IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 9 September 2022, accepted 17 October 2022, date of publication 28 October 2022, date of current version 8 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3217799

RESEARCH ARTICLE

Immersive Authoring of 360 Degree Interactive Applications

FLÁVIO MIRANDA DE FARIAS[1], DOUGLAS PAULO DE MATTOS[1], GHEORGHITA GHINEA[2], (Member, IEEE), AND DÉBORA CHRISTINA MUCHALUAT-SAADE[1]

[1]MídiaCom Laboratory, Fluminense Federal University, Niterói, Rio de Janeiro 24210-240, Brazil
[2]Department of Computer Science, Brunel University London, UB8 3PH London, U.K.

Corresponding author: Flávio Miranda De Farias (flavio.farias@midiacom.uff.br)

This work was supported in part by the Brazilian National Council for Scientific and Technological Development (CNPq), Carlos Chagas Filho Foundation for Research Support in the State of Rio de Janeiro (FAPERJ), Coordination for the Improvement of Higher Education Personnel (CAPES), and Program for Institutional Internationalization (CAPES PrInt).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Research Ethics Committee of Fluminense Federal University.

ABSTRACT Although there are proposals in the literature for authoring mulsemedia (*mul*tiple *se*nsorial *media*) applications with 2D content, there are no suitable solutions when it comes to 360° content. Moreover, little consensus on 360° mulsemedia authoring methodology exists. Aiming at filling this gap, we propose the concept of immersive authoring of 360° multisensory applications. Our proposal comprises an immersive 360° authoring environment to bring the author closer to the final user presentation environment. We implemented our proposal in AMUSEVR, a virtual-reality (VR) environment for authoring 360° mulsemedia applications. We see it as an alternative or a possible complement to available 2D mulsemedia authoring tools. AMUSEVR provides creation and editing of interactive multiple sensorial media scenes by directly arranging objects in a 3D space using VR technology. Also, the tool allows users to run their applications through AMUSEVR viewer mode. We used the Goal Question Metric (GQM) approach to plan our tests and a group of users evaluated the tool with the SUS and UEQ questionnaires, obtaining a SUS score of 82.25 and an excellent UEQ benchmark, which are very promising results.

INDEX TERMS 360 authoring, event-based relations, interactive scenarios, mulsemedia, AMUSEVR.

I. INTRODUCTION

The evolution of multimedia applications has brought new kinds of content, such as 360° image and video, which provide the consumer an enhanced sense of immersion within the media content. Therefore, they bring the promise of fresh narratives in communications, education and marketing, among other fields [23]. It should be noted that there are a few limitations to this technology in that the freedom of orientation found in the VR (Virtual Reality) environment introduces unpredictability in the user experience. The audience is at risk of losing essential elements or details, due to the methods of interaction with content, as well as the lack of experience of content producers and the eventual audience [23].

The associate editor coordinating the review of this manuscript and approving it for publication was Usama Mir.

Considering 360° media, one limitation is the fact that the user is usually fixed at the center of the application. This can work for cinema content (movies and alike), where one sits in a fixed location, but does not always work for interactive museums, for example. To meet the demand for freedom of movement, VR has been used, as it offers a mix of different kinds of content along 6 axes of movement. Furthermore, multiple media objects can be located spatially in the presentation and, using controller devices, users can interact with content, allowing them to evolve from passive spectators to active participants [7], [11].

For VR to proliferate as a technology that offers an enhanced user experience [23], we need to provide tools for producing that kind of content to authors who are not programming experts. However, available tools still use authoring techniques inherited from traditional technologies aimed at 2D media. Some use the approach focused on building presentations based on timelines [4], [5], others on relationships among media objects [10]. Unsurprisingly, providing an authoring environment that allows manipulating 360° media content in 3D space is still a challenge.

Based on the videogame development industry, which is a current environment for VR content creation [23], the work of [18] recommends the use of VR to produce and place objects within the content in order to improve production and provide greater fidelity in the final product. For this particular type of technology, game engines, such as Unity,[1] offer good production support for VR content, as well as support for standard media, such as image, video and audio, user interaction and even integration with sensory effects [20], [21].

Indeed, integration with sensory effects becomes imperative, given that most human communication is non-verbal, and we frequently use other senses, besides sight and hearing, to understand the world around us, to socialize, and to entertain ourselves [8]. Additionally, we can better interpret content offered by an application combining different senses [23]. We have proposals for authoring sensory effects and interactive events synchronized with traditional multimedia content, such as MultiSEM [10] and STEVE [1]. However, there are fewer applications and methodologies for this purpose targeting VR and 360° content.

In this paper, we propose and investigate the concept of immersive authoring of 360° interactive applications. We aim at answering the following research question: *"What is the user experience of creating 360° interactive multisensory applications using an immersive authoring environment?"* To investigate and answer this research question, we propose an immersive 360° authoring environment to bring the author closer to the final user presentation environment. The immediate goal is to establish a working environment and interface. However, in a near future, we will integrate multisensory rendering modules. Therefore, detailed descriptions of sensory effects are not the focus of this article.

We implemented our proposal in a VR environment named AMUSEVR (*Authoring 360° MUltimedia and Sensory Effects in VR*). The main contribution of this work is that it provides immersion for the application author and not just for the final users. Our goal is to facilitate the production of 360° interactive applications by placing the author in the same environment that the end user will experience. This work is an extension of [19], in addition, we go deeper discussing the XML language we developed as an interchange format for AMUSEVR. We also present new AMUSEVR user experiments detailing the methodology used and results obtained when evaluating AMUSEVR with SUS (System Usability Scale) and UEQ (User Experience Questionnaire). To the best of our knowledge, we are not aware of a similar solution that provides authoring in real time in a VR environment, supports interactive media objects and a flexible XML language as an interchange format.

[1]https://unity.com/

The remainder of this paper is organized as follows. Section II discusses related work. Section III presents our proposal for immersive authoring for 360° interactive applications. Section IV presents AMUSEVR, the implementation of our proposal in a VR platform. Section IV-B briefly presents MultiSEL, the XML language we propose as interchange format. Section V describes evaluation experiments based on the Goal Question Metric (GQM) approach. Conclusions and future work are given in Section VI.

II. RELATED WORK

In regards to related work, we defined, as our search goal, studies about authoring solutions, which used authoring tools supporting mulsemedia (multimedia and multisensory) content, 360° content, user interaction, all in a fully immersive environment. As a totally equivalent solution was not found, our search filters were set for post-2017 results, using the following keywords: "interactive authoring", "360 authoring", "immersive authoring", "mulsemedia authoring", "sensory effect authoring", "virtual reality authoring", and "virtual reality environment". The sum of our searches turned up a catalogue of more than a thousand results. Duplicate results, poorly relevant academic and erroneous work were removed through filters, leaving us with forty-seven relevant studies in total, albeit none matching our research query fully. Of the forty-seven, the following noteworthy results are mentioned as contributing to our research.

STEVE 2.0 [10] is a graphical authoring tool that uses its own event-based model called MultiSEM (Multimedia Sensory Effect Model) to represent mulsemedia applications and to synchronize their nodes, whether traditional media or sensory effects. The model uses an event-based temporal synchronization paradigm, which allows STEVE's users to define interaction and temporal relations among media items. Since the tool provides a temporal view in the authoring GUI (Graphical User Interface), users do not need to have programming skills to define a mulsemedia application using the STEVE 2.0 environment. It also gives authors feedback about temporal synchronization inconsistencies and provides automatic extraction of sensory effects from audiovisual content analyses by applying machine learning-based methods [1]. STEVE 2.0 provides NCL (Nested Context Language)[2] as a final format that can be run in the Ginga-NCL player, an extension of the Brazilian Digital TV middleware. On the other hand, STEVE does not explore 360° or VR content. In regards to STEVE's underlying model, MultiSEM [10] represents sensory effects as nodes following the approach presented in [15], which models sensory effects as first-class entities. That representation approach uses a high-level abstraction so that the spatio-temporal synchronization of mulsemedia applications can be specified regardless of devices used for implementing them in the real world. We follow this same approach in our proposal for immersive

[2]https://www.itu.int/rec/T-REC-H.761

authoring, with a multisensory module set to be addressed in the future.

With regards to studies that support 360° content, the work of [18] proposes a declarative authoring model that allows authors to design and create 360° interactive videos, focusing on their own declarative language based on XML. As their solution projects media content onto a spherical object, the VR environment creates limitations due to the fact that the user is able to move around freely, which can break the immersion of their 360° video objects. In addition, the creative process is limited to direct editing in XML, to be later viewed through a VR HMD (Head Mounted Display).

In [4], the authors propose a multisensory 360° video authoring tool that allows the author to edit content linked to the video with a basic level of customization. That authoring tool allows the editing and viewing of 360° video with the particularity of allowing the video to be complemented with other sensory stimuli such as tactile and olfactory effects. The same authors in [5] propose an authoring tool with three different authoring interfaces (desktop, immersive and haptic interface) for creating multisensory 360° videos, updating previous work. In [6], they updated the tool to provide some support for collaborative authoring. In their work, the authors were limited to direct desktop editing for later viewing the result through VR HMDs. In addition, sensory effects are not specified using the virtual authoring environment.

The work described in [12] has two main contributions: (1) it investigates existing highlighting methods for VR scene settings that are based on 3D models and explores their suitability for a 360° video VR setting; (2) it proposes immersive authoring methods suitable for non-expert users to create highlights within a 360° video VR application. However, although the tool provides authoring in real time, it is limited to creating menus and inserting highlights in video objects.

