DOI 10.31891/2307-5732-2023-327-5-121-125
UDC 004.6

YEVGENIYA SULEMA
Igor Sikorsky Kyiv Polytechnic Institute
https://orcid.org/0000-0001-7871-9806
email: sulema@pzks.fpm.kpi.ua
DMYTRO RVACH
Igor Sikorsky Kyiv Polytechnic Institute
https://orcid.org/0000-0002-1461-1086
email: rvach@pzks.fpm.kpi.ua

HIGH-LEVEL ARCHITECTURE OF MULSEMEDIA SOFTWARE SYSTEM

This paper delves deep into the creation of a unified, high-level architecture for an emerging category of software known as mulsemedia software. Mulsemedia, a term that connects 'multimedia' and 'sensory', refers to an extended version of multimedia. Unlike traditional multimedia which primarily focuses on visual and auditory data, mulsemedia encompasses a broader spectrum of temporal multimodal data. A core component of this concept is the mulsemedia object, essentially a physical entity monitored by a computer system. This is done through the integration of various sensors, aimed at capturing a comprehensive digital representation of the object. The goal is to have this digital description be so intricate and detailed that it mirrors the human sensory experience when perceiving the object. Mulsemedia software, therefore, can be understood as an application that specializes in handling this rich, temporal multimodal data, providing a holistic view of the mulsemedia object. The software landscape already features categories that align with the mulsemedia software system, notably the digital twin platforms and metaverse applications. However, mulsemedia stands out as a relatively new concept, and thus there's a pressing need to establish novel methodologies to develop software that can efficiently and effectively interact with diverse modalities of mulsemedia data. The primary focus of this research is centred on conceptualizing and formulating a high-level architecture tailored for mulsemedia software. This architecture, while robust in its design principles, boasts versatility, making it seamlessly adaptable across a myriad of application scenarios and use cases. Such a foundational architecture will pave the way for future advancements in mulsemedia software development, fostering innovation in this burgeoning field.

Keywords: mulsemedia, software architecture, metaverse, digital twins.

ЄВГЕНІЯ СУЛЕМА
Київський політехнічний інститут імені Ігоря Сікорського
ORCID ID: 0000-0001-7871-9806
email: sulema@pzks.fpm.kpi.ua
ДМИТРО РВАЧ
Київський політехнічний інститут імені Ігоря Сікорського
ORCID ID: 0000-0002-1461-1086
email: rvach@pzks.fpm.kpi.ua

ВИСОКОРІВНЕВА АРХІТЕКТУРА МУЛЬСЕМЕДІЙНОЇ ПРОГРАМНОЇ СИСТЕМИ

Ця стаття присвячена розробленню уніфікованої високорівневої архітектури нового класу програмного забезпечення – мульсемедійного програмного забезпечення. Цей клас програмного забезпечення працює з даними, які визначають розширене мультимедіа, також відоме як мульсемедіа. Концепція мульсемедіа передбачає роботу з темпоральними мультимодальними даними, що описують мульсемедійний об'єкт, який є фізичним об'єктом, що контролюється комп'ютерною системою за допомогою набору сенсорів для формування мультимодального цифрового опису, який визначає об'єкт настільки комплексно, як його може сприйняти людина через свої органи чуття. Таким чином, мульсемедійне програмне забезпечення – це програмна система, яка працює з темпоральними мультимодальними даними, що визначають мульсемедійний об'єкт. Спорідненими до мульсемедійних програмних систем класами програмного забезпечення є платформи цифрових двійників та застосунки метаверсу. Мульсемедіа є відносно новою концепцією, і вона вимагає створення нових підходів до розроблення програмних застосунків, які працюють з мульсемедійними даними різних модальностей. Дослідження, представлене у цій статті, зосереджено на розробленні архітектури високого рівня, яку можна легко застосувати до різних випадків використання.

Ключові слова: мульсемедіа, архітектура програмного забезпечення, метаверс, цифрові двійники.

Introduction

Recently, there has been an expansion of application of software systems in fields that were traditionally considered to belong exclusively to human mental activity. A clear example is the medical industry, where the decision is always made by a human – a medical doctor. However, the use of artificial intelligence for pattern recognition and primary diagnostics demonstrates that software solutions can be effectively applied in areas of human activity where they were not used before. This tendency also has other manifestations. So, for example, the creation of the Metaverse concept is actually the beginning of deep digitalization, which involves the transfer of part of human activity to the digital world [1, 2]. This will require new approaches to the presentation of information about the physical world for the reproduction of real objects in a virtual environment. At the same time, there will be a need for the reverse, that is, the reproduction of the properties of virtual objects in the real environment so that the user can feel these properties with the help of his senses. This can be done with the help of mulsemedia technology, that is an extension of multimedia technology [3, 4].

