A Conceptual Architecture and a Framework for Dealing with Variability in Mulsemedia Systems

Estêvão Bissoli Saleme[1], **Celso Alberto Saibel Santos**[1], **George Ghinea**[2]

[1]PPGI – Universidade Federal do Espírito Santo (UFES), Vitória-ES, Brazil

[2]Dep. of Computer Science – Brunel University London, U.K.

estevaobissoli@gmail.com, saibel@inf.ufes.br, george.ghinea@brunel.ac.uk

Abstract. *The candidate's thesis is within the emerging topic of mulsemedia systems, which encompass audiovisual content associated with multisensory effects, users' quality of experience, and human-computer interaction. It presents a conceptual architecture and a framework that take into account the challenges and requirements for mulsemedia delivery systems identified from the gaps and shortcomings in related work. Furthermore, the work's outcome brings multiple and valuable contributions relying on a complex and solid experimental setup, which have been shared and published in relevant journals and conferences.*

1. Introduction

The increasing interest in digital immersive experiences has drawn the attention of researchers into understanding human perception whilst adding sensory effects to multimedia systems such as Virtual Reality (VR) and Augmented Reality (AR) applications, content players, and games. These so-called mulsemedia—multiple sensorial media—systems [Saleme et al. 2018a, Saleme et al. 2019c] are capable of delivering wind, smell, vibration, among others, along with audiovisual content to enhance users' Quality of Experience (QoE) and to improve immersiveness in applied areas such as entertainment, healthcare, education, culture, and marketing.

To support the researchers' investigation, many standalone software solutions and incipient architectural proposals have been developed to bind these applications to sensory effects devices, such as wind fans, scent emitters, and vibration chairs. These devices, in turn, are constantly evolving, making it difficult to update applications to be compatible with them [Saleme et al. 2019a]. There is little or no interoperability between software and hardware in this realm, hindering reuse in other contexts. Every time a mulsemedia application is needed, new software is built mostly from scratch. This process has proven to be demanding, time-consuming, and costly mainly because it requires researchers and developers alike to gain knowledge about new devices, connectivity, communication protocols, and other particulars. The fact is that building such systems imposes several challenges and requirements (discussed in the thesis) due mainly to their ever-evolving and heterogeneous traits [Saleme et al. 2019d]. As a result, few mulsemedia systems have remained reusable to be applied to different research purposes as opposed to the use of open mulsemedia datasets [Saleme et al. 2019e].

Therefore, the thesis introduces a decoupled conceptual architecture to deal with variability of scenarios in mulsemedia delivery systems, which includes recommendations to cope with the variation of end-user applications and sensory effect devices through the

support and reuse of even unforeseen communication and connectivity protocols, and Sensory Effects Metadata (SEM). The conceptual architecture is grounded on the separation of concerns by dividing a mulsemedia system into five main layers: (1) *Mulsemedia Applications*, which isolates multimedia issues; (2) *Communication Broker*, which provides the reuse of mulsemedia services provided by sensory effect renderers through different communication protocols; (3) *Sensory Effects Processing*, which copes with metadata processing regardless of the SEM standard; (4) *Connectivity Layer*, which allows the reuse of different connectivity protocols for sensory effect devices independently of implementation of commands for devices; and (5) *Sensory Effects Rendering*, to represent devices that deliver sensory effects to end-users.

To evaluate it, an open-source and robust mulsemedia framework was developed. It allows researchers and developers to either undertake mulsemedia experiments or to take it as a reference to follow the conceptual architecture. The framework supports multi-communication and multi-connectivity protocols, multi-standards, and allows the accommodation of new technology relying on its set of architectural and design patterns to be applied successfully. It provides services to mulsemedia applications through a transparent communication broker and is able to work with timeline- and event-based applications. Moreover, the framework presents a set of configurable parameters to customize the solution according to the needs imposed by mulsemedia applications, offers a mechanism to calibrate delay for each component it works with, and provides a debug and flexible simulation scheme [Saleme et al. 2019e]. Then, a performance assessment was carried out on communication protocols for the integration between event-based applications, whereby temporal restrictions play a role, and the framework. Results indicated statistically significant differences in response time providing directions for optimized integrations. Finally, a user QoE subjective evaluation comparing a monolithic mulsemedia system with this framework was undertaken with results suggesting no evinced statistically significant differences in user-perceived QoE between the systems under different aspects. Therefore, it is hoped that the work fosters the area of Multimedia and Human-Computer Interaction (HCI) in the sense that researchers can leverage either the conceptual architecture to design mulsemedia delivery systems or the framework to carry out their experiments.