Demonstrating the versatility of VR technology, the work of [13] reports the evaluation of an educational toolkit called the Immersive Nugget[3] Tiles (IN-Tiles). With IN-Tiles, the authors report that manipulating VR nuggets and authoring VR learning content can be directly accomplished within a virtual environment allowing authors to immediately experience the results of their authoring efforts in VR. Because this tool is focused on learning, the authors made comparisons with 2D teaching tools. However, the toolkit does not support multimedia content, presenting static virtual content entirely in polygonal 3D.

Citing a popular non-academic tool, we also have the Google Tour Creator,[4] a versatile tool that focuses on the production of 360° environments. It was designed to provide presentations based on tours through Google Maps. Although it is very versatile and enables interactive content, it does not support sensory effects, whilst production and editing via web browsers is limited, in addition to Google having announced the end of the tool support in June 2021.[5]

TABLE 1. Comparison of related studies.

Related studies	3D immersive content	Interactive Content	360° Media	Sensory Effects	Authoring Tool	Fully Immersive Editing
Saleme et al. [21]	No	No	No	Yes	Yes	No
Mattos et al. [10]	No	Yes	No	Yes	Yes	No
Mendes et al. [18]	Yes	No	Yes	No	Yes	No
Coelho et al. [5]	Yes	Yes	Yes	Yes	Yes	No
Horst et al. [13]	Yes	Yes	No	No	Yes	Yes
Coelho et al. [6]	Yes	Yes	Yes	Yes	Yes	No
Comşa et al. [7]	Yes	No	Yes	Yes	No	No
Horst et al. [12]	Yes	Yes	No	No	Yes	Yes
Coelho et al. [4]	Yes	Yes	Yes	Yes	Yes	No
Park and Han [20]	Yes	Yes	Yes	Yes	Yes	No
Fassold and Takacs [11]	No	No	Yes	No	No	No
AMUSEVR	Yes	Yes	Yes	Yes	Yes	Yes

Table 1 summarizes the comparison among related work. Notwithstanding related tools, we are not aware of any available solution which offers unique support to all features like AMUSEVR does, such as: authoring in real time within a virtual reality environment, including media objects and sensory effects in any location in the 3D environment, in addition to supporting interaction with media objects, transforming the presentation from passive to interactive, alongside supporting a flexible XML language such as MultiSEL as an interchange format, as will be discussed in the following sections.

III. PROPOSAL FOR IMMERSIVE AUTHORING OF 360° MULSEMEDIA

While there are a myriad of 2D authoring tools currently available, including a few which support mulsemedia, when it comes to immersive mulsemedia content, either the number of options is greatly diminished, or the support options for multiple media types is reduced. Various factors are responsible for limiting the number of existing proposals. These factors include difficulty in implementing a management system to control diverse immersive media content and limitations in editing and viewing the content in real time.

Taking into consideration some related work cited in Section II and other studies [9], as well as our desired goal, we defined the following desirable requirements for an immersive authoring environment and the execution of 360° mulsemedia applications. The requirements related to 360° content, layout and scenes were based on requirements proposed in studies such as [18] and [6]. In respect of the requirement to use an interchangeable format to import and export projects, in addition to relationship synchronization

[3]Nugget is the author's reference to small segments of content.

[4]https://arvr.google.com/tourcreator/

[5]https://support.google.com/edu/expeditions/answer/10086878?hl = en

tools, we were inspired by [1] and [18]. Regarding the support for sensory effects, we got some ideas from [6] and [4]. Accordingly, the environment should:

- offer authoring functionalities and also run 360° mulsemedia applications in a VR platform;
- allow the author to edit and preview his/her 360° interactive applications in a similar way that the end-user will experience it, so that the author will have as close an experience as possible to that of the end-users;
- support a 360° background, which can be disabled if so desired;
- allow the author to insert and manipulate media in a number of formats, including traditional media objects (video, audio, still images, and text), 360° media (sill images and video), sensory effects (light, wind, scent, temperature changes, etc.);
- allow six-degrees of movement (6DoF) to the user;
- provide preview options for the full 360° content without the need to save or leave the immersive environment;
- support a number of mulsemedia content formats in simultaneous playback according to the author specification;
- provide support for user interactivity in any application to be created;
- support event-based relationships [9] between media items to allow for flexible media synchronization and interactions associated with media objects;
- support physical actuator-operated sensory effect systems while also rendering the visual counterparts of the said sensory effects in the VR environment. This requirement is left for future implementation; and
- provide and use an interchangeable format to allow application interchange with other authoring environments. XML is a strong candidate due to its flexibility and ease of adaptation to new technologies.

Aiming at satisfying those requirements, we propose an immersive environment as illustrated in Figure 1. Observe how the author (or end-user) is free to move within the virtual environment using an HMD and interacting with diverse media objects (2D or 360°) and sensory effects, within an open virtual environment.

To provide an improved form of separation and organization of several 360° content in the same immersive application, we propose organizing the application into scenes. One *scene* can be composed of a single 360° media item and several other 2D media items and sensory effects, which are positioned in the virtual 3D space. Using scenes improves distribution and storage of media content to be processed in the VR equipment memory. Each scene is independent and has its own media objects, enabling the author to create multiple scenes with different themes, which allow for diverse scenarios in the same presentation. The end-user can navigate from one scene to another interacting with media objects in the previous scene. Figure 2 shows the class diagram of our proposal, where a scene is defined as a composition of media objects and sensory effects. Links allow defining interaction relationships between media objects or other scenes.

FIGURE 1. Illustration of a scene in our proposal.

To synchronize the presentation of media objects in a scene, a hybrid methodology was adopted, unifying the concepts of timeline and event-based paradigms to create relationships between media objects. Each scene, when started, serves as a time reference for all its media and sensory effect components.

IV. AMUSEVR

The implementation of our proposal of an immersive authoring system for 360° interactive multisensory applications is named AMUSEVR[6] and was developed using the Unity engine, the HTC Vive[7] virtual-reality HMD and the Windows platform. For demonstration and usability tests, we used media content available at free websites.[8]

As seen in the use case diagram shown in Figure 3, the system was developed with two profiles in mind. For the end-user profile, we have the ''Viewer Mode'', where it is possible to import a saved project and thereby experience 360° interactive applications. For the author profile, there is also the ''Author Mode'', where the author creates and edits the content in a truly immersive environment. In this particular mode, as seen in Figure 4, the headset and controls allow the author to insert media objects, edit them by changing their properties individually, or remove them from the project. For this purpose, we designed for right and left-handed people a

[6]A video about AMUSEVR is available at https://youtu.be/qtAWrNpSuXQ

[7]https://www.vive.com/us/

[8]https://www.pexels.com/ and https://www.mettle.com/360vr-master-series-free-360-downloads-page/

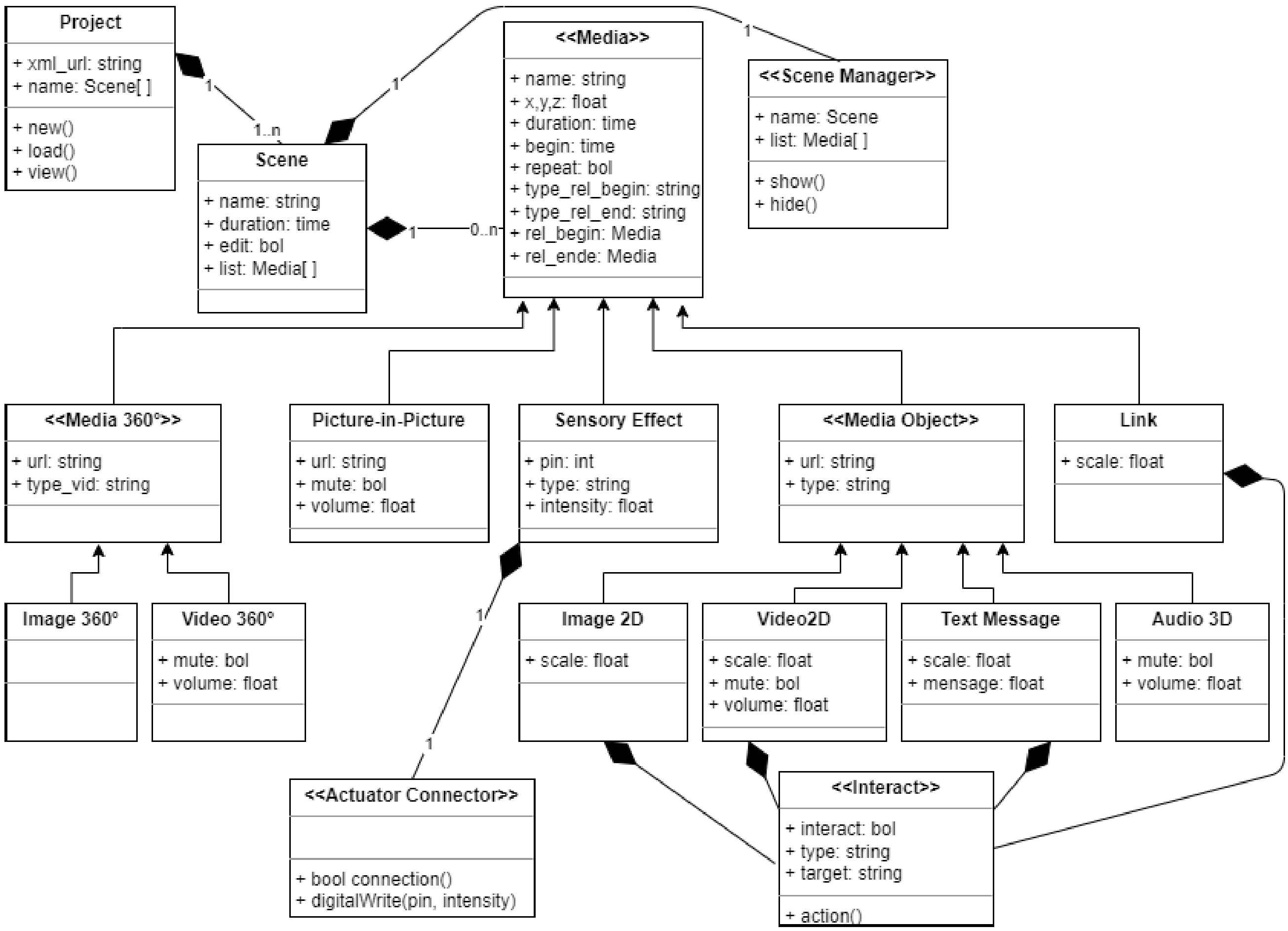

FIGURE 2. Class diagram of our proposal.

