Exploring an AR-based User Interface for Authoring Multimedia Presentations

Paulo Renato Conceição Mendes
Federal University of Maranhão
São Luís, Brazil
mendesdpn@gmail.com

Roberto Gerson de Albuquerque Azevedo
PUC-Rio
Rio de Janeiro, Brazil
razevedo@inf.puc-rio.br

Ruy Guilherme Silva Gomes de Oliveira
Federal Institute of Maranhão
São Luís, Brazil
ruy.oliveira@ifma.edu.br

Carlos de Salles Soares Neto
Federal University of Maranhão
São Luís, Brazil
csalles@deinf.ufma.br

ABSTRACT

This paper describes the BumbAR approach for composing multimedia presentations and evaluates it through a qualitative study based on the Technology Acceptance Model (TAM). The BumbAR proposal is based on the event-condition-action model of Nested Context Model (NCM) and explores the use of augmented reality and real-world objects (markers) as an innovative user interface to specify the behavior and relationships between the media objects in a presentation. The qualitative study aimed at measuring the users' attitude towards using BumbAR and an augmented reality environment for authoring multimedia presentations. The results show that the participants found the BumbAR approach both useful and easy-to-use, while most of them (66,67%) found the system more convenient than traditional desktop-based authoring tools.

CCS CONCEPTS

• Information systems → Multimedia content creation; **• Human-centered computing** → *Mixed / augmented reality*; *User interface design*;

KEYWORDS

Augmented Reality, Multimedia, Authoring Tool, User Interface

ACM Reference Format:
Paulo Renato Conceição Mendes, Roberto Gerson de Albuquerque Azevedo, Ruy Guilherme Silva Gomes de Oliveira, and Carlos de Salles Soares Neto. 2018. Exploring an AR-based User Interface for Authoring Multimedia Presentations. In *DocEng '18: ACM Symposium on Document Engineering 2018, August 28–31, 2018, Halifax, NS, Canada.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3209280.3209534

1 INTRODUCTION

Multimedia authoring processes support the development of multimedia products (such as interactive movies and presentations) by integrating media objects of different types (e.g., pictures, sounds, and videos) in a synchronized and meaningful way. Indeed, even though the value of a multimedia content is greatly dependent on the quality of the individual media objects that compose it, it is only with an appropriate composition of all the media elements into a well-suited presentation that users will actually enjoy them and have a good quality of experience [28]. A common way of representing interactive multimedia presentations is through *multimedia documents*, such as those represented in HTML [13], SMIL [29], and NCL [34].

On the one hand, most of the state-of-the-art multimedia document models favor a clear separation of the specification of the media objects that compose the multimedia presentation and the synchronism specification —which defines the temporal inter-dependence between the media objects— among these media objects. On the other hand, there are also some differences in the authoring metaphors for synchronism specification currently in use by the multimedia document models. Some of them explicitly specified a first-class entity for the synchronism specification, called links (e.g., in NCL and HTML). In other cases, media objects are organized into hierarchical structures, called temporal compositions, that implicitly defines the synchronization of media objects (e.g., the *par* and *seq* containers in SMIL). Also, some of these document models support other types of compositions (e.g. the *context* and *div* elements in NCL and HTML, respectively) as a way to ease the task of authoring and structure complex documents. In any case, the synchronism specification is one of the most time-consuming tasks in the multimedia authoring process.

Although the concepts used in the aforementioned document models are usually easy to learn, their textual representation is not easy to be handled by non-programmers, who are usually more willing to use a visual representation of the model. Multimedia authoring tools are software that support the design and implementation of multimedia presentations aiming at simplifying the authoring process and allowing non-programmers to create such presentations. Indeed, the development of a multimedia authoring tool that simplifies the authoring process may increase user's engagement in developing multimedia presentations, opening new possibilities to creative people draft and create new multimedia experiences. Since the most common way to interact with a computer today is still through a GUI (Graphical User Interface) following a

DocEng '18, August 28–31, 2018, Halifax, NS, Canada

ACM ISBN 978-1-4503-5769-2/18/08…$15.00
https://doi.org/10.1145/3209280.3209534

WIMP (windows, icon, menus, pointers, typically a mouse), most of the current multimedia authoring tools use such paradigm.