Mulsemedia enables digitalization of information that humans receive through their senses. If audiovisual information is always in use for interaction between a computer system and a user, other types of sensorial information such as olfactory, tactile, gustatory, kinesthetic and theroceptic are new forms of information to be exchanged between

a computer system and a user.

Application of the mulsemedia concept requires the use of specialized hardware which allows to capture multisensory information. Nowadays, there is a wide range of such a hardware either available on the market or still under the development on research labs [5-14]. However, mulsemedia software still requires further development. This research is aimed at filling the gap between a large number of possible applications, a wide range of available hardware and limited examples of mulsemedia software.

Literature Review

In [3], the authors formulate the main challenges for delivering multisensory effects to heterogeneous systems and suggest an interoperable mulsemedia framework to address the indicated challenges. For this purpose, they propose an open distributed mulsemedia system where architectural and design patterns are used for meeting the requirements on communication, connectivity, and sensory effects metadata standard. The approach takes into consideration availability of corresponding mulsemedia devices in the user's environment.

The paper [4] is devoted to the development and evaluation of a mulsemedia software-hardware system for high education in medicine. The system uses virtual reality elements. The authors use this system to study the impact of using virtual reality simulators on students' motivation to learn.

In [10], the authors provide an overview of recently developed wearable haptic devices. These devices classified by their ability to operate with different modalities which characterize an object's properties such as texture, weight, size, shape, etc. The research shows that a combination of tactile and kinaesthetic feedback, in particular, vibrotactile and tendon-driven mechanisms, enables more extensive and detailed exploration of the object's properties.

In [13], the authors present the technology of taste information processing. Taste sensors, which are also named electronic tongues or bioelectronic tongues, are devices designed to be applied in food industry (for evaluation of food and beverages) and healthcare (for medical diagnostics).

The paper [14] shows the design and application of the electronic tongue. This device consists of sensors, data acquisition system, multiplexers, and a data analysis system. The processing of data received from the device includes data acquisition, scaling, multivariate data analysis, classification and cross validation.

The paper [15] presents a generative approach to the development of mulsemedia applications. This approach is demonstrated on the use case of education. For this purpose, the authors modelled a family of multisensory education applications that explore the education scenario. They also focused on extension of the MML metamodel to support the modelling of sensory effects.

In [16], the authors provide the general information on the mulsemedia systems, which encompass audiovisual content associated with multisensory effects, users' quality of experience, and human-computer interaction. The framework enables different scenarios including: (1) video clip enriched with external light, smell, vibration, and wind; (2) smell-intensive system, and (3) 360° VR mulsemedia system. The implementation of the mulsemedia system according to these scenarios can facilitate further wider use of such systems in different areas of human activity.

The analysis of the existing literature sources allows us to conclude that the further research related to the concept of mulsemedia must be focused on the designing a reference architecture for the software development.

Thus, the *research objective* is to design a unified high-level architecture for the development of mulsemedia software applications for different areas of human activity. This architecture must enable the possibility of using different devices that operate with multisensory information.

Presenting Main Material

To present the designed unified high-level architecture, we need to formulate two fundamental notions.

Definition 1. *Mulsemedia object* is a physical object which is supervised by a computer system using a set of sensors in order to form multimodal digital description that defines the object as complex as if it can be perceived by humans through their organs of sensing.

Definition 2. *Mulsemedia software* is an application which operates with temporal multimodal data defining a mulsemedia object.

The developed architecture is presented in Fig. 1.

This architecture assumes that a mulsemedia object is described through several data modalities. For example, if the mulsemedia object is a gas-distribution unit, it can be represented by three data modalities: (1) it can be supervised visually by a video camera; (2) it can be under control through a haptic interface; (3) its current state can be registered using a smell detector. Using sensors and actuators for these modalities, both monitoring and operation of the gas-distribution unit can be provided in a remote mode on a qualitatively new level by developing and employing appropriate mulsemedia software. In case, if a mulsemedia object is a technical object, mulsemedia software can be implemented as a digital twin platform. In case, if a mulsemedia object is expected to be implemented as a virtual object, mulsemedia software can be a part of the metaverse platform.