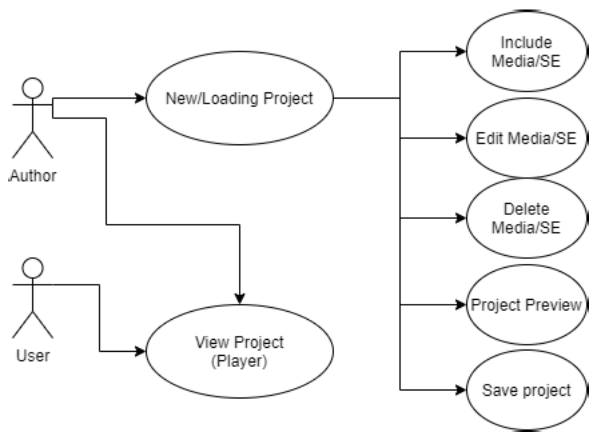

FIGURE 3. AMUSEVR-Use case diagram.

menu that contains options and methods for the scene to be positioned above the left control device by default. The right control device is designed to point at the menus and further manipulate objects.

A notable difference in the authoring methodology proposed by our work is that traditional applications work first by selecting the media object file, then editing or deleting them. However, in our solution, each media object type presents different behavior and properties. For example, a 2D image can be selected through pointing, then edited or deleted. On the other hand, a 360° image must not be selected through pointing, as it could be selected unintentionally due to how the author's field of view is fully taken up by the 360° image. In that case, the 360° object can be selected via a dropdown menu in the scene editing menu shown in Figure 4. Therefore, the proposed authoring steps follow the flowchart shown in Figure 5, where the command functions seen in the menu in Figure 4 are chosen first, so the appropriate order of commands is followed, depending on the type of media object.

Figure 6 illustrates a very important aspect of the proposed 3D spatial concept, the ability to position 2D media objects anywhere in the x, y and z axes within the 3D environment, simply with a drag-and-drop sequence. Additionally in Figure 6, on the left we have a photo of a tourist spot (Sugar Loaf) in the city of Rio de Janeiro, which is a 2D image-type media object. On the right side of the figure, there is another 2D-video media object, portraying a seal at a beach. Below the Sugar Loaf photo, the scene contains yet a third text

FIGURE 4. AMUSEVR-Author mode menu and controls.

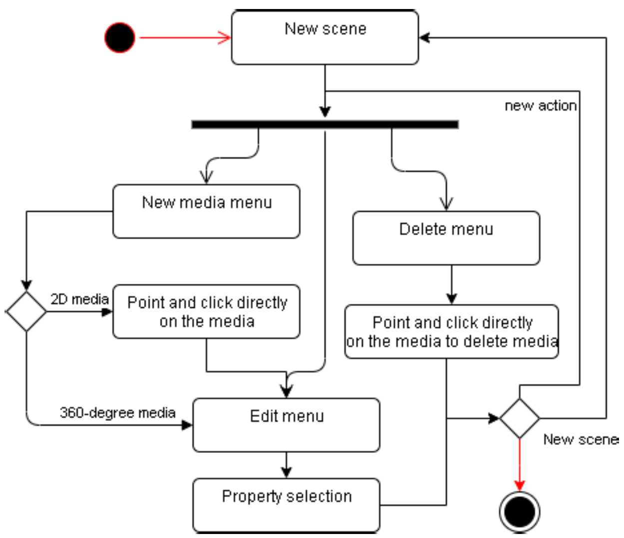

FIGURE 5. AMUSEVR-Activity diagram.

FIGURE 6. 2D media objects in a 360° scene.

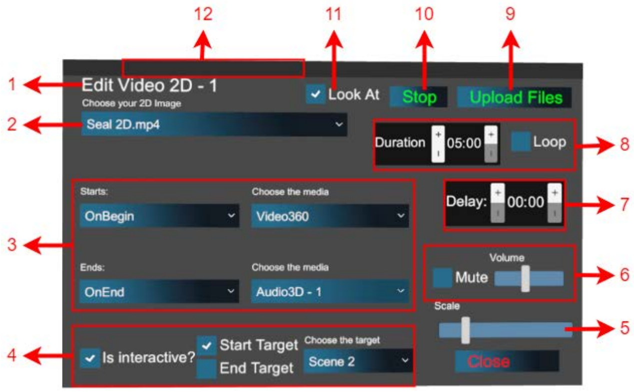

FIGURE 7. Video object editing in detail.

type media object, containing the phrase ''YOUR MESSAGE HERE'', and finally, a fourth 360° video media object of a beach scene in the background. All media objects have editable positions and attributes within the scene, allowing authors to fine tune their characteristics.

In Figure 7 a sample 2D-video editing window can be seen. Its properties are as follows:

1) Displays the media ID;
2) A dropdown list of any compatible media content file found within the installation folder.
3) Event-based relationships specifying how this media object begins or ends. Inside the ''Starts'' combo box, we can find the options ''OnBegin This Scene'', ''OnBegin'', ''OnEnd'', and ''Not defined'', each being related to another media object within the same scene. The ''Ends'' combo box options are ''End of This Media Time'', ''OnBegin'', and ''OnEnd'', which are also related to another media object within the scene;
4) User interaction options group. Upon enabling an object to be interactive, it can start or end another scene or media object when the end-user interacts with it;
5) Scale slide bar to change the video media object size;
6) Group of audio options related to the object;
7) Defines a delay for the object start time;
8) Controls the duration of still media items and/or sensory effects. Controls video/audio looping;
9) Imports new media files into the installation folder to populate the dropdown menu;
10) Starts or stops media during authoring mode;
11) The ''Look at'' option keeps media oriented towards the user; and
12) A strip allowing the user to drag and drop the object to another position.

Figure 8 shows the Editing Mode for a new scene in which we see a scenario composed of a 3D audio media object, a 360° panoramic image and a 2D video with its editing properties being displayed. In this sample scene shown in Figure 8, we highlight how AMUSEVR provides editing methods, such as displaying or hiding media elements in the scene, as well as creating links to other scenes, thereby enabling transitions between scenes of the same 360° project.

FIGURE 8. Scene editing mode.

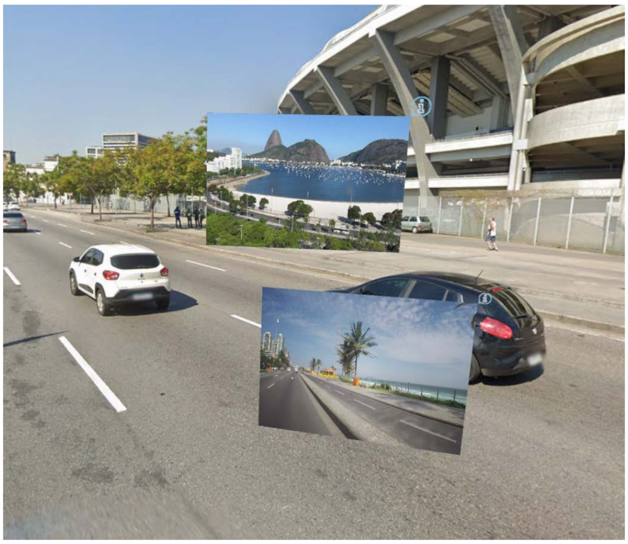

FIGURE 9. A scene with an interactive video and still image.

This feature transforms projects from passive to interactive presentations, providing options to define interactive objects and create hyperlinks.

Noteworthy is the "Is interactive?" option in Figure 8. When the author specifies that an object is interactive, he/she can specify which other object should be started ("Start Target") or ended ("End Target") when the end-user interacts with it. Upon activating the "Is interactive?" attribute, the correspondent object gains button-like properties along with an icon in the shape of an "*i*" in the upper-right corner, as seen in the Sugar Loaf 2D-video in the middle of Figure 9. These functions grant creative freedom and functionality within diverse applications where the user can, for example, move between the rooms of a virtual museum or scenes for 360° interactive movies, or to switch between spoken or subtitled language. The interactive function can be run at any time with a click on the object during preview or playback mode. As the actions are activated by interactive functions, one can show or hide content as previously set up in authoring mode. Thereby, one is able to transition from one scene to the next by clicking on media objects, show or hide subtitles, enable ASL (American Sign Language) PIP (Picture-in-Picture), change the audio or version of a video, all through user interaction. For example, in Figure 8, object "Video 2D - 1" is interactive, and when the end-user clicks on it, it will start "Scene 2".

In AMUSEVR Viewer Mode, each scene, upon being accessed, runs its own discrete clock, independent of the other scenes, which serves as time reference for all media contained inside it. However, media objects also respond to interaction or synchronous events defined by event-based relations. Each media object has its own start, end and delay time properties, which are displayed or not depending on the author's preferences.