Different from previous works, this work proposes the BumbAR approach for composing multimedia presentations. Instead of using a traditional WIMP/GUI approach, BumbAR explores the use of augmented reality (AR) as an innovative user interface for the authoring process of multimedia presentations. AR is defined by Carmigniani et al. [7] as "a real-time direct or indirect view of a physical real-world environment that has been enhanced/augmented by adding virtual computer-generated information to it." Indeed, in recent years, AR has been used in a diverse range of research fields such as games [24] and education [41]. In addition, the use of AR has demonstrated to increase users' enjoyment in performing tasks [5, 15, 41]. In the proposed approach, AR-based interaction techniques are used to create the temporal relationships between media objects that are part of a multimedia presentation.

To detail the BumbAR approach, the rest of the paper is organized as follows. Section 2 presents some related work, discussing how they compare with BumbAR. Section 3 details the BumbAR concepts, its interface, and a prototype implementation of the BumbAR approach. Section 4 presents a subjective evaluation of BumbAR using the Technology Acceptance Model (TAM). The main goal of the evaluation is to gather evidence whether or not the potential users will well receive the proposed interaction model. Finally, Section 5 discusses our final considerations and point out future work.

2 RELATED WORK

The works related to this paper and that BumbAR builds on can be broadly classified in traditional authoring tools for multimedia presentations, AR-based authoring tools, and tangible user interfaces. We briefly discuss some of them in what follows.

Authoring Tools for Multimedia Presentations

Similar to traditional authoring tools for multimedia presentations —e.g. GRiNS [6], LimSee3 [11], Adobe Flash/Adobe Animate [2], NEXT [26], and NCL Composer [4]— the primary goal of BumbAR is to allow users without programming skills to create multimedia content.

Different from those tools, BumbAR does not use the traditional GUI/Desktop-based interface, but explores the use of AR and real-world objects (markers) to specify the behavior and relationships between the media objects in the presentation. By doing so, we aim at exploring new metaphors for authoring multimedia content, which can also engage and entertain the user at the same time.

In particular, even though using different interaction techniques, the BumbAR underlying conceptual model is based on NCM (Nested Context Model) [35], which is also followed by the GUI-based approach of NCL Composer [4]. BumbAR also exports the authored content to NCL (Nested Context Language) [34], allowing the final application to run on digital TV sets or web browsers (e.g., using WebNCL [25] or NCL4Web [32]).

AR-Based authoring tools

Although today most of the AR content is still created by system developers, some authoring tools have been proposed to create augmented reality content. They usually provide complex tools for designers to model 3D objects and associate them with physical objects and markers. Most of them, however, still follows a GUI/desktop-based approach (e.g. [19, 22–24, 30, 33]). Others, more related to this work, provide an immersive or mixed GUI/immersive approach for creating AR content.

For instance, Langlotz et al. [21] present a system that targets an inexperienced audience to create AR content based on mobile devices. By directly interacting with the mobile device using a touchscreen, the system allows authors to generate registered 3D primitives, modify registered 3D primitives, apply textures to the registered 3D objects, and generate registered 2D objects (which can serve as annotations). While the proposal is interesting, there is no evaluation of the target audience. Thus it is hard to know how good it is from the end users' point of view.

Vera and Sánchez [39] and Vera et al. [40] also present an on-going work on the development of an AR-based authoring tool, named SituAR, for creating in-situ AR content. The main goal of the approach is to enhance points of interest (POIs) for smart cities. As features envisioned to SituAR the authors mention: creating and visualizing annotations, viewing a map, and expanding nearby annotations. As previously mentioned, the proposed approach is still an on-going project, and the cited papers do not provide enough details for a more detailed comparison.

