



Special issue on Sentient Multimedia Systems

Loredana Caruccio¹ · Giuseppe Polese¹ · Shi-Kuo Chang²

Published online: 1 January 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

It is a pleasure for us to introduce this special issue of the *Multimedia Tools and Applications* journal, on *Sentient Multimedia Systems*. This issue includes eleven research articles on novel aspects of sentient multimedia systems that significantly benefit from the incorporation/integration of multimedia data (e.g., visual, audio, pen, voice, image, etc.).

Sentient Multimedia Systems are distributed systems capable of actively interacting with the environment by gathering, processing, interpreting, storing, and retrieving multimedia information originated from sensors, robots, actuators, websites, and other information sources. This Special Issue has been inspired by the 26th International Conference on Distributed Multimedia Systems, Visualization, and Visual Languages (DMSVIVA2020) [5], and includes a selection of the best papers presented at the conference, plus some additional papers focusing on sentient multimedia systems. In particular, the papers presented in this Special Issue investigate different aspects of such a theme. All of the articles went through a rigorous procedure of review, involving two expert reviewers for each article. After two/three rounds of review, we selected eleven articles that made relevant contributions to both research and practice.

Other than a paper on the compression of multimedia data, the remaining papers can be classified in two macro-categories, namely *Visualization and Visual Languages*, and *Intelligent Multimedia Systems*. The paper on multimedia compression by Luo et al. concerns the compression of 3D mesh animation sequences [11]. In their work, the authors present a novel spectral clustering-based dynamic reshaping model for the Principal Component Analysis (PCA) elements of spatio-temporal segments, in order to effectively compress 3D mesh sequences. A comparative evaluation with state-of-the-art methods demonstrates the effectiveness of the proposed model, which substantially improves the compression performance of 3D mesh sequences.

The set of papers on *Intelligent Multimedia Systems* include two papers on *Action Recognition* [3, 6], and four papers on *Machine Learning Architectures and Models for Multimedia Applications* [1, 2, 8, 9].

The first paper on Action Recognition by Breve et al. proposes a novel algorithm for mapping human movements into MIDI music, in order to investigate the user's perception associated with the interpretation of sounds [3]. It exploits real-time tracking methodologies, together with a sample-based synthesizer using different types of filters to modulate

✉ Loredana Caruccio
lcaruccio@unisa.it

¹ University of Salerno, Fisciano, Italy

² University of Pittsburgh, Pittsburgh, PA, USA

frequencies. The evaluation performed through a room experience reports significant results on the user perception with respect to the environment they are immersed. Instead, the paper by Coccoli et al. proposes a solution to cope with the problem of illegal dumping prevention in a smart city [6]. It relies on the use of cognitive computing technologies to analyze videos provided from cameras installed in urban areas, aiming to identify illegal trash and bulky waste, triggering an alarm to the municipality.

The category of *Machine Learning Architectures and Models for Multimedia Applications* includes a paper on the automatic diagnosis of cervical cancer [9], one on a machine learning framework to increase safety at work [2], one on the classification of user transportation modalities [1], and finally, one on the prediction of Cyber-Attacks [8]. The paper by Elakkiya R et al. proposes a hybrid deep learning algorithm for cervical localization and precancerous/cancerous lesion detection, in order to provide an end-to-end application for the early diagnosis and prognosis of cervical cancer, which represents one of the curable cancers when it is diagnosed in the early stages [9]. The algorithm accurately spots the cervix without manual annotations and interventions, classifying the cervical cells as normal, precancerous, and cancerous lesions, also identifying the type and stage of cervical cancer. Experimental results demonstrate good accuracy achieved in all training, validation, and testing stages. The paper by Bonifazi et al. proposes a framework using both sentient multimedia systems and machine learning to monitor safety issues in a workplace [2]. It consists of three distinct levels, namely Personal Devices, which are smart objects worn by workers, Area Devices, which are fixed smart objects associated with a specific area, and finally, the Safety Coordination Platform, which monitors the safety of the working environment, and if necessary, it activates suitable alarms. Moreover, the authors illustrate the specialization of the proposed framework on a typical scenario of safety at work, i.e., fall detection, by showing its effectiveness in monitoring the work environment, activating alarms in case of falls, and sending appropriate advice to help workers involved in falls. The paper by Badii et al. tackles the sustainable mobility problem in smart cities [1], providing a solution for delivering useful personalized assistance messages to the user. In particular, they defined a classification system using data originating from sensors embedded in mobile phones, and GIS data (user contextual information), to identify the transportation mean of users. The authors also provide the subset of features related to the GPS, the distance, the accelerometer data, and the temporal windows data that in real operating conditions do not significantly reduce the precision of the classification. Finally, the paper by Cuzzocrea et al. focuses on the detection of Distributed Denial of Service (DDoS) attacks [8], by analyzing and mining sequences of IP addresses, yielding an anomaly detection mining method for predicting the next IP address. It exploits non-linear analysis based on Volterra Kernels and Hammerstein models by considering the sequence of IP addresses as a numerical sequence. Experimental results show that Hammerstein outperforms other models in terms of prediction error.