As previously commented, another way to control the presentation of media objects in the scene is through the event-based relationship paradigm, in which it is possible to associate the beginning and end of the presentation of a media object to another media object. In Figure 7, the "Starts" menu defines a start relationship with media object "Video360" through the option "OnBegin". That is, the object being edited, called "Video 2D - 1", will only start when object "Video360" has started. For the same "Video 2D - 1" object, we have another event-based relationship called "Ends", associated to the end ("OnEnd") of object "Audio3D - 1". Thus, "Video 2D - 1" will only end when object "Audio3D - 1" ends. Through this, it is possible for the author to schedule the presentation of objects in the scene in an event-based manner and to remove them from the scene automatically as needed. The author can also specify hyperlinks on interactive objects to enable presentation of other media objects in a scene together with synchronized events.

As shown in Figure 7, the object called "Video 2D - 1" has further options, such as duration, loop, delay, volume and scale. These other properties may or may not be available according to the object type. See Table 2 for detail about the extra properties of media objects and sensory effects.

More information about supported file types, codecs, etc., can be found at the website of the associated documentation of the Unity engine.[9]

Regarding the communication between AMUSEVR and physical actuators for rendering sensory effects, this issue is left for future work. Moreover, freedom of movement in AMUSEVR will be discussed in detail in Section IV-A. In respect to the method for importing and exporting AMUSEVR projects, XML was chosen as the interchange format, as will be discussed in Section IV-B.

A. SIX DEGREES OF FREEDOM

AMUSEVR has both authoring and visualization modes that support movement in six degrees of freedom (6DoF), aiming at supporting creative freedom and flexibility for authors and greater immersion for participants in a presentation. This is possible because of the ability of the user to walk around the virtual scenario freely, limited only by one's physical environment in the real world.

[9]https://docs.unity3d.com/Manual

TABLE 2. Specific properties of media objects and sensory effects.

Object types	Properties
3D audio	Loop, volume and mute
2D image	Scale
360° image	Loop
Interactive object	Target object/scene or related start/end event
Picture in picture	Loop, volume and mute
Sensory Effects	Loop, duration and intensity
Text message	Loop and text content
2D video	Loop, volume, mute and scale
360° Video	Loop, volume and mute

In AMUSEVR Author Mode, media and sensory effects being edited provide an option named ''Look at'', shown as item 11 in Figure 7. Whenever it is enabled, it causes the 2D media, audio or even a sensory effect, to rotate on its central axis and focus its front face or the direction of emission towards the center of the virtual reality HMD of the author (or end-user), yet still maintaining its anchor at the X, Y, and Z position defined by the author. When the option is disabled, the 2D media, audio or even sensory effect, stops updating the focus in the virtual reality HMD and retains its most recently updated facing direction.

By means of the ''Look at'' media property, along with participants being allowed complete freedom of movement within the presentation, the author can design immersive scenarios where media positioning is coordinated with the users' participation.

Considering that audio objects in AMUSEVR support 3D audio, the sound is spread spatially from approximately 1 meter of the point of emission (approximate distance within the virtual world), with linear volume loss until reaching approximately 10 meters of radius from the point of emission. When 3D audio objects produce sounds recorded in a mono format, they distribute the sound equally in all directions; whereas, when the content is recorded in multichannel mode, it uses the visual icon's face as a central reference for sound emission.

Also taking 2D images and videos into consideration, the content will be presented through the front of the frame (layer). This means that a participant observing laterally at 90 degrees from the front or directly behind the 2D image or video media object will not be able to see the content being displayed. Following the considerations above, Figure 10 describes customized micro-scenarios which exemplify the use of media, such as 3D audio and 2D images, with the position of the participant being taken into account: In frame A - we have audio which possibly has the ''Look at'' function enabled and the participant is in the sound range of the audio, he will perceive the audio as being transmitted in front of him. In frame B - we have audio which possibly has the ''Look at'' function enabled and only the left side of the participant is in the audio's sound radius. He will perceive the audio as being transmitted from his left direction. In frame C - we

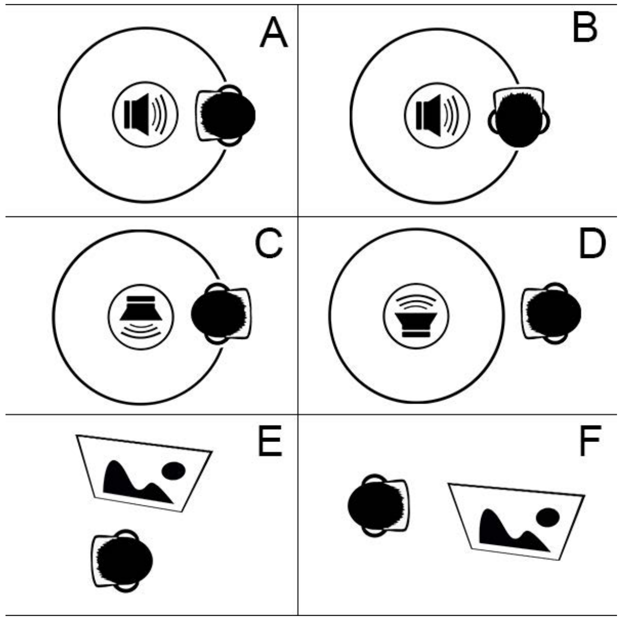

FIGURE 10. Example of 6DoF.

have audio which has the ''Look at'' function disabled and the participant is in the sound radius of the audio, he will perceive the audio as being transmitted behind him. In frame D – we have audio which has the ''Look at'' function disabled and the participant is outside the sound range of the audio, he will not hear that particular audio object. In frame E – we have an image which possibly has the ''Look at'' function enabled and the participant will be able to view the image as soon as the participant's field of vision is rotated to the right side. In frame F – we have an image which possibly has the ''Look at'' function disabled and the participant may have difficulty viewing the 2D image object, unless he/she is positioned further in front of the image and consequently look towards it. Considering the collection of scenarios in Figure 10 we can see how the author, having the freedom to place different content to be displayed at different times in different locations, would take advantage of the multitude of options during the creative process of his immersive projects.

B. THE MultiSEL AUTHORING LANGUAGE

MultiSEL (Multimedia Sensory Effect Language) is an XML-based language for specifying mulsemedia applications with 360° visual content for virtual reality environments. It is also based on the MultiSEM model [10] and uses the declarative approach to allow authors to specify 360° multisensory applications. MultiSEL also aims at aiding the interoperability between 2D/3D mulsemedia authoring tools.

In our proposal, AMUSEVR uses MultiSEL as an interchange format between the Author Mode and the Viewer Mode. Therefore, an author can create or edit an immersive 360° application using the Author Mode and distribute it later to end users to experience it using the Viewer Mode. We chose this language as it has proven its value for editing,

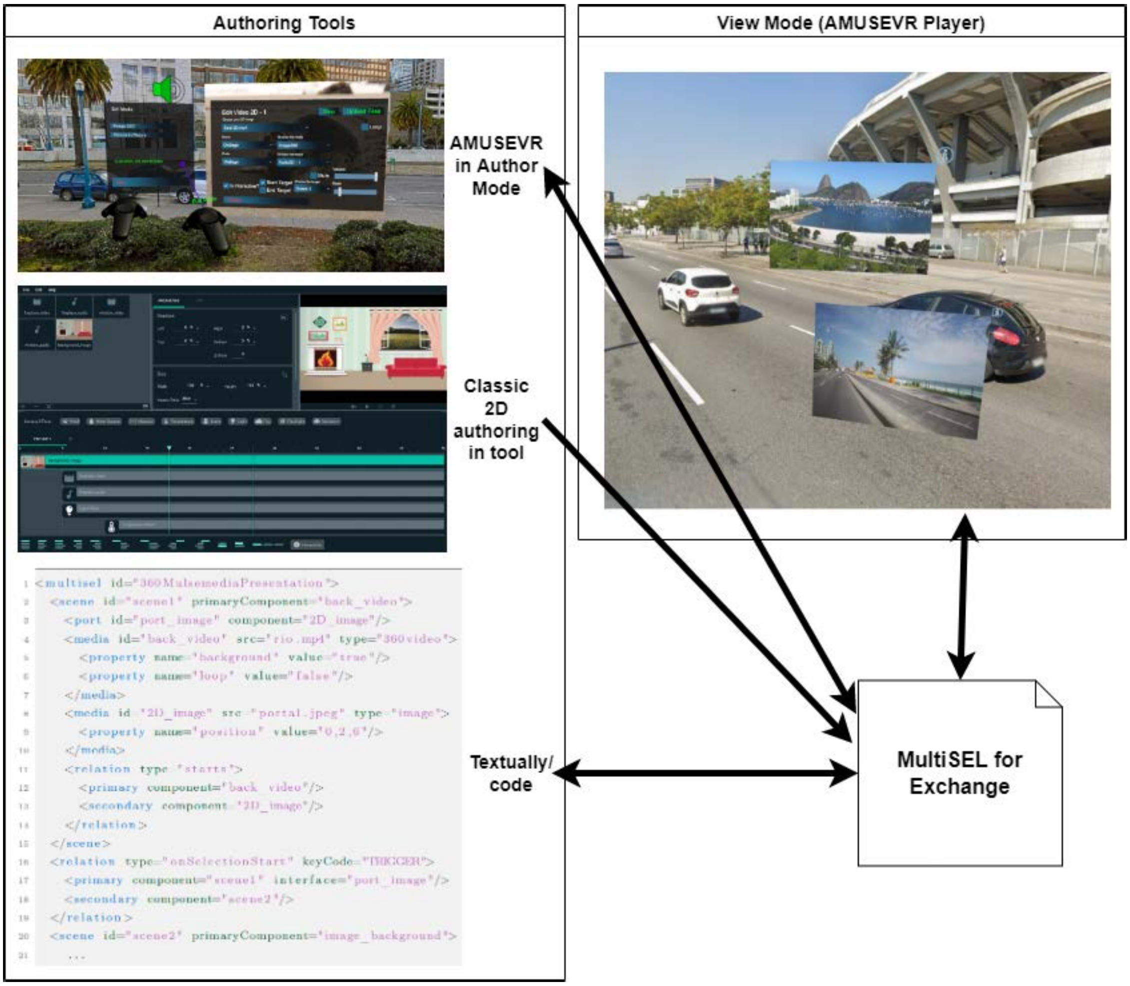

FIGURE 11. **MultiSEL document example of portability.**

as well as benefiting from the clarity provided by the XML structure. MultiSEL was projected as the basis of future support for tools which work with multiple 3D scenes, media, and sensory effects simultaneously, as well as providing concepts such as interaction relations. As our proposal is an authoring and editing approach using controls within an immersive environment, it becomes possible for authors to start developing their application in a 2D editor, such as STEVE [9], to export an intermediary project in MultiSEL, and then import the MultiSEL project into AMUSEVR. Alternatively, expert authors can simply create their MultiSEL project directly in XML and then refine and finish their project in AMUSEVR. The possibilities provided can be seen in Figure 11.