Different from the authoring tools for AR content mentioned above, the content created by our proposal is not itself AR content. Our main goal is to explore AR as an innovative user interface for creating traditional multimedia presentations (i.e., composed of images, texts, videos, etc.).

Closer to our work, Shin et al. [31] also aim at creating an AR-based authoring tool for a non-AR content. The AR-based tool proposed by Shin et al. aims at creating storyboards. By using the proposed *item block* (AR markers), it is possible to specify the characters, background information, stage properties, action, and facial expressions. By placing character-based items in the camera view, it is possible to determine the pose of the corresponding 3D models in the scene. Dynamic components (e.g., actions and facial expressions) can be placed anywhere within the camera view. Although of different nature (we are mainly interested in the temporal behavior of a multimedia presentation whereas Shin et al. are interested in storyboards for 3D scenes) the linking phase (detailed in Section 3) proposed in this paper is related to the dynamic components proposed by Shin et al.. One of the drawbacks of Shin et al. [31], however, is that there is no user-related evaluation work.

Finally, although we are mainly interested in specifying the behavior of traditional multimedia presentations, some of the concepts and interaction techniques used in this paper can also be useful for other AR-based authoring tools (to both AR and non-AR content). In special, the event-based NCM concepts, in which BumbAR is based on, could also be helpful for the specification of the behavior of AR content. A more in-depth exploration of such a possibility, however, is beyond the scope of this paper, and we left it as future work.

Tangible User Interfaces

Another area related to our research is the *Tangible User Interfaces* [17]. The main goal of this area is to provide tangible interfaces so that users can use traditional physical objects as interactive elements with the computer. Indeed, AR technologies can be used as a basis for "easing" the development of many tangible user interfaces, and there are plenty of examples of integrating both. For instance, Ha et al. [14] provide a mixed GUI and TUI (Tangible User Interface) to create AR content. In the work, the authors provide an authoring tool for eBooks enhanced with AR features. The tangible interface is composed of a cubical box attachable to a mouse input and with multiple markers printed on the box. Similarly, as will be clear during the remainder of the paper, we are also providing a tangible (AR marker-based) interface in which authors create content by interacting with physical objects. But, differently, our main focus is on an interface for creating multimedia presentations. To the best of our knowledge, this is the first work exploring such approach for creating traditional multimedia presentations.

3 THE BUMBAR APPROACH

The main contribution of this paper is the proposal and evaluation of an AR-based authoring process for creating multimedia presentations. The proposed authoring process is implemented in the BumbAR authoring tool, which allows authors to create the multimedia content, preview its final result, and export it to be presented on web browsers or digital TV sets.

In what follows, we discuss the proposed authoring process (Section 3.1) and its implementation in the BumbAR tool (Section 3.2).

3.1 Authoring Process

The authoring process of multimedia presentations using BumbAR consists of two distinct phases: *media configuration* and *linking*.

The media configuration phase

In the media configuration phase, the authors must define the media objects that are part of the presentation and its properties. Also, each media object is associated with a unique custom AR marker, which is dynamically chosen by the author.

A *media object* is identified by its *name* and its *URL* (Uniform Resource Locator), which points to its media content file. Common media object types supported by BumbAR are the following: text, image, video, and audio.

Moreover, a media object contains *properties*, which define how it is presented. In the BumbAR approach, the supported media object properties are the common ones of 2D multimedia presentations, such as the media object position in the screen (x and y), its size (width and height), transparency, and volume.

With regards to the spatial properties, the position, and size of the media objects are measured in a 2D Cartesian coordinate system varying from 0 to 1, as shown in Figure 1 The origin of the system $(0, 0)$ is at the left-bottom of the screen, and $(1, 1)$ is at the right-top of the screen. For instance, if the height of a media object $m1$ is 0.5, that means its height is 50% of the screen's height. Also, the zIndex property defines the overlap order of the media object. A media object with greater zIndex value is in front of a media object with a lower zIndex value.