According to the designed architecture, a mulsemedia software system includes the following components.

The core of the system is the Mulsemedia Object's Data Processing Logic module. It represents the essential part of temporal multimodal data processing for the given task for the operation with the mulsemedia object. This module is expected to implement the operations on TJSON-objects.

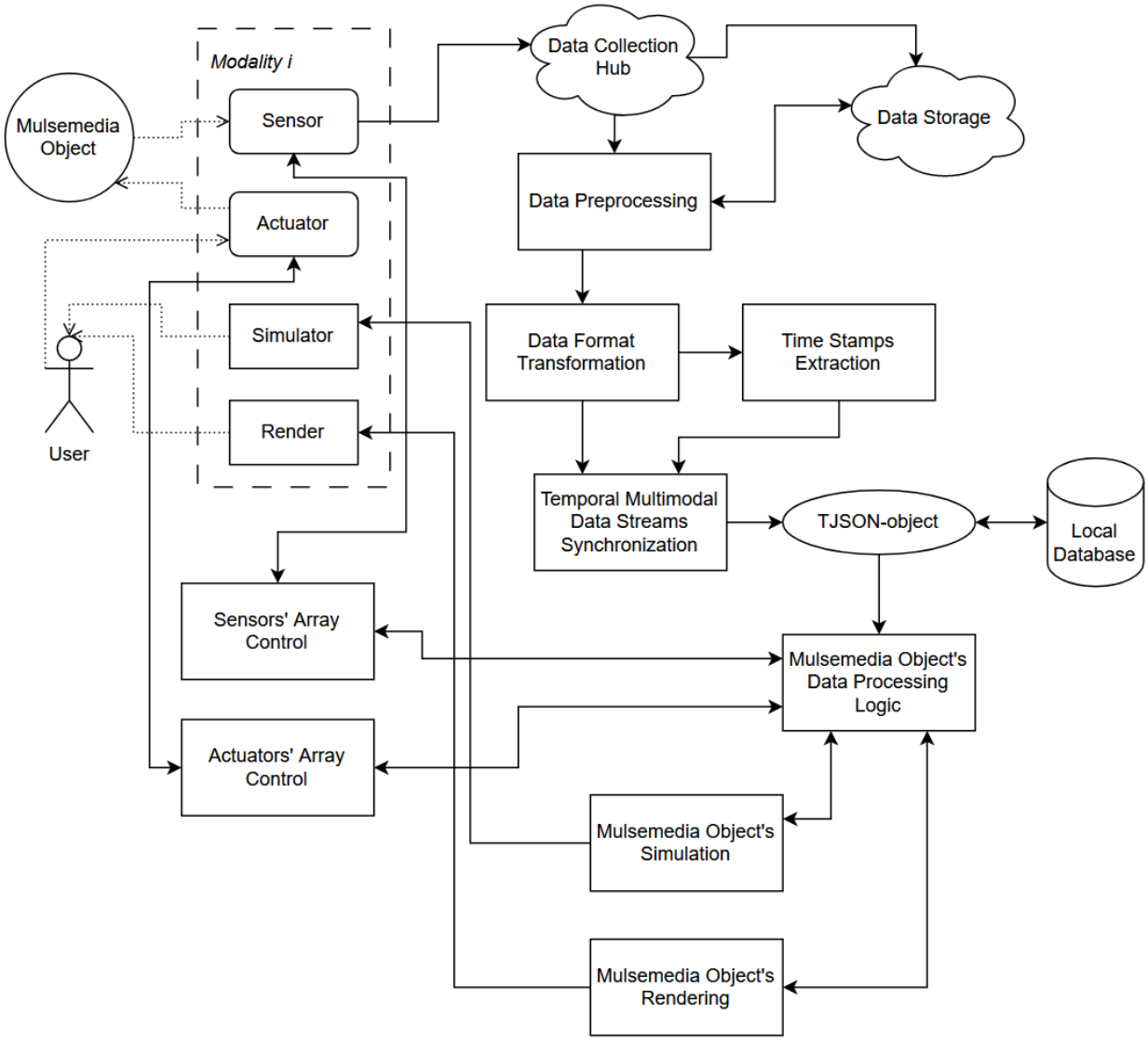

Fig. 1. Mulsemedia Software High-level Architecture

A TJSON-object is an extension of a JSON-object. JSON (JavaScript Object Notation) is widely used as a format for storing and transporting data as it is easy to understand and use [17]. Thus, a JSON-object is a description of a certain object, for example, the object of observation based on JSON format. However, if the object of observation is a mulsemedia object, its description requires temporal information of different modalities. That is why JSON format needs to be extended for representing of temporal multimodal data describing the same object that is mulsemedia object. This can be done by introducing a special version of JSON format that includes time stamps. This advanced JSON format enables defining a TJSON-object that is a Timeline JSON-object [18].