A MultiSEL document can be composed of several scenes, declared in the *body* element, as shown in Listing 1. The first scene to be presented is defined by the *body primaryComponent* attribute.

In MultiSEL, a 360° scene is represented by the *scene* element, which has the attributes *id* and *primaryComponent* attributes. The latter defines the first media node to be presented when the scene starts. The *scene* element also contains the following child elements: the document nodes (media and sensory effects) and their relations, the *port* and *property* elements. The *port* element provides external access to the scene's inner elements. It defines interface points that relations use to refer to inner nodes of a scene. To do that, the *port* element defines the *id* attribute for identifying the interface point and the *component* attribute for referring to an inner node of the scene to which the interface point maps. This *component* attribute can also refer to a scene property. In addition, the *port* element may define the *interface* attribute to reference media and sensory effect properties. Concerning the *property* element, it defines scene variables that can be used in temporal relations among the document nodes.

MultiSEL also provides the *relation* element to specify temporal relations among components, such as *scene*, *media* and *effect*. The language defines two basic attributes for the *relation* element: *type* and *delay*. The *type* attribute allows authors to give the temporal semantic among the relation elements. To do that, MultiSEL is based on the predefined

TABLE 3. MultiSEL synchronous temporal relation types.

Relations	Symbols	Descriptions
Starts		Nodes begin when the primary node starts
Starts_Delay		Nodes begin with delay when the primary node starts
Finishes		Nodes end when the primary node finishes
Finishes_Delay		Nodes end with delay when the primary node finishes
Meet		*Nodes begin when the primary node finishes*
Meets_Delay		Nodes begin with delay when the primary node finishes
Met_By		*Nodes end when the primary node starts*
Met_By_Delay		Nodes end with delay when the primary node starts
Before		*Present nodes sequentially with a delay between them when the primary node finishes*

temporal relations defined in MultiSEM temporal synchronization model [10]. Therefore, the type attribute can assume the following values to define synchronous temporal relations: *starts*, *startsDelay*, *finishes*, *finishesDelay*, *meet*, *meetDelay*, *metBy*, *metByDelay* and *before*. The *delay* attribute may be necessary according to *type*. For example, if the *type* attribute is *startsDelay* then the *delay* attribute should be defined. In this case, the relation starts the secondary nodes with that delay after the primary node starts. The delay can also be specified for the relation's secondary participants separately. Table 3 shows the graphical representations and descriptions for those relation types. Those synchronous relations comprise one primary node (green rectangle) and one or more secondary nodes (gray rectangles). The primary node plays the condition role (*starts* or *stops*) and the secondary plays the action role (*start* or *stop*) [10].

MultiSEL also provides asynchronous relation types to specify user interaction. The relation type *onSelectionStart* specifies that when the user interacts with the relation primary node, the *secondary* one will be started. In this case, the primary node represents the media item end-users need to interact with using an interactivity key to trigger the associated action. The relation type *onSelectionStop* specifies that when the user interacts with the relation primary node, the *secondary* one will be stopped.

An example of MultiSEL code is given in Listing 1. This MultiSEL document defines two scenes: *scene1* and *scene2*. *scene1* starts with *video_background* media as

```
1  <?xml version="1.0" encoding="utf-8"?>
2  <multisel id="360MulsemediaPresentation" xmlns="
     DefaultProfile">
3    <head>
4      <meta name="author" value="anonymous" />
5      <meta name="year" value="2022" />
6    </head>
7    <body primaryComponent="scene1">
8        <scene id="scene1" primaryComponent="
     video_background">
9          <port id="port_image" component="2D_image"/>
10         <media id="video_background" src="rio.mp4" type="
     360video">
11           <property name="background" value="true"/>
12           <property name="loop" value="false"/>
13         </media>
14         <media id="2D_image" src="portal.jpeg" type="
     image">
15             <property name="x" value="0" />
16             <property name="y" value="2" />
17             <property name="z" value="6" />
18         </media>
19         <relation type="starts">
20           <primary component="video_background"/>
21           <secondary component="2D_image"/>
22         </relation>
23       </scene>
24       <scene id="scene2" primaryComponent="
     image_background">
25         ...
26         <effect id="wind_effect" type="windEffect">
27           <property name="intensity" value="50"/>
28             <property name="x" value="0" />
29             <property name="y" value="6" />
30             <property name="z" value="12" />
31         </effect>
32         <relation type="starts" delay="10">
33           <primary component="image_background"/>
34           <secondary component="wind_effect"/>
35         </relation>
36       </scene>
37       <relation type="onSelectionStart" keyCode="trigger"
     >
38         <primary component="scene1" interface="port_image
     "/>
39         <secondary component="scene2"/>
40       </relation>
41    </body>
42  </multisel>
```

LISTING 1. MultiSEL document example.

its 360° background. It also contains a relation element whose type is *starts*. This relation defines that the secondary component (*2D_image*) begins when the primary one (*video_background*) starts. Included in the *body* element, the document specifies a relation of type *onSelectionStart* between both scenes. It defines that *scene2* begins when media *2D_image*, accessed through the *port* element in *scene1*, is selected using the *trigger* key. *scene2* contains a wind sensory effect that begins with a 10s delay when *image_background* starts. *image_background* is a media object contained in *scene2*, but is not shown in Listing 1.

In demonstration of the equivalence between AMUSEVR and MultiSEL, using different scenarios as examples, the equivalent of the AMUSEVR relation category "Starts" shown in item 3 of Figure 7 can be seen in lines 19-22 in Listing 1. Another case, this one involving interactivity, can be found in item 4 of Figure 7, which represents the relation found in lines 37-40 in Listing 1.

When importing a MultiSEL document into AMUSEVR, sensory effects are considered in the virtual environment in

TABLE 4. The goals which were defined for the AMUSEVR experiments.

Goal	Description
G1	Analyze the tool with the goal of evaluating its usability from the user's perspective.
G2	Analyze the tool with the goal of evaluating its quality of experience from the user's perspective.
G3	Analyze the tool with the goal of evaluating the authoring effort required for both expert and novice users.
G4	Analyze the tool with the goal of evaluating the authoring effort compared to an alternative method.

the current implementation. As future work, we are going to implement physical rendering of sensory effects.

V. PROPOSAL EVALUATION

This section discusses the evaluation process for this proposal via experiments where users utilized the AMUSEVR tool as a case study prototype. To that end, we used the Goal Question Metric (GQM) [2] evaluation techniques and methods, the User Experience Questionnaire (UEQ)[10] [16], [22], and the System Usability Scale questionnaire - SUS[11] [3].

For the purpose of describing the process, the methodology used to structure and to elaborate our evaluation are discussed in Section V-A. The user experiment is discussed in Section V-B and a quantitative evaluation is described in Section V-C. In Section V-D, we present the obtained results and finally in Section V-E, an analysis and discussion of final conclusions and current limitations are given.

A. METHODOLOGY

We used the Goal Question Metric (GQM) [2] approach to structure our evaluation. The GQM mechanism can be visualized as a flowchart in which, following the top-down concept, the flow starts in goal nodes, continues to the question nodes and ends in the metric nodes. Furthermore, the GQM guidelines aims at defining the proposal and the perspective of the goals. The proposal defines the object being studied and motives for its study. The perspective defines the personal aspect or point of view given by the one performing the evaluation.

Applying the GQM method results in a three-tier hierarchy model: Conceptual (goals), Operational (questions), and Quantitative (metrics). The model definition has the establishment of goals as an initial step. Each defined goal is decomposed into questions that aim at answering them, making them quantifiable. The questions are, in turn, refined into objective or subjective metrics.

Table 4 shows the goals defined for the evaluation of AMUSEVR. For each goal G1-G4, the "Description" column describes the object being analyzed.

[10]https://www.ueq-online.org/

[11]https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

TABLE 5. SUS questionnaire.

Items	Questions
SUS-Q1	I think that I would like to use this system frequently.
SUS-Q2	I found the system unnecessarily complex.
SUS-Q3	I thought the system was easy to use.
SUS-Q4	I think that I would need the support of a technical person to be able to use this system.
SUS-Q5	I found the various functions in this system were well integrated.
SUS-Q6	I thought there was too much inconsistency in this system.
SUS-Q7	I would imagine that most people would learn to use this system very quickly.
SUS-Q8	I found the system very cumbersome to use.
SUS-Q9	I felt very confident using the system.
SUS-Q10	I needed to learn a lot of things before I could get going with this system.

TABLE 6. UEQ-S questionnaire.