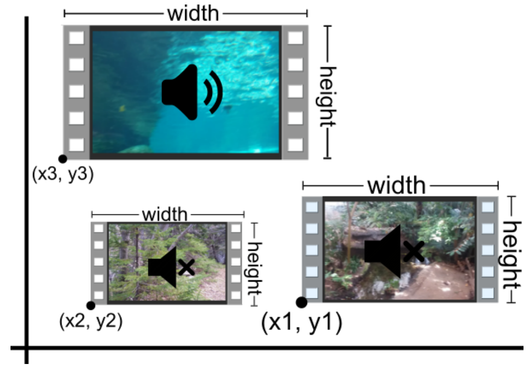


Figure 1: Media configuration

The linking phase

In the *linking phase*, the authors define how the multimedia presentation evolves over time, and how it reacts to users' interaction. To do so, they use an AR-based user interface in which the previously configured AR markers for the media objects are dynamically combined with behavior-related AR markers. Such a combination is mainly performed by physically dragging and colliding media objects and behavior-related AR markers in the real world, which results in creating relationships between the media objects, as detailed below. In addition, to highlight the different AR marker and provide feedback for the users, each AR marker is rendered in the AR view as a cube-based 3D object, which we call an *AR block*.

The relationships defined between the media objects follows the Nested Context Model (NCM) [35]. By dragging and colliding the media objects and behavior-related AR blocks, the authors can create *link*s (more specifically, *causal link*s). A *causal link* determines that when a condition (e.g., an event is triggered) is satisfied, one or more actions are executed. Each supported condition and action is represented by a different AR marker and a different AR block.

Table 1 shows the templates used for the supported AR block types. (The symbols used for individual AR markers and blocks, however, vary based on the specific condition, action, or media object. For instance, although following the same template, an *onEnd* condition AR block has a different symbol and underlying name than the *onBegin* AR block shown on the first line of Table 1.) Table 2 shows the individual conditions and actions currently supported by BumbAR. The icons used for each of them are based on the visual notation created by Laiola Guimarães et al. [20].

To create a *causal link* between two media objects, the authors must:

(1) mark the *media object block*s that are part of the condition of the link, creating a *media-condition block*;
(2) mark the *media object block*s that are part of the actions of the link, creating a *media-action block*; and
(3) associate the *media-condition block*s with the *media-action blocks*, creating the *causal link*.

The step (1) above is performed by dragging and colliding a *condition block* with a *media block*. By doing so, an icon representing the used condition is attached to the right-top corner of the *media block*, and it becomes a *media-condition block*. A *media-condition block* is

Table 1: The BumbAR augmented reality block templates

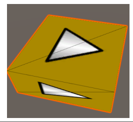
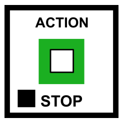
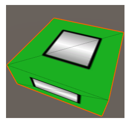
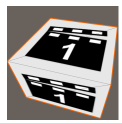
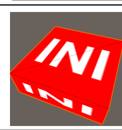
Block type	Usage	AR marker	AR block
Condition	It is used to represent conditions.		
Action	It is used to represent actions.		
Media	It is used to represent a media object.		
Initial	It is used to define the media object that starts when the multimedia presentation starts.		

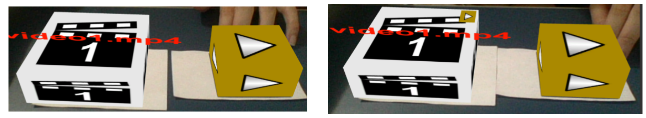
Table 2: Conditions and actions used in BumbAR

Name	Role	Meaning
onBegin	condition	Triggered when the media object starts.
onEnd	condition	Triggered when the the media object stops or ends.
onAbort	condition	Triggered when the media object is aborted.
onPause	condition	Triggered when the media object is pause.
onResume	condition	Triggered when the media object continues after being paused.
start	action	Starts the media object.
stop	action	Stops the media object.
abort	action	Aborts the media object.
pause	action	Pauses the media object.
resume	action	Resumes the media object (makes it continues) after being paused.

the same *media-block* that was used to create it but marked with a condition (to be part of a *link*). Figure 2 shows this process (see the icon that appears at right-top of the *media block*).