2. Evaluations

2.1. Case Studies on Variability

The case studies carried out present heterogeneous real-world scenarios of usage in which the conceptual architecture was materialized into a framework to empirically validate it [Saleme et al. 2019e]. These studies include different profiles of usage successfully undertaken to adjust the framework for (a) a mulsemedia environment comprised of video clips enriched with external light, smell, vibration, and wind delivered a system composed of a media player compatible with MPEG-V, DIY devices, smartphones, and a commercial scent device; (b) a scenario of usage whereby a proprietary device is handled by the framework to deliver different intensities of smell; and (c) a 360° VR mulsemedia system composed of an HMD display with a portable scent diffuser attached to it and a wind device to engage users in a 360° environment augmented with smell and airflow. Figure 1 portrays these different environments using the same framework with distinct configurations. The results showed the frame-

work's capability to adapt itself to different scenarios of usage through configuration, without requiring code change. These scenarios turned into a setting for other studies [Saleme et al. 2019d, Covaci et al. 2019, Comsa et al. 2020, Mesfin et al. 2020a].

(a) (b) (c)

Figure 1. Different scenarios using the framework: (a) Video Clip Enriched with External Light, Smell, Vibration, and Wind, (b) Smell-intensive System, and (c) 360° VR Mulsemedia System [Saleme et al. 2019e].

2.2. Performance Assessment

This study made a comparison between different protocols (CoAP, MQTT, UPnP, and WebSocket) that convey information between mulsemedia applications to the implemented framework. Four mock applications were created for each one. The interaction between the two sides is given by two steps: (1) a loop sequence to send SEM from the mulsemedia application to the framework; and (2) a real-time communication to start rendering sensory effects after an occurrence of an event in the mulsemedia former. This communication strategy was an improvement in the face of works that transmit SEM whenever an event happens. SEM is processed once and then invoked through a cue over the network. It does save resources by reducing the overhead of repeated information transmitted. By adding lightweight communication protocols, such as CoAP and MQTT, the study concludes that response time is boosted in relation to UPnP and WebSocket and are recommended for networked event-based mulsemedia systems. There were statistically significant differences in response time between those protocols. Moreover, these time references can be used for the design of mulsemedia systems with temporal restrictions. This study is available in the thesis but preliminary research was published in [Saleme et al. 2017, Saleme et al. 2018b].

2.3. Users' QoE Using the Framework

With the aim to compare user QoE whilst consuming mulsemedia from a monolithic system (SEMP) and a decoupled one (the proposed framework), 40 participants (20 male and 20 female) divided into 2 groups were invited to take part in this subjective experiment. Each group watched 5 different videos in a randomized order on the framework and SEMP without being aware of it. The Philips amBX Premium Kit (wind fans, vibration bar, and external lighting) was employed to render the sensory effects for both systems. The study concludes that no statistically significant difference in user-perceived QoE was found between the two systems under different perspectives such as overall experience watching the videos with sensory effects, audiovisual experience and content, wind effects, vibration effects, external lighting effects. The levels of perceived quality, reality, synchronization, distraction, annoyance, and enjoyment were not remarkably affected when the system was changed. Moreover, an analysis of the interaction between system and video did not evince that there is a statistically significant change in user-perceived QoE when changing the videos for each system. Therefore, the framework's architecture, which allows the possibility of combining different technologies in a customizable fashion, can be leveraged in mulsemedia environments without decreasing levels of user QoE

in comparison with a monolithic system. Parallel user QoE research using the framework has been made available in multiple studies [Jalal et al. 2018a, Covaci et al. 2019, Comsa et al. 2020, Mesfin et al. 2020a, Jalal et al. 2018b, Jalal and Murroni 2019].

3. Research Accomplishments

3.1. Sandwich PhD

The candidate had the opportunity to be an Academic Visitor at Brunel University London, UK, for six months, supported by CAPES (88881.187844/2018-01). Alongside his coworkers, he was able to build software and hardware prototypes[1] for mulsemedia, which was his object of study. From these developments, we could carry out many subjective experiments[2] interviewing several participants at the University of West London, University of Wales Trinity Saint David, Middlesex University, and University of Kent. Researchers at those universities further became the candidate's coauthors in many relevant papers contributing to his formation.