Items	The AMUSEVR experience was	
UEQ-Q1	Obstructive	Supportive
UEQ-Q2	Complicated	Easy
UEQ-Q3	Inefficient	Efficient
UEQ-Q4	Confusing	Clear
UEQ-Q5	Boring	Exciting
UEQ-Q6	Not interesting	Interesting
UEQ-Q7	Conventional	Inventive
UEQ-Q8	Usual	Leading edge

1) G1 - QUESTIONS AND METRICS

In order to reach goal G1, i.e., analyze the tool with the goal of evaluating its usability from the user's perspective, the System Usability Scale (SUS) questionnaire [3] was used, as it is a well-known tool for measuring usability. SUS specifies a questionnaire of ten items shown in Table 5, where each one accepts responses in the Likert scale, ranging from "Strongly disagree" ranked at 1 to "Strongly agree" ranked at 5 [17]. We used the SUS score as metric for G1 questions.

2) G2 - QUESTIONS AND METRICS

To satisfy goal G2, i.e., analyze the tool with the goal of evaluating its quality of experience from the user's perspective, the User Experience Questionnaire (UEQ-S) [16], [22] was used. As can be seen in Table 6, the questionnaire contains eight items (UEQ-Q1 to UEQ-Q8) with negative feedback in the middle column and positive feedback in the right column in response to the sentence "The AMUSEVR experience was." The participants answers were on a scale of one (negative feedback) to seven (positive feedback). We used UEQ pragmatic and hedonic scores as metrics for G2 questions.

3) G3 - QUESTIONS AND METRICS

To meet goal G3, i.e., analyze the tool with the goal of evaluating the authoring effort required for both expert and novice users, a set of specific additional questions were used as can be seen in Table 7. Q1 ascertained if the participant

TABLE 7. Custom questions.

Q#	Questions
Q1	Did you manage to finish the task?
Q2	Have you used virtual reality before?
Q3	Do you have experience with multimedia authoring tools?
Q4	How many minutes did you take to complete the task?

was able to complete a task, and if so, the time required for task completion was annotated in Q4. Questions Q2 and Q3 were intended to generate correlation towards which users had previous experience in virtual reality and authoring tools, separating experts and novices.

Question Q2 was designed to separate the participants into two groups, those who had already used similar technologies for immersion and those who had not. The hypothesis is that participants with experience and greater comfort in the use of virtual reality equipment would do better than those who have no experience.

Question Q3 was designed to separate the participants into two groups, those who had already worked on multimedia authoring tasks and those who had not. The hypothesis is that participants with knowledge of multimedia authoring would do better than those with no experience.

We used the editing time taken to complete the task as metric to answer G3 questions comparing expert and novice users.

4) G4 - QUESTIONS AND METRICS

For goal G4, i.e., analyze the tool with the goal of evaluating the authoring effort compared to an alternative method, we used MultiSEL as the alternative authoring method. Therefore, one question Q5 is defined, which is "What is the required number of clicks on the AMUSEVR tool compared to the number of lines of a correspondent MultiSEL file with a similar specification?". In order to answer that question, a comparative test for AMUSEVR was done against the declarative authoring method based on the MultiSEL language, considering different taks as seen in Table 8. We used a quantitative metric based on the number of clicks between AMUSEVR and the number of lines of MultiSEL code for each task, which will be detailed in Section V-C.

We recognize that this comparison between AMUSEVR and MultiSEL may not be so fair, but as there is no other available alternative method or tool to compare AMUSEVR authoring effort with, we used MultiSEL for this initial evaluation.

5) GQM GRAPH

Figure 12 summarizes and illustrates the structure of the GQM model built. The experiments performed for data collection will be presented in the following subsections. Our experiments were divided into two parts, one experimentation with participants to meet G1-G3 and the second that was

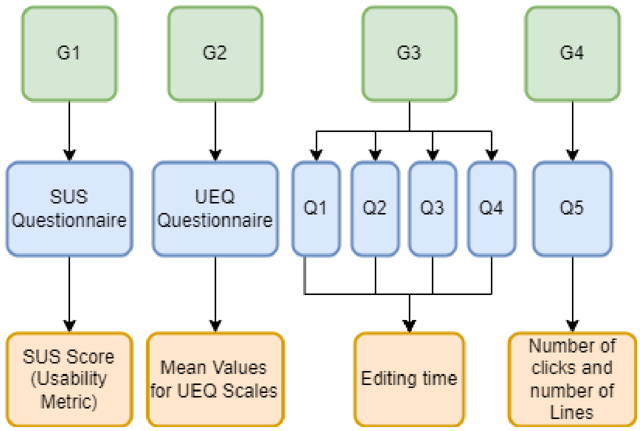

FIGURE 12. Directed graph for our goals, questions and metrics (GQM).

carried out at another time in order to analyze quantitative data to reach G4.

B. USER EXPERIMENTS

User experiments were done to investigate our main research question: *"What is the user experience of creating 360° interactive multisensory applications using an immersive authoring environment?"*. Due to the COVID-19 pandemic and resulting health restrictions on the use and sharing of rooms and equipment, it has not been possible to carry out experiments with a large number of candidates; yet tests have been carried out with ten (10) users, a number sufficient to highlight most usability issues [14]. They were eight men and two women aged between 23 to 47 years old, with education levels ranging from undergraduate to post-graduate. Only five had not used virtual reality equipment previously, while four had not had any experience of multimedia authoring systems; a single participant was left-handed.

Participants were given the option to choose whether to sit or stand during the immersive authoring experience, as illustrated in Figure 13, with a test supervisor monitoring the task. The test was monitored and displayed on an additional screen in real time. A video about the experiment is available[12].

All participants were given time to familiarize themselves with the environment and tools used (five minutes on average), after which, each one started the tasks and had their results recorded. The tasks aimed at creating an interactive 360° presentation of a tourism showcase of Rio de Janeiro.

The tasks were defined according to some requirements:

- the project should be short with a maximum of two scenes, Scenes 1 and 2, as participants may have no experience in authoring tools or in virtual reality. Also, predicting possible future comparisons with other tools or even a MultiSEL project;
- to cover some highlights of the tool, there should be an example of background media, some classic 2D media

[12]https://youtu.be/KPFfSj-5_qk

FIGURE 13. Experiment with a participant (seated position).

and a sensory effect, as well as an example of a synchronization relationship and interaction between scenes;
- participants must activate the preview mode in order to assess whether they are satisfied with their project and test the programmed interactivity.

To guide users participating in the experiment, we provided a table in the control menu describing the tasks to be performed, which are:
- insert a 360° video or image in Scene 1 and Scene 2;
- insert a 2D image or video in Scene 1;
- create a synchronization relationship between the 2D image or video with the 360° video or image, in Scene 1;
- edit the video or 2D image to make it interactive and start Scene 2;
- in Scene 2, use a sensory effect, and
- upon completion, return to Scene 1 in Viewer Mode to experience the final result.

After completing the tasks, participants were asked to fill out a survey which covered the SUS and UEQ questionnaires, as well as questions about age, education level, and experience with multimedia authoring tools.

C. AUTHORING EFFORT - AMUSEVR X MultiSEL

As an initial method of quantitatively comparing authoring effort between AMUSEVR and MultiSEL, we focused on the number of clicks necessary to complete tasks in AMUSEVR against the number of lines within a MultiSEL document

```
<?xml version="1.0" encoding="UTF-8"?>
<multisel id="Project-1" title="MultiSEL for AMUSEVR"
    xmlns="DefaultProfile">
  <head>
    <meta name="author" value="anonymous" />
    <meta name="year" value="2022" />
  </head>
  <body primaryComponent="scene 1">
    <scene id="scene 1" primaryComponent="video360_s1">
      <media id="video360_s1" src="demo.mp4" type="video"
      >
        <property name="duration" value="259" />
        <property name="background" value="true" />
        <property name="loop" value="false" />
        <property name="volume" value="0.72" />
        <property name="mute" value="false" />
      </media>
      <relation id="Relation-1" type="finishes">
        <primary component="scene 1" />
        <secondary component="video360_s1" />
      </relation>
    </scene>
  </body>
</multisel>
```

LISTING 2. Example of an AMUSEVR-compatible T1 task done in MultiSEL.

with the same specification. In order to run this comparison, we created different applications with AMUSEVR. After editing each project in AMUSEVR, we exported it as a MultiSEL project for the purpose of establishing a comparative standard for each different level of specification.

For the first comparative task, we set up the 360-degree background video demo.mp4. Completing this task in AMUSEVR involved: a right-hand pointer click on the left-hand menu button "New Media" as shown in Figure 4, followed by a click on the "New 360 Video" button in the newly-opened menu. A new click was done on the media list menu contained within the floating media editing menu (see Figure 7), which lists all available media files to be used as a 360 background. To export the project, a click on the left-hand menu button "Back" by which the user is taken back to the main menu. Another click on "Exit" closes the project, and to complete the task, a final click on "Yes" to confirm the export. In summary, six clicks were required to complete the task. The exported file resulting from the AMUSEVR task produces Listing 2, a 22-line MultiSEL document. In this example, a 360 degree video was used in the background, but for each type of media, there is an evaluation of the type of properties, resulting in a variation in the number of lines.

To avoid bias, we established six identical yet progressively more difficult tasks to compare AMUSEVR and MultiSEL authoring, as seen in Table 8 (T1 to T6).

D. RESULTS

This section analyzes and discusses the data obtained from the experiments and the questionnaires which answer the goals defined previously.

1) G1—ANALYSIS OF RESULTS—SUS SCORE

Considering goal G1, for analyzing the usability of AMUSEVR, the SUS questionnaire returned an A-grade average

TABLE 8. Comparative Tasks.