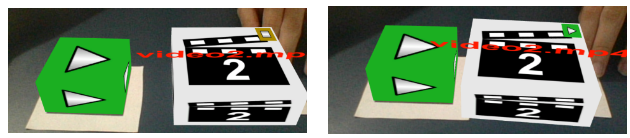
The step (2), i.e., the creation of a *media-action block*, follows the same procedure of the creation of the *media-condition*, but the author uses an *action block* colliding with a *media object block*.

After having a *media-condition* or *media-action* block, it is also possible to change it easily. To do so, the author just needs to collide another *condition* or another *action* block with the *media-condition* or *media-action* block. This way, the same media object can have different roles in different links. Figure 3 shows a *media-condition block* with the condition *OnEnd* being changed into a *media-action block* with the action *Start*.



(a) Before collision **(b) After collision**

Figure 2: *Media block* becoming a *media-condition block*



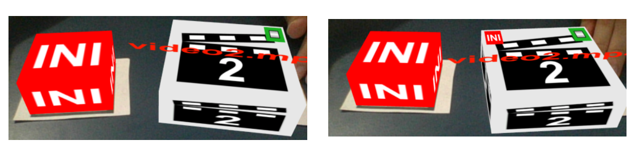
(a) Before collision **(b) After collision**

Figure 3: *Media-condition block* becoming a *media-action block*

Finally, the step (3) above, i.e., the creation of the *link* itself, is performed by colliding a *media-condition block* with a *media-action block*. Figure 4 exemplifies this process by the creation of a link *onBegin media1 stop media2*.

Figure 4: *Link* being created

Besides the above blocks, BumbAR also provides the *initial block*. That block is used to mark a *media block* informing that the associated media object must be started when the multimedia presentation starts. Similar to the other cases, an author marks a *media block* to start when the application starts, but dragging and colliding a *media block* with an *initial block*. By doing so, an icon with the symbol of the *initial block* is attached to the left-top corner of the *media block*. Figure 5 exemplifies this process.



(a) Before collision **(b) After collision**

Figure 5: *Media-action block* collides with *initial block*

3.2 Prototype Implementation

The above authoring process is concretized in the BumbAR authoring tool. The BumbAR tool is implemented as an asset [1] in the Unity game engine [36]. By reusing the Unity engine's user interface and components, BumbAR also provides an environment for both composing multimedia presentations using augmented reality and playing them.

Indeed, the BumbAR tool has two modes: *editing mode* and *presentation mode*. The former is the mode in which multimedia presentations are composed using the authoring process previously described, and the latter is the mode in which the developed presentations are presented. Figure 6 highlights the two modes and the integration with the Unity environment.

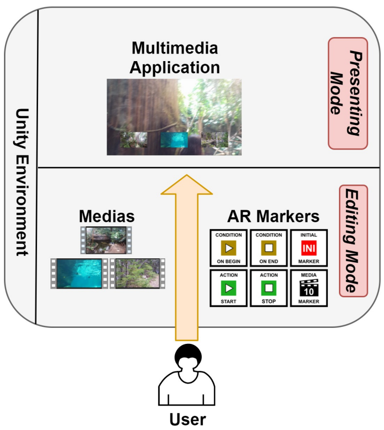


Figure 6: The BumbAR tool architecture.

The BumbAR tool was developed using C#, one of the languages supported by Unity. The Vuforia package for Unity [16] is used for the recognition of AR markers. The AR blocks 3D models, the implementation of collisions, and the media playback control uses Unity's native components. C# events and methods are used to implement the BumbAR *causal link* systems; *conditions* are implemented using C# events, and *actions* using its methods.