3.2. Publications

- Journals (Impact Factor 2019):
  - [Saleme et al. 2019a] ACM Computing Surveys (**IF 7.990**)
  - [Saleme et al. 2019c] Journal of Universal Computer Science (**IF 0.701**)
  - [Saleme et al. 2019d] IEEE Multimedia (**IF 4.962**)
  - [Saleme et al. 2019e] Multimedia Systems (**IF 1.563**)
  - [de Amorim et al. 2019] Multimedia Tools and Applications (**IF 2.313**)
  - [Comsa et al. 2020] IEEE Multimedia (**IF 4.962**)
  - [Covaci et al. 2020b] IEEE Transactions on Multimedia (**IF 6.051**)
  - [Mesfin et al. 2020b] Multimedia Tools and Applications (**IF 2.313**)
  - [Mesfin et al. 2020a] IEEE Transactions on Multimedia (**IF 6.051**)
- Conferences (h5-index Google Scholar metrics):
  - [Saleme et al. 2017] ACM MMSys'17 (**h5-index=31**)
  - [Saleme et al. 2018a] ACM MEDES'18 (**h5-index=9**)
  - [Saleme et al. 2018b] ACM MMSys'18 (**h5-index=31**)
  - [Covaci et al. 2019] ACM Multimedia'19 (**h5-index=58**)
  - [Saleme et al. 2019b] SBC IHC'19 (**h5-index=9**)

In addition to journals and conferences, a summary of the candidate's thesis was published at ACM SIGMM (Special Interest Group in Multimedia) Records, Volume 11, Issue 4, December 2019 (ISSN 1947-4598)[3]. Furthermore, the papers [Saleme et al. 2020] and [Covaci et al. 2020a] have been respectively submitted to the International Journal of Human-Computer Studies (IF 3.163) and IEEE TVCG (IF 4.558).

3.3. Student Travel Awards

- SIGMM Student Travel Award for ACM MMSys'17[4], Taipei, Taiwan - 500.00 USD, ACM SIGMM.
- SIGMM Student Travel Award for ACM Multimedia'19[5], Nice, France - 1500.00 USD, ACM SIGMM. The candidate was also a session volunteer[6].

[1] Mulsemedia Infrastructure summary available at https://bit.ly/32NdIw9
[2] QoE in Mulsemedia Systems summary at available https://bit.ly/2vBZr9l
[3] ACM SIGMM Records PhD summary available at https://bit.ly/39jYj8Z
[4] ACM MMSys'17 presentation photo available at https://bit.ly/2wmJ3tE
[5] ACM Multimedia'19 presentation photo available at https://bit.ly/32K84Lo
[6] ACM Multimedia'19 session volunteer report available at https://bit.ly/2TkYKus

4. Final Remarks and Impact

The thesis dealt with a hot topic: a conceptual architecture and a framework to support how the research community can leverage reusable mulsemedia systems in different contexts. The originality of the work relies on the multisensory excitation in addition to the usual multimedia dimension (image and sound). We highlight the nature of the work that requires not only software engineering skills, but also psychology background in addition to other side disciplines (statistics, signal processing, etc.). The work tackles and validates the problem throughout the stack, from the system to the user perception. It addresses a well-known and accepted problem and can be an impactful work. For instance, independent researchers [Jalal et al. 2018a, Jalal et al. 2018b, Jalal and Murroni 2019] have taken advantage of our work to carry out their research using our set of tools.

The manuscript is composed of six chapters developed on 147 pages. It is supplemented with a meaningful and detailed annex to secure reproducibility of the experiments conducted along the research work. The research contributions rely on a complex and solid experimental setup. The work results are discussed around research questions along with the development of most advanced multimedia technology. The references try to represent the interdisciplinary dimension of the conducted research (advanced techniques in software engineering and systems, HCI, and psychology), which is inherent with the nature of the emerging sciences of Mulsemedia Systems and Quality of Experience.

From the thesis, the dissemination efforts are substantial with relevant direct and indirect outputs. These contributions are important for the community as they are addressing hot but poorly addressed topics. They have the potential to constitute a milestone for other researchers by fostering the area of HCI and mulsemedia in the sense that multisensory research can be leveraged to pave new ways and boost several application areas, such as entertainment, healthcare, education, culture, and marketing. Furthermore, the work was exposed to different supervision culture. This has not only enriched the candidate skills, but also made him find his role/place in such context successfully.

References

Comsa, I.-S., Saleme, E. B., Covaci, A., Mesfin, G., Trestian, R., Santos, C. A. S., and Ghinea, G. (2020). Do I Smell Coffee? The Tale of a 360° Mulsemedia Experience. *IEEE MultiMedia*, 27(1):27–36.

Covaci, A., Saleme, E. B., Mesfin, G., Comsa, I.-S., Trestian, R., Santos, C. A. S., and Ghinea, G. (2020a). Multisensory 360° Videos Mediate Presence. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–9. Submitted.