The set of papers on *Visualization and Visual Languages* includes papers on the visualization of data and of visual languages, other than a paper on visual language parsing, and one on the visualization of the parsing process. The first paper by Kalamaras et al. concerns the visualization of sensitive medical data [10]. The authors present a graph-based approach for the visualization of one or multiple patient cohorts. In particular, a graph, i.e., a network of patients, is constructed from the raw data to encode the similarities among patients, and is positioned on the screen by using force-directed methods. The individual characteristics of each patient are presented by using multivariate glyphs, i.e., small visual representations of multi-dimensional vectors, enabling the determination of the type of patients occupying the different areas of the graph. Three use cases have been used, by

which authors verify that the end results provide insights into the raw data, enabling the detection of patterns and outliers. The second paper by Dang Bui and Ogata [4] proposes a new state picture design of the mutual exclusion protocol invented by Mellor-Crummey and Scott (MCS protocol), aiming to support improved visualizations in a state machine graphical animation (SMGA) tool. The new state picture design of a state machine formalizing the MCS protocol has been assessed based on Gestalt principles, which revealed that the new state picture design and the SMGA tool largely contributed to the successful completion of the formal proof that the MCS protocol enjoys the mutual exclusion property. The third paper by Zou et al. proposes a general parsing algorithm for context-sensitive graph grammars, relying on the RGG formalism [12]. The algorithm embeds two essential strategies, namely context matching and production-set partitioning. The former is utilized to re-examine the found redexes so as to exclude the unnecessary redexes, whereas the latter can be used to precisely choose the relevant productions to narrow down the search space for redexes. Therefore, the two strategies can considerably improve the computational efficiency of a general parsing algorithm, even though the worst-case time complexity is not reduced. Good performances have also been demonstrated through a case study, an experiment, and a qualitative analysis. Finally, the paper by Costagliola et al. focuses on the visualization of the textual and visual language parsing processes. In particular, the authors propose the ParVis visual system for the animated visualization of logged parser trace executions [7]. They aim at developing a system providing a simple interface to parser implementers, applicable to any type of parser, describing two possible uses of ParVis: to visualize the execution of CUP generated parsers, for educational purposes, and to visualize the behavior of a generalized hypergraph-based parser to support researchers during its implementation. A user study involving 14 participants provide encouraging results in the usage of ParVis.

We thank the Editor-in-Chief Borko Furht for his guidance throughout this process and Dr. Margaret Rahmati, administrative assistant, for her help in coordinating this special issue. Finally, we express our thanks to the authors and reviewers, without whose input the issue would not have been possible.

References

1. Badii C, Difino A, Nesi P, Paoli I, Paolucci M (2021) Classification of users' transportation modalities from mobiles in real operating conditions. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-10993-y>
2. Bonifazi G, Corradini E, Ursino D, Virgili L, Anceschi E, De Donato MC (2021) A machine learning based sentient multimedia framework to increase safety at work. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-10984-z>
3. Breve B, Cirillo S, Cuofano M, Desiato D (2021) Enhancing spatial perception through sound: mapping human movements into MIDI. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-11077-7>
4. Bui DD, Ogata K (2021) Better state pictures facilitating state machine characteristic conjecture. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-10992-z>
5. Chang S (ed) (2020) The 26th International DMS Conference on Visualization and Visual Languages, DMSVIVA 2020, KSIR Virtual Conference Center, USA, July 7–8, 2020. KSI ResearchInc. <https://doi.org/10.18293/DMSVIVA2020>
6. Coccoli M, De Francesco V, Fusco A, Maresca P (2021) A cloud-based cognitive computing solution with interoperable applications to counteract illegal dumping in smart cities. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-11238-8>
7. Costagliola G, De Rosa M, Fuccella V, Minas M (2021) Visual exploration of visual parser execution. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-10624-6>

8. Cuzzocrea A, Fadda E, Mumolo E (2021) Cyber-attack detection via non-linear prediction of IP addresses: an innovative big data analytics approach. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-11390-1>
9. Elakkiya R, Teja KSS, Deborah LJ, Bisogni C, Medaglia C (2021) Imaging based cervical cancer diagnostics using small object detection-generative adversarial networks. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-10627-3>
10. Kalamaras I, Glykos K, Megalooikonomou V, Votis K, Tzovaras D (2021) Graph-based visualization of sensitive medical data. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-10990-1>
11. Luo G, Zhao X, Chen Q, Zhu Z, Xian C (2021) Dynamic data reshaping for 3D mesh animation compression. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-10629-1>
12. Zou Y, Zeng X, Zhu Y (2021) A general parsing algorithm with context matching for context-sensitive graph grammars. *Multimed Tools Appl.* <https://doi.org/10.1007/s11042-021-11076-8>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.