Currently, the AR markers discussed in the previous section are supported. Adding new markers (e.g. to support all the NCM events) is possible and relatively easy for one with knowledge of the Unity environment. Extra markers can be created by adding a new *database* of markers in the Vuforia Developer Portal [2]. A *database* consists of images uploaded and converted into markers in the Vuforia Developer Portal. Then, a new database can be added in the BumbAR tool by downloading and importing it as a package on Unity. After that, a new marker is added by selecting the desired *block* (for instance, a *media block*) and changing its *database* (selecting the new one that was created) and the desired marker. This process only describes the addition of a new marker. In order to implement new *actions* and *conditions*, it is necessary to implement their behavior using C# methods and events as cited before.

[1]In the Unity context, assets add extra content and functionalities to the engine. It may be easily reused as an component.
[2]https://developer.vuforia.com/target-manager

Besides presenting the created multimedia presentation, the BumbAR tool also supports exporting it to *Nested Context Language* (NCL). NCL is the standard language of the Brazilian Terrestrial Digital TV System (SBTVD-T) [1] and ITU-T H.761 recommendation from the International Telecommunications Union (ITU-T) [18]. There are open-source and commercial NCL players for both digital TV [12] and the Web [25, 32]. Thus, a multimedia presentation composed with the BumbAR tool can also be executed in those environments. Since BumbAR and NCL use similar concepts from the NCM model, such as media objects and links, exporting from the internal representation of BumbAR to NCL is straightforward.

4 EVALUATION

To evaluate the BumbAR approach, we designed a qualitative empirical study in which participants were asked to develop a multimedia presentation using the system. Afterwards, they were invited to fulfill a questionnaire with their perceptions about the authoring tool. The main goal of the study was to analyze the users' attitude towards the development of multimedia presentations using BumbAR, and, more generally, an AR-based user interface system for creating such presentations.

In the remainder of this section, we detail the methodology we used (Section 4.1), the participants' profiles (Section 4.2), and the results of the study (Section 4.3).

4.1 Methodology

The empirical study was conducted with one participant at a time and it was composed of 5 distinct phases:

- **Consent term presentation**: The consent term is presented, verbally emphasizing the goals of our research. The participant is invited to consent or not to participate in the research, signing the consent term in case of consenting with it.
- **Pre-experiment interview**: The goal of this phase is to elaborate the participant's profile, especially about his or her knowledge about multimedia presentations and augmented reality. A detailed characterization of the participants' profile is presented in Section 4.2.
- **BumbAR presentation**: The goal of this phase is to get the participant to know BumbAR essential concepts and how to compose multimedia presentations with it. To achieve that, the BumbAR tool and its authoring process are presented to the participant.
- **Tasks execution**: The participant is invited to read the scenario. After that, the multimedia presentation is played emphasizing the expected results.
- **Post-experiment interview**: The goal of this phase is to identify the participant's perceptions about the tasks, and the facilities and difficulties he or she had while using BumbAR. Some observations made by the researchers during the task were also discussed with the participants.

The tasks used in the *Task execution* phase were chosen regarding the completeness of the experiment. For that reason, we envisioned a multimedia presentation in which all the markers and links available in the current version of the BumbAR tool are used. A video describing the authoring process of this presentation can be seen in our auxiliary material[3]. Figure 7 shows the expected behavior of the multimedia presentation to be developed by the participants.

We divided the steps to compose the presentation into four tasks, which the participants could follow as a guideline. The chosen tasks were displayed on a monitor during the experiment, and they are the following:

- **Task 1**: configure for four video objects with maximum volume:
  - video 1: it has the URL *video1.mp4* and fills the full screen;
  - video 2: it has the URL *video2.mp4* and fills the left half of the screen;
  - video 3: it has the URL *video3.mp4* and fills the right half of the screen; and
  - video 4: it has the URL *video4.mp4* and fills the full screen.
- **Task 2**: Compose a multimedia presentation that plays video 1.
- **Task 3**: Extend the multimedia presentation developed in Task 2 to do the following: when video 1 ends, play two other video objects, video 2 and video 3, at the same time, knowing that video 3 starts when video 2 starts.
- **Task 4**: Extend the multimedia presentation developed in Task 3 to do the following: when video 2 ends, stops video 3 and plays video 4.