Covaci, A., Saleme, E. B., Mesfin, G. A., Hussain, N., Kani-Zabihi, E., and Ghinea, G. (2020b). How do we experience crossmodal correspondent mulsemedia content? *IEEE Transactions on Multimedia*, 22(5):1249–1258.

Covaci, A., Trestian, R., Saleme, E. B., Comsa, I.-S., Mesfin, G., Santos, C. A. S., and Ghinea, G. (2019). 360° Mulsemedia: A Way to Improve Subjective QoE in 360° Videos. In *Proc. of the 27th ACM Int. Conf. on Multimedia*, pages 2378–2386. ACM.

de Amorim, M. N., Saleme, E. B., de Assis Neto, F. R., Santos, C. A. S., and Ghinea, G. (2019). Crowdsourcing authoring of sensory effects on videos. *Multimedia Tools and Applications*, 78(14):19201–19227.

Jalal, L., Anedda, M., Popescu, V., and Murroni, M. (2018a). Qoe assessment for broadcasting multi sensorial media in smart home scenario. In *2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–5.

Jalal, L., Anedda, M., Popescu, V., and Murroni, M. (2018b). QoE Assessment for IoT-Based Multi Sensorial Media Broadcasting. *IEEE Trans. Broadcast.*, 64(2):552–560.

Jalal, L. and Murroni, M. (2019). The impact of multi-sensorial media in smart home scenario on user experience and emotions. In *2019 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, pages 1–6.

Mesfin, G., Saleme, E. B., Ademoye, O., Kani-Zabihi, E., Santos, C. A. S., and Ghinea, G. (2020a). Less is (Just as Good as) More - an Investigation of Olfactory Intensity and Hedonic Valence in Mulsemedia QoE using Heart Rate and Eye Tracking. *IEEE Transactions on Multimedia*.

Mesfin, G. A., Hussain, N., Kani-Zabihi, E., Covaci, A., Saleme, E. B., and Ghinea, G. (2020b). QoE of Cross-modally Mapped Mulsemedia: an Assessment Using Eye Gaze and Heart Rate. *Multimedia Tools and Applications*.

Saleme, E. B., Celestrini, J. R., and Santos, C. A. S. (2017). Time Evaluation for the Integration of a Gestural Interactive Application with a Distributed Mulsemedia Platform. In *Proc. of the 8th ACM on Multimedia Systems Conference*, pages 308–314. ACM.

Saleme, E. B., Covaci, A., Assres, G., Comsa, I.-S., Trestian, R., Santos, C. A. S., and Ghinea, G. (2020). The Influence of Human Factors on 360° Mulsemedia QoE. *International Journal of Human-Computer Studies*, pages 1–39. Submitted.

Saleme, E. B., Covaci, A., Mesfin, G., Santos, C. A. S., and Ghinea, G. (2019a). Mulsemedia DIY: A Survey of Devices and a Tutorial for Building Your Own Mulsemedia Environment. *ACM Computing Surveys*, 52(3):58:1–58:29.

Saleme, E. B., Covaci, A., Mesfin, G., Santos, C. A. S., and Ghinea, G. (2019b). Sumarização de dispositivos de efeitos multissensoriais para interações humano-computador. In *Extended Proceedings of IHC '19: XVIII Brazilian Symposium on Human Factors in Computing Systems*, IHC '19. SBC.

Saleme, E. B., Santos, C. A. S., Falbo, R. A., Ghinea, G., and Andres, F. (2018a). Towards a reference ontology on mulsemedia systems. In *Proceedings of the 10th International Conference on Management of Digital EcoSystems*, MEDES '18, pages 23–30. ACM.

Saleme, E. B., Santos, C. A. S., Falbo, R. A., Ghinea, G., and Andres, F. (2019c). MulseOnto: a Reference Ontology to Support the Design of Mulsemedia Systems. *Journal of Universal Computer Science*, 25(13).

Saleme, E. B., Santos, C. A. S., and Ghinea, G. (2018b). Improving response time interval in networked event-based mulsemedia systems. In *Proceedings of the 9th ACM Multimedia Systems Conference*, MMSys '18, pages 216–224. ACM.

Saleme, E. B., Santos, C. A. S., and Ghinea, G. (2019d). Coping With the Challenges of Delivering Multiple Sensorial Media. *IEEE MultiMedia*, 26(2):66–75.

Saleme, E. B., Santos, C. A. S., and Ghinea, G. (2019e). A mulsemedia framework for delivering sensory effects to heterogeneous systems. *Multimedia Systems*, 25(4):421–447.