technology, communication and networking have resulted in a large amount of multimedia data. It is estimated that multimedia data including image, video, and audio accounts for 60% of Internet traffic, 70% of mobile phone traffic, and 70% of all available unstructured data [1]. Multimedia data have increasingly become a valuable source for insight and information in applications ranging from business forecasting, healthcare, to science and hi-tech. The sheer volume and heterogeneity of multimedia data and applications present great challenges and unique opportunities to the research on multimedia computing. Generally speaking, research on multimedia computing is concerned with presentation, integration and computation of one or more media (such as text, image, graphics, audio, video, social data, and data collected from various sensors, etc.) using computing techniques [2].

As a premium conference on multimedia, the 2015 IEEE International Symposium on Multimedia (ISM 2015) continues to provide a popular forum for researchers and practitioners to exchange information regarding advances in the state of the art and practice of multimedia computing, as well as to identify the emerging research topics and define the future of multimedia computing. The papers accepted in ISM 2015 provided novel ideas, new results, and state-of-the-art techniques in the field of multimedia computing. Following this successful event, this special issue is to provide yet another forum for the researchers of the top symposium papers to further

present their research results, thus potentially increasing their impact on the community.

Summary of Articles

This special issue is the collaboration between the *International Journal of Semantic Computing* (IJSC) and ISM, which facilitates the publication of the extended versions of the top symposium papers through a fast-track review and publication process. It contains work from experts in the multimedia and knowledge technology research communities, presenting theoretic framework and practical implementations, and identifying challenges and open issues in multimedia computing.

3D motion capture data is a specific type of multimedia data that is mainly used to record movements of humans, animals, or objects over time. This type of data has found widespread utilization in academia and industry, for instance, for entertainment purposes, medical applications, film-making, and video game development. One of the major advantages of 3D motion capture data is the capability of expressing spatio-temporal dynamics with the highest possible accuracy [3]. This property makes 3D motion capture data particularly useful for research into the domain of gestural pattern analysis. In "Efficient Query Processing in 3D Motion Capture Gesture Databases," Beecks *et al.* discuss the issue of efficiently accessing gestural patterns in 3D motion capture databases using spatiotemporal similarity and present a distance-based approach to gestural pattern analysis. This approach models gestural patterns using gesture signatures and investigates a lower bound approximation of the Gesture Matching Distance, which is able to improve the overall efficiency by more than one order of magnitude while maintaining compatible accuracy with the state-of-the-art method. It therefore enables the semi-automatic investigation of large heterogeneous motion capture data archives.

The digitization of music and the distribution of content over the web have greatly increased the number of musical pieces that listeners can access but also makes finding music of interests more difficult. In the article titled "Musical Similarity and Commonness Estimation Based on Probabilistic Generative Models of Musical Elements," Nakano *et al.* presents an approach to estimate musical similarity and commonness, which provides a basis for retrieving musical pieces that closely match a listener's favorites. It is based on a unified framework of probabilistic generative modeling of four musical elements (vocal timbre, musical timbre, rhythm, and chord progression). In their work, the commonness can be estimated by using song-set models, which is easier than estimating the musical similarities of all possible pairs of songs. In the experiments, the estimated musical commonness is evaluated on basis of the Pearson product-moment correlation coefficients, which shows this method can work effectively on song sets of different sizes.

Pedestrian segmentation task is considered a special case of person/human segmentation, which can be used in several applications, including robotics, surveillance

systems, driver assistance models, among others. The article, "Shape-based Pedestrian Segmentation in Still Images," illustrates a new shape-based model for pedestrian segmentation. The proposed model is scale invariant that generates a graph around the detected person. The estimated contour is defined by a path in the graph with maximal cost combined with a selection scheme. The model works well in nontrivial images as compared to the state-of-the-art. In addition, the proposed model is color/texture independent, which can be useful in nighttime applications (when such information could be weakened due to illumination conditions) or when infrared/thermal cameras are employed.

Recent progress in Web Real-Time Communication (WebRTC) promotes multi-apps environment while creating islands of communication apps where users of one website or service cannot easily communicate with those of others. In the article entitled "User Reachability in Islands of WebRTC Communication Apps," K. Singh discusses the architecture and implementation of a multi-platform system to support user reachability in multiple communication services. The work adopts user and endpoint driven reachability policies and cross-app interactions, instead of pair-wise service federation or global location service. Therefore, in the system, users decide how they want to be reached on multiple apps. This architecture separates the user contacts from reachability apps, and has several independent and non-interoperable WebRTC-based apps for two-way and multi-party multimedia communications. Its flexible implementation can be used for enterprise or personal communications, or as a white-labeled app for consumers of a business.

Future Directions

Together, these articles begin to address several important issues in multimedia computing areas. However, the researches in this field toward understanding the semantics of multimedia data sources and building ubiquitous user access are still far from mature. We thus envision several future research directions in this field as summarized below:

- ***Semantic interoperability*:** Multimedia semantic integration considers the descriptions derived from the semantic analysis that is possibly presented in different formats and needs to be integrated before being used collectively. Many approaches have been proposed, such as database schema integration, ontology integration and mapping [4]. However, semantic interoperability among heterogeneous information sources remains a challenging and important research topic.
- ***Social multimedia management and retrieval*:** Increasingly, multimedia collections are associated with networked communities consisting of interconnected groups of users who create, annotate, browse, search, share, view critique and remix collection content. Social multimedia management and retrieval (i.e. retrieval of social multimedia contents uploaded on social networks) thus possess another level of challenge on top of its traditional counterpart. That is how

to best model process user-contributed information and user interactions within the social network [5].

- ***Efficient and joint compression*:** It is estimated that the volume of multimedia big data approximately doubles every two years [6]. Consequently, there is an urgent need for highly efficient compression and coding technologies that can keep pace with the fast growth of the multimedia data. Moreover, it is also highly desirable to develop the joint compression and coding technology for multiple media data, such as video, audio and virtual reality data, which is essential to some attractive ubiquitous multimedia applications in unmanned aerial vehicles, self-driving cars and augmented reality products [2].

Researches on these areas will help spur the development of novel theories and practices in multimedia computing, and may bring new opportunities and become driving forces to the research in the related fields. We hope that readers will find this special issue informative and enjoyable.

Acknowledgments

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 Dr. Phillip Sheu (the EiC) of IJSC for giving us the opportunity to organize this special issue. Yonghong Tian is partially supported by grants from the National Natural Science Foundation of China under contract No. 61390515 and No. 61425025.

References

[1] J. R. Smith, Riding the multimedia big data wave, in *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2013, pp. 1–2.

[2] Y. Tian, M. Chen and L. Sousa, Ubiquitous multimedia: emerging research on multimedia computing, *IEEE MultiMedia* (2016).

[3] C. Beecks *et al.*, Spatiotemporal similarity search in 3D motion capture gesture streams, in *Proceedings of the 14th International Symposium on Spatial and Temporal Databases*, 2015, pp. 355–372.

[4] P. Sheu, Semantic computing, in P. Sheu, H. Yu, C. V. Ramamoorthy, A. K. Joshi and L. A. Zadeh (eds.), *Semantic Computing* (Wiley-IEEE Press, 2010), pp. 1–9.

[5] M. Chen, H. Luo and C. Zhang, Multimedia semantic computing, *International Journal of Multimedia Data Engineering and Management* **2** (2011) pp. 1–4.

[6] J. Gantz and D. Reinsel, The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East, Technical report, Int'l Data Corp. (IDC), 2012; www.emc.com/collateral/analyst-reports/idc-digital-universe-united-states.pdf.