(a) Presenting video 1 once the application starts

(b) Presenting videos 2 and 3 simultaneously

(c) Presenting video 4 after videos 2 and 3 end

Figure 7: Expected artifact from experiment

To compose the presentation, the participants used an AR installation that consisted of one notebook with the BumbAR tool, one webcam and the AR markers used in the authoring tool. The markers were restricted to the *initial* marker, *condition* markers, *action* markers and four media object markers.

For the *Post-experiment interview* phase, we designed a questionnaire based on the Technology Acceptance Model (TAM) [8–10, 38], which is extensively used by works aiming to evaluate the acceptance of new technologies [3, 27, 41]. The TAM model proposes the conception of statements for technologies evaluation based on two concepts: *perceived usefulness (PU)* and *perceived ease-of-use (PEOU)*. Davis et al. [10] define PU as "the degree to which a person believes that using a particular system would enhance his or her job performance" and PEOU as "the degree to which a person believes that using a particular system would be free of effort."

Table 3 shows the TAM-based questionnaire we designed for the qualitative evaluation. It is composed of seven PU statements and seven PEOU statements and two open-ended questions for the participants to give opinions about the tool. Each of the statements had options based on a 7-point Likert scale of level of agreement [37], with the following options: 1 - Strongly disagree; 2 - Disagree; 3 - Somewhat disagree; 4 - Neither agree or disagree; 5 - Somewhat agree; 6 - Agree; and 7 - Strongly Agree.

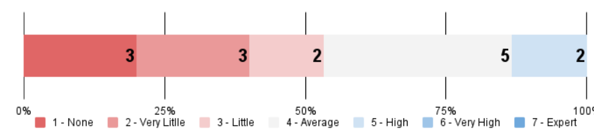
[3]The video can also be seen in the following web address https://youtu.be/GoLuOYLIJaQ

4.2 Participants

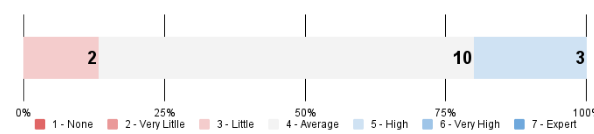
The empirical study involved undergraduate students in two courses: Computer Science and Design. This group was chosen because the development of multimedia applications is often present in the curriculum of these professionals. A total of 15 students with ages varying from 20 to 27 years participated in the study.

The participants' expertise level in the development of multimedia presentations and in augmented reality applications was also taken into consideration. They were measured on a seven-point Likert scale ranging from 1 (None) to 7 (Expert).

Figure 8 shows, in percentage and absolute values, the participants' expertise in both the development of multimedia presentations and augmented reality.



(a) Participants' expertise in the development of multimedia applications



(b) Participants' expertise in augmented reality

Figure 8: Participants' expertise

4.3 Results

There have been no technical problems during the experiment so that the participant could focus on the merits of the study. All of them completed the proposed tasks and, consequently, developed the requested multimedia presentation. The mean time taken by the participants to develop the multimedia presentation was 13 minutes and 49 seconds, with a standard deviation of 3 minutes and 17 seconds.