T#	Tasks
T1	Scene 1 containing a 360-degree video background (like Listing 2).
T2	Scene 1 containing a 3D audio.
T3	Scene 1 containing a 360-degree video background and a 3D audio.
T4	Scene 1 containing a 360-degree video background and an interactive button for scene 2. Scene 2 with one 360 degree image.
T5	Scene 1 containing a 360-degree video background and an interactive button for scene 2. Scene 2 with one 360 degree image and one 2D video.
T6	Scene 1 containing a 360-degree video background and an interactive button for Scene 2. Scene 2 with one 360 degree image and an interactive 2D video for Scene 3. Scene 3 with one heat sensory effect.

TABLE 9. Number of users and their answers to the SUS questionnaire statements (1 - strongly disagree; 5 - strongly agree).

SUS Questions	1	2	3	4	5
I think that I would like to use this system frequently.	0	0	1	5	4
I found the system unnecessarily complex.	6	1	3	0	0
I thought the system was easy to use.	0	0	1	4	5
I think that I would need the support of a technical person to be able to use this system.	3	3	1	1	2
I found the various functions in this system were well integrated.	0	0	0	4	6
I thought there was too much inconsistency in this system.	7	2	1	0	0
I would imagine that most people would learn to use this system very quickly.	0	0	2	1	7
I found the system very cumbersome to use.	6	2	1	1	0
I felt very confident using the system.	0	0	0	4	6
I needed to learn a lot of things before I could get going with this system.	6	1	0	2	1

score of 82.25 points, as illustrated in Figure 14. This score was obtained through the data generated by G1 questions shown in Table 9. The SUS score result is promising for system usability, as it is higher than 68 points that represent the baseline average [3].

2) G2—ANALYSIS OF RESULTS—UEQ METRICS

To achieve goal G2, the UEQ was carried out in a complementary manner to evaluate user experience. The mean values of the answers to the questions and the methodology already established for the UEQ were used, with its data analysis tool.[13] According to UEQ result interpretation [22], our results indicate an average around 2.0 for "Pragmatic Quality"' and average around 2.5 for "Hedonic Quality", as it can be seen in Table 10. Figure 15 shows the benchmark achieved, which classifies the user experience as excellent according to the evaluation requirements of UEQ creators [16], [22].

3) G3 - ANALYSIS OF RESULTS—AUTHORING EFFORT FOR EXPERT AND NOVICE USERS

Questions Q1, Q2, Q3 and Q4 were used for analyzing goal G3, as presented in Table 7. Participants' answers to those questions are shown in Table 11. We can notice that all

[13]https://www.ueq-online.org/Material/Short_UEQ_Data_Analysis_Tool.xlsx

TABLE 10. UEQ average results.

UEQ Quality	Negative	Positive	Aver. Sc.	Stand. Deviat.
Pragmatic	Obstructive	Supportive	+1.8	0.8
	Complicated	Easy	+1.5	1.2
	Inefficient	Efficient	+2.4	0.5
	Confusing	Clear	+2.0	1.1
Hedonic	Boring	Exciting	+2.5	0.5
	Not interesting	Interesting	+2.6	0.5
	Conventional	Inventive	+2.6	0.3
	Usual	Leading edge	+2.5	0.3

TABLE 11. G3 question answers.

Participants	Q1	Q2	Q3	Q4 (min)
1	Yes	Yes	Yes	14
2	Yes	No	No	6
3	Yes	Yes	No	7
4	Yes	Yes	Yes	6
5	Yes	No	Yes	7
6	Yes	Yes	Yes	5
7	Yes	No	No	10
8	Yes	No	No	4
9	Yes	No	Yes	9
10	Yes	Yes	Yes	8

TABLE 12. VR expert and novice users time to complete the experiment.

VR User	Time (minutes)	
	Average	Standard deviation
Expert	7.27	4.83
Novice	6.54	4.11

TABLE 13. Multimedia authoring tool expert and novice users time to complete the experiment.

Authoring Tool User	Time (minutes)	
	Average	Standard deviation
Expert	8.90	4.84
Novice	4.90	3.77

participants have finished the proposed tasks (Q1); 50% of the participants had used virtual reality technology before and the other 50% had not (Q2); 60% of the participants had previously used authoring tools and the other 40% had not (Q3); and the average time taken to complete the tasks with AMUSEVR (Q4) was 7.6 minutes with a standard deviation of 2.87 minutes for the whole group. The time taken for distinct groups of VR expert and novice users is presented in Table 12. Table 13 presents the time taken considering expert and novice users in multimedia authoring tools.

4) G4 - ANALYSIS OF RESULTS—QUANTITATIVE COMPARISON

In regards to achieving goal G4, the testing process was described in Section V-C. Figure 16 shows the comparison of number of clicks and lines required for diverse tasks T1-T6. We can see that, as the complexity of the task increases from T1 to T6, it is clear that there is an almost linearly progressive

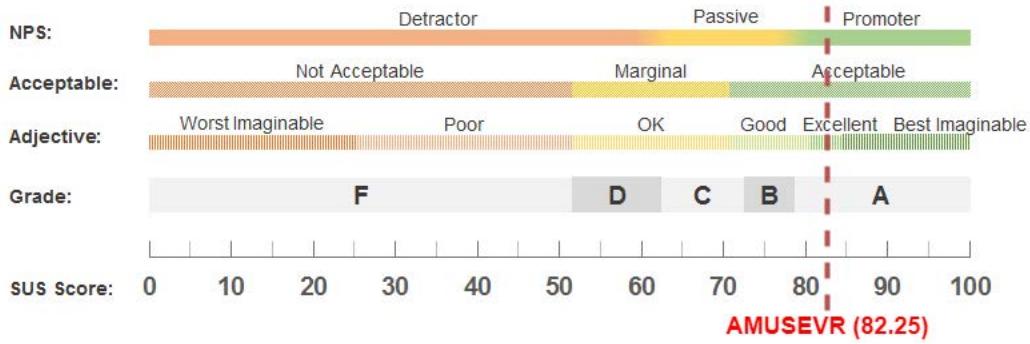

FIGURE 14. Acceptability and adjective scales. (SUS score.)

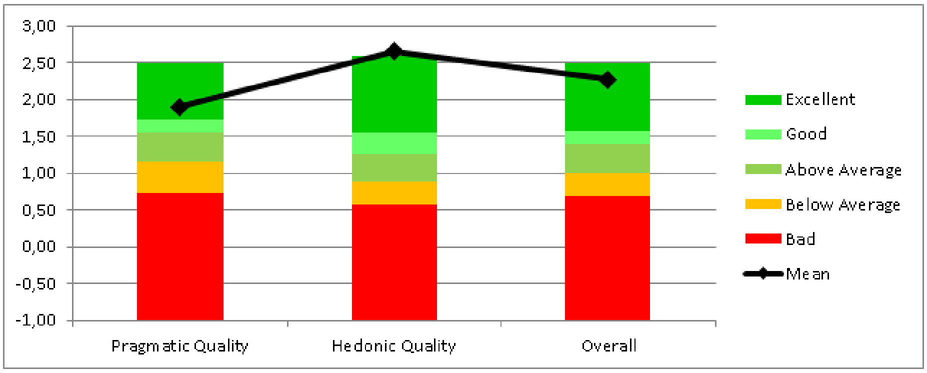

FIGURE 15. UEQ benchmark.

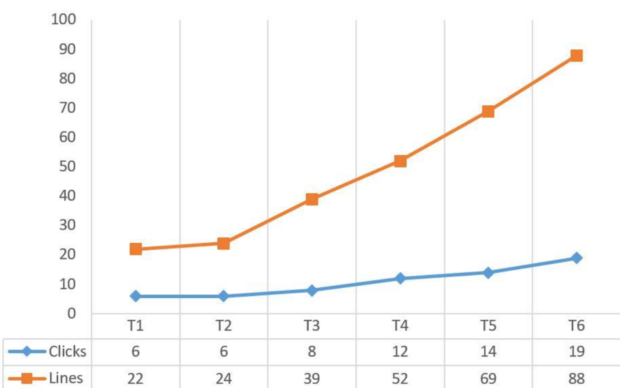

FIGURE 16. AMUSEVR clicks versus MultiSEL lines.

distance in relation to the number of AMUSEVR clicks (blue) and number of MultiSEL lines (orange) in Figure 16.

E. DISCUSSION AND LIMITATIONS

We recognize the small number of users in the authoring experiment as a current limitation of our research, as due to COVID-19 restrictions, we were not able to recruit a larger sample. Nonetheless, our exploratory work has resulted in some very promising results. We also recognize that the evaluation to analyze goal G4 is limited, but unfortunately we found no other alternative authoring method or tool available to compare the AMUSEVR authoring effort with. Therefore, we can conclude that:

- G1 goal was met as the SUS score of 82.25 points shows that usability of the AMUSEVR tool was excelent from the user's perspective as shown in Figure 14;
- G2 goal was met by the UEQ results, which indicated that user experience with the AMUSEVR tool was excelent considering UEQ benchmark as shown in Figure 15;
- Considering the G3 goal, we conclude that AMUSEVR can be used by novice users as they were able to create a 360° interactive experience. On the other hand, when comparing expert and novice users, our results were different than we expected, as participants with more experience with VR and authoring tools took in average more time to complete the authoring tasks. We believe that expert participants were more interested in exploring AMUSEVR facilities and knowing the immersive environment in more detail than novice users. Therefore, that explains why they took more time using the authoring system.
- G4 goal was also met considering our test results, as AMUSEVR proved to cause less authoring effort than MultiSEL for creating different 360 interactive experiences.