In the designed architecture, a TJSON-object is produced by the module of Temporal Multimodal Data Streams Synchronization, and it is stored in the Local Database.

The module of Temporal Multimodal Data Streams Synchronization implements the method of multimodal mulsemedia data consolidation which enables the combination and synchronization of multimedia data using the principles of multithreading [18]. The module receives both data of different modalities and their time stamps as input data. In their turn, data of a certain modality and its time stamps are the outputs of two other modules: Data Format Transformation module and Time Stamps Extraction module. The purpose of the Data Format Transformation module is to convert data of certain modality from its original format to a format implemented as the basic format in a particular mulsemedia software system. Criteria for the selection of a proper format depend on several aspects including the task for the mulsemedia object investigation and hardware used for sensing this object. The challenge here is that currently mulsemedia hardware is not standardized when it comes to such specific modalities as haptics, olfaction and gustatory. It means that there is not a predefined way of data representation format to be used in the mulsemedia software system. At the same time, it gives flexibility in operating with multimodal data to the developers. The Time Stamps Extraction module is aimed at extraction of temporal information from the multimodal data presented either as a stream or as a file in a certain format. It receives the data of a specific modality from the Data Format Transformation module.

Each modality of information defining the mulsemedia object is supported by a set of devices: sensors for collecting data of this modality and actuators for providing the feedback. The control of these devices is carried out

through the Sensors' Array Control module and the Actuators' Array Control module. These modules exchange instructions with the Mulsemedia Object's Data Processing Logic module.

Data streams of all modalities are transferred from the arrays of sensors for different modalities registration. These streams are collected by the Data Collection Hub which is a cloud component of the mulsemedia software system. It directs data to both the cloud Data Storage and the Data Preprocessing module. The purpose of the Data Preprocessing module is to enhance the data quality before the data stream is processed by the Data Format Transformation module.

The interaction of the user with the mulsemedia object or its digital representation (digital twin) is provided by using the set of simulators, renders, and actuators operating with a certain modality.

A simulator either can be implemented as a component of the mulsemedia software system or it can be an external application compatible with this system. The use of simulations is supported by the Mulsemedia Object's Simulation module which exchange instructions with the Mulsemedia Object's Data Processing Logic module.

A render is both a device and its driver which enable reproduction of information for certain modality. Interconnection between the render's driver and the Mulsemedia Object's Data Processing Logic module is fulfilled through the Mulsemedia Object's Rendering module.

Thus, the designed architecture supports all aspects of processing the temporal multimodal data describing the mulsemedia object. It can be implemented by using either general-purpose languages or the domain-specific programming language ASAMPL [19, 20].

Conclusions

The presented high-level architecture of a mulsemedia software system is aimed at facilitating the development of mulsemedia-based software solutions for a wide range of applications. The designed reference architecture can be adapted to different sets of temporal multimodal data describing a mulsemedia object. The essential feature of this architecture is that it assumes: (1) extraction of hidden time stamps from data presented in different formats; (2) composing data of all modalities as one TJSON-object; (3) possibility of various scenarios for operating with information about a mulsemedia object.

Further research can be focused on the development of specific software libraries which implements procedures of working with sensors, actuators, renders, and simulators for different modalities representing a mulsemedia object.