To find the internal consistency of the gathered data, we calculated the Cronbach's alpha coefficient for the two groups of statements measured on a Likert scale, *PU* and *PEOU*. The Cronbach's alpha for the PU statements is 0.7637, while for the PEOU statements is 0.8433. (A Cronbach's alpha greater than 0.7 usually

Table 3: The TAM-based questionnaire used in the evaluation

	Questionnaire Statements
PU1	In general, the presented concepts (media objects, links, conditions, and actions) are useful for specifying multimedia presentations.
PU2	In general, the development of multimedia presentations using Augmented Reality is useful.
PU3	In general, the process of developing multimedia presentations using Augmented Reality with the chosen markers is useful.
PU4	I consider that the system is more convenient than traditional ones.
PU5	I believe that the system makes the development of multimedia presentations fast.
PU6	I think that the system is useful.
PU7	In general, I believe the system is advantageous.
PEOU1	In general, the authoring process with augmented reality is simple and comprehensible.
PEOU2	In general, the authoring process with augmented reality using the chosen markers is simple and comprehensible.
PEOU3	In general, the adopted paradigm for syncing media objects(links, conditions, actions) is simple and comprehensible.
PEOU4	Learning to use the system was easy to me.
PEOU5	I think the system is easy to use.
PEOU6	I think the system makes the development of multimedia presentations easy.
PEOU7	I believe I could become skillful in developing multimedia presentations with the system.
Free1	Name positive aspects of the tool, if any.
Free2	Name negative aspects of the tool, if any.

indicates that the reliability of the statements is at a satisfactory level.)

Figure 9 shows the percentage of each statements’ answer to the *PU* and *PEOU* statements. All the statements had mostly positive answers. The statements *PU6*, which says “I think the system is useful” and *PEOU5* which says “I think the system is easy to use” have a significant level of agreement, having answers *5 - somewhat agree* or higher. These results reveal that, in the participants’ opinion, the BumbAR proposal is both useful and easy to use.

Also, as previously mentioned, we took some notes of interesting informal commentaries made by the participants about the authoring process. Indeed, together, the informal commentaries and the open-ended questions have been proved to be a very interesting data, which is usually lost in a pure quantitative experiment. From the positive side, most of the participants mentioned the intuitiveness aspect of the BumbAR approach and that it makes the authoring process more enjoyable. From the negative side, most of them highlighted the necessity of showing the history of links already created, something that is currently not supported by the tool.

A particular participant said the following about the BumbAR approach: *“it is faster and less complicated than other methods I know”*. According to the results, most of the participants agreed with him/her: 86,67% of the participants believe that the system makes the development of multimedia presentations fast, while 66,67% found it more convenient than the traditional ones.

5 FINAL CONSIDERATIONS

In this work, we proposed and AR-Based approach for composing multimedia presentations. The proposed approach is based on the event-condition-action model of Nested Context Model (NCM). Moreover, we presented the BumbAR tool, which implements the BumbAR approach with an implementation based on Unity, allowing users to compose, preview, and generate an NCL document of their multimedia presentations. To evaluate our proposal, we designed a qualitative study based on the TAM model. The participants of the study found the approach both useful and easy-to-use.

Although our current implementation of the BumbAR tool discussed in Section 3.2 is capable of developing a large number multimedia presentations, it still does not fully implements all the conditions and actions present in the NCM model, such as the ones related to *selection* and *attribution*. Those types of conditions and actions allow the dynamic modification of properties of the *media objects*, such as their volume and transparency. The implementation of other conditions and actions is one of our future work.

In the BumbAR approach, we used NCM *causal links* to specify relationships and the temporal synchronism between media objects. As another future work, we also intend to explore other approaches for defining that synchronism. For instance, the Synchronized Multimedia Integration Language (SMIL)[29] approach proposes the synchronization between *media objects* based on two concepts: *par* and *seq*. Media objects in a *seq* container play sequentially: when one ends the next starts; whereas media in a *par* container play in parallel: they start and end at the same time. For many types of multimedia presentations, it is possible to ease its authoring process using such abstractions. Those additional concepts can be implemented by extending the current BumbAR approach for supporting the composition of complexes behaviors based on the existent ones.

Finally, even though the qualitative study based on the TAM model helped us reaching significant conclusions about the BumbAR approach, further analysis is still necessary for evaluating the