Finally, we also asked users for free comments and suggestions about AMUSEVR. We received suggestions considering:

- User Interface (UI) aspects - Some improvements to the UI were suggested by users, such as enlarging some fonts and message pop-ups which alert the user to important actions, such as delete functions, for example. One example of user comment about UI was "…Some menu itens were small and the pointing was difficult.";
- Positive feedback - We got compliments on immersion, the leading-edge factor, and satisfaction of use. One example of user compliment was "The system is very interesting to use. You can have fun while creating content. I enjoyed it a lot, I'm even thinking about studying similar topics in the future. I loved the innovation! I felt like "The Tech Guy" in a sci-fi movie.";
- Portability aspects - suggestions to port AMUSEVR to other more readily available VR platforms, such as those similar to Google Cardboard.[14] The goal is to provide easier access to the tool using smartphones. One example of user comment about portability was "…I would like to use this system with my smartphone…".

VI. CONCLUSION

This paper investigated the use of immersive authoring environments for creating 360° interactive experiences. We presented an immersive approach for authoring this type of applications in a VR environment. The approach was implemented in an authoring system named AMUSEVR, which provides author immersion in a VR environment, creative freedom and the availability of using a wide variety of media object and sensory effect types.

AMUSEVR was evaluated with the GQM approach considering user and quantitative experiments. Usability experiments reported a viable user experience, where all users completed their tasks of creating a 360° interactive mulsemedia presentation for a tourism application. Quantitative experiments showed that the authoring effort using AMUSEVR is smaller than the effort using a declarative authoring approach with MultiSEL.

AMUSEVR is extensible and designed to receive updates with new types of media with minimal effort. In addition, we are planning to provide support for other brands of VR headsets, such as Oculus Quest 2 (now called Meta Quest 2) and Google Cardboard, along with a standalone Viewer Mode. As future work, we will implement physical rendering of sensory effects to provide fully 360° multisensory experiences. We also expect to run further user experiments to evaluate discomfort when using the system. We also expect to run further user experiments involving MultiSEL XML format for authoring and exchanging multisensory applications. We are particularly interested in the education domain, where teachers could use AMUSEVR for creating 360° interactive applications for their students. To the end, we would like to investigate how multisensory experiences could be used for improving education across different levels - and will form the thrust of our future research efforts.

[14]https://arvr.google.com/cardboard/

REFERENCES

[1] R. S. D. Abreu, D. Mattos, J. D. Santos, G. Ghinea, and D. C. Muchaluat-Saade, "Toward content-driven intelligent authoring of mulsemedia applications," *IEEE MultimediaMag.*, vol. 28, no. 1, pp. 7–16, Jan. 2021.
[2] V. R. Basili, "Software modeling and measurement: The goal/question," Metric Paradigm, Chicago, IL, USA, Tech. Rep. UMIACS-TR-92-96, 1992.
[3] J. Brooke, "SUS: A 'quick and dirty' usability scale," *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1996.
[4] H. Coelho, M. Melo, L. Barbosa, J. Martins, M. Sergio, and M. Bessa, "Immersive edition of multisensory 360 videos," in *Advances in Intelligent Systems and Computing*, vol. 746. Cham, Switzerland: Springer Verlag, 2018, pp. 309–318.
[5] H. Coelho, M. Melo, L. Barbosa, J. Martins, M. S. Teixeira, and M. Bessa, "Authoring tools for creating 360 multisensory videos—Evaluation of different interfaces," *Expert Syst.*, vol. 38, no. 5, Aug. 2021, doi: 10.1111/exsy.12418.
[6] H. Coelho, M. Melo, J. Martins, and M. Bessa, "Collaborative immersive authoring tool for real-time creation of multisensory VR experiences," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 19473–19493, Jul. 2019.
[7] I.-S. Comşa, E. B. Saleme, A. Covaci, G. M. Assres, R. Trestian, C. A. S. Santos, and G. Ghinea, "Do i smell coffee? The tale of a 360° mulsemedia experience," *IEEE MultimediaMag.*, vol. 27, no. 1, pp. 27–36, 2019.
[8] A. Covaci, G. Ghinea, C.-H. Lin, S.-H. Huang, and J.-L. Shih, "Multisensory games-based learning-lessons learnt from olfactory enhancement of a digital board game," *Multimedia Tools Appl.*, vol. 77, no. 16, pp. 21245–21263, 2018.
[9] D. P. D. Mattos and M.-S. D. C. Gheorghita, "Beyond multimedia authoring," *ACM Comput. Surveys (CSUR)*, vol. 54, no. 7, pp. 1–31, Jul. 2021, doi: 10.1145/3464422.
[10] D. P. D. Mattos, D. C. Muchaluat-Saade, and G. Ghinea, "An approach for authoring mulsemedia documents based on events," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2020, pp. 273–277.
[11] H. Fassold and B. Takacs, "Towards automatic cinematography and annotation for 360° video," in *Proc. ACM Int. Conf. Interact. Experiences TV Online Video*, Jun. 2019, pp. 157–166.
[12] R. Horst, S. Diez, and R. Dörner, "Highlighting techniques for 360° video virtual reality and their immersive authoring," in *Proc. Int. Symp. Vis. Comput.* (Lecture Notes in Computer Science), vol. 11844. Cham, Switzerland: Springer, Oct. 2019, pp. 515–526.
[13] R. Horst, R. Naraghi-Taghi-Off, L. Rau, and R. Dorner, "Bite-sized virtual reality learning applications: A pattern-based immersive authoring environment," *J. Universal Comput. Sci.*, vol. 26, no. 8, pp. 947–971, Aug. 2020.
[14] W. Hwang and G. Salvendy, "Number of people required for usability evaluation: The 10±2 rule," *Commun. ACM*, vol. 53, no. 5, pp. 130–133, May 2010.
[15] M. Josué, R. Abreu, F. Barreto, D. Mattos, G. Amorim, J. D. Santos, and D. Muchaluat-Saade, "Modeling sensory effects as first-class entities in multimedia applications," in *Proc. 9th ACM Multimedia Syst. Conf.*, Jun. 2018, pp. 225–236.
[16] B. Laugwitz, T. Held, and M. Schrepp, "Construction and evaluation of a user experience questionnaire," in *Proc. HCI Usability Educ. Work, 4th Symp. Workgroup Hum.-Comput. Interact. Usability Eng. Austrian Comput. Soc.*, vol. 5298, 2008, pp. 63–76.
[17] R. Likert, "A method of constructing an attitude scale," in *Scaling: A Sourcebook for Behavioral Scientists*. London, U.K.: Taylor & Francis, 1974, pp. 233–243.
[18] P. R. C. Mendes, A. L. V. Guedes, D. D. S. Moraes, R. G. A. Azevedo, and S. Colcher, "An authoring model for interactive 360 videos," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, Jul. 2020, pp. 2–7.
[19] F. M. De Farias, D. P. De Mattos, and D. C. Muchaluat-Saade, "AMUSEVR: A virtual reality authoring environment for immersive mulsemedia applications," in *Proc. 1st Workshop Multisensory Experiences (SensoryX)*, Jun. 2021, pp. 1–10.

[20] N.-S. Park and M. Han, "Sensory effect authoring tool for virtual reality simulation," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2018, pp. 1497–1499.
[21] E. B. Saleme, C. A. Santos, and G. Ghinea, "A mulsemedia framework for delivering sensory effects to heterogeneous systems," *Multimedia Syst.*, vol. 25, no. 4, pp. 421–447, 2019, doi: 10.1007/s00530-019-00618-8.
[22] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Construction of a benchmark for the user experience questionnaire (UEQ)," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 4, no. 4, p. 40, 2017.
[23] M. Slater and M. V. Sanchez-Vives, "Enhancing our lives with immersive virtual reality," *Frontiers Robot. AI*, vol. 3, pp. 13–18, Dec. 2016.

FLÁVIO MIRANDA DE FARIAS received the bachelor's degree in information systems from the Federal University of Acre (UFAC), Brazil, the master's degree in computer science from the Federal University of Rio Grande do Sul (UFRGS), and the Ph.D. degree in computer science from Fluminense Federal University (UFF). He is currently a Professor at the Federal Institute of Science and Technology of Acre–IFAC. He is also a Researcher at the MidiaCom Laboratory, focusing on multimedia in virtual reality; the Amazonian Music Research Center-NAP, focusing on ubiquitous musical computing and musical creativity and researcher; and a the Informatics Group for research in computing at IFAC with a focus on educational systems.

DOUGLAS PAULO DE MATTOS received the master's degree in multimedia applications for digital TV and Web and the Ph.D. degree in mulsemedia applications from Fluminense Federal University (UFF), Niterói, Brazil. He was a Visiting Researcher in mulsemedia area at Brunel University London. His work focuses on an approach for authoring mulsemedia applications based on events. He proposed a mulsemedia conceptual model, an XML-based language, and a graphical tool for enhancing mulsemedia authoring.

GHEORGHITA GHINEA (Member, IEEE) received the first B.Sc. degree in computer science, the second B.Sc. degree (Hons.) in mathematics, and the M.Sc. degree in computer science from the University of the Witwatersrand, Johannesburg, South Africa, in 1993, 1994, and 1996, respectively, and the Ph.D. degree in computer science from the University of Reading, U.K., in 2000. He is currently a Professor of mulsemedia computing with the Department of Computer Science, Brunel University, London, U.K. His research interests include confluence of computer science, media and psychology, and focusing on building adaptable cross-layer end-to end communication systems incorporating user multisensorial and perceptual requirements.

DÉBORA CHRISTINA MUCHALUAT-SAADE received the Ph.D. degree in computer science from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), in 2003. She is currently a Full Professor at the Department of Computer Science, Fluminense Federal University (UFF), Niterói, Brazil. She is also one of the Founder and the Head of the MídiaCom Research Laboratory (www.midiacom.uff.br). She has contributed to the design and development of the Ginga-NCL middleware, used in the Brazilian Digital TV Standard (ABNT NBR 15606-2) and IPTV services (ITU-T H.761). Her main research interests include multimedia systems, computer networks, the IoT, smart grids, and e-health.