References

1. Mystakidis, S. Metaverse. Encyclopedia 2022, 2, 486–497. https://doi.org/10.3390/encyclopedia2010031

2. Ian F. Akyildiz, Hongzhi Guo, Rui Dai, Wolfgang Gerstacker. Mulsemedia Communication Research Challenges for Metaverse in 6G Wireless Systems. Multimedia. arXiv:2306.16359. https://doi.org/10.48550/arXiv.2306.16359

3. Saleme, E.B., Santos, C.A.S. & Ghinea, G. A mulsemedia framework for delivering sensory effects to heterogeneous systems. Multimedia Systems 25, 421–447 (2019). https://doi.org/10.1007/s00530-019-00618-8

4. Kumar, A.; Srinivasan, B.; Saudagar, A.K.J.; AlTameem, A.; Alkhathami, M.; Alsamani, B.; Khan, M.B.; Ahmed, Z.H.; Kumar, A.; Singh, K.U. Next-Gen Mulsemedia: Virtual Reality Haptic Simulator's Impact on Medical Practitioner for Higher Education Institutions. Electronics 2023, 12, 356. https://doi.org/10.3390/electronics12020356

5. The Best Surround-Sound Systems 2023. https://www.nytimes.com/wirecutter/reviews/best-surround-sound-system/

6. ASUS created new 3D display technology for ProArt Studiobook 16 3D OLED and Vivobook Pro 16X 3D OLED laptops. https://mezha.media/en/2023/01/06/asus-created-new-3d-display-technology-for-proart-studiobook-16-3d-oled-and-vivobook-pro-16x-3d-oled-laptops/

7. The 13 Best 3D Cameras 2023. https://www.3dsourced.com/rankings/best-3d-camera/

8. The Best 3D Scanners in 2023. https://all3dp.com/1/best-3d-scanner-diy-handheld-app-software/

9. The Best 3D Printers in 2023. https://all3dp.com/1/best-3d-printer-reviews-top-3d-printers-home-3-d-printer-3d/

10. Van Wegen, M.; Herder, J.L.; Adelsberger, R.; Pastore-Wapp, M.; van Wegen, E.E.H.; Bohlhalter, S.; Nef, T.; Krack, P.; Vanbellingen, T. An Overview of Wearable Haptic Technologies and Their Performance in Virtual Object Exploration. Sensors 2023, 23, 1563. https://doi.org/10.3390/s23031563

11. Haptic Feedback Device and Method. Patent WO2023178428. World Intellectual Property Organization. https://patentscope.wipo.int/search/ja/detail.jsf;jsessionid=3007B52868220BEE47DEF98AD2B24067.wapp2nB?docId=WO2023178428&_gid=202339

12. Recent Advancements in Olfaction and Electronic Nose. A special issue of Sensors journal. 2023. https://www.mdpi.com/journal/sensors/special_issues/6G27QVZF8A

13. Jin Wang, Kenji Sakai, Toshihiko Kiwa. All-in-one terahertz taste sensor: integrated electronic and bioelectronic tongues. Sensors & Diagnostics. 2023, 2, 620. https://pubs.rsc.org/en/content/articlepdf/2023/sd/d3sd00038a

14. Leon-Medina, J.X.; Anaya, M.; Tibaduiza, D.A. New Electronic Tongue Sensor Array System for Accurate Liquor Beverage Classification. Sensors 2023, 23, 6178. https://doi.org/10.3390/s23136178

15. Marcelo Fernandes de Sousa, Raoni Kulesza, Carlos André Guimarães Ferraz, and Matheus Lima. 2019. A Generative Software Development approach for mulsemedia application domain. In Brazilian Symposium on Multimedia and the Web (WebMedia '19), October 29-November 1, 2019, Rio de Janeiro, Brazil. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3323503.3360290

16. Estêvão Bissoli Saleme, Celso A. S. Santos, Gheorghita Ghinea. A Conceptual Architecture and a Framework for Dealing with Variability in Mulsemedia Systems. ACM SIG Multimedia Records. Volume 11, Issue 4, 2019. Article No. 9. https://doi.org/10.1145/3530839.3530848

17. Tom Marrs. JSON at Work: Practical Data Integration for the Web. O'Reilly Media, 2017. 374 p.

18. Rvach Dmytro, Yevgeniya Sulema. Consolidation Method for Multimodal Mulsemedia Data. System Technologies, 2022. Vol. 6.

19. Yevgeniya Sulema. ASAMPL: Programming Language for Mulsemedia Data Processing Based on Algebraic System of Aggregates. Advances in Intelligent Systems and Computing, Vol. 725, Springer, 2018, pp. 431-442.

20. Ivan Dychka, Yevgeniya Sulema, Dmytro Rvach, Liubov Drozdenko. Programming Language ASAMPL 2.0 for Mulsemedia Applications Development. Advances in Computer Science for Engineering and Education. Lecture Notes on Data Engineering and Communications Technol., Vol. 134, pp. 107-116. Springer, 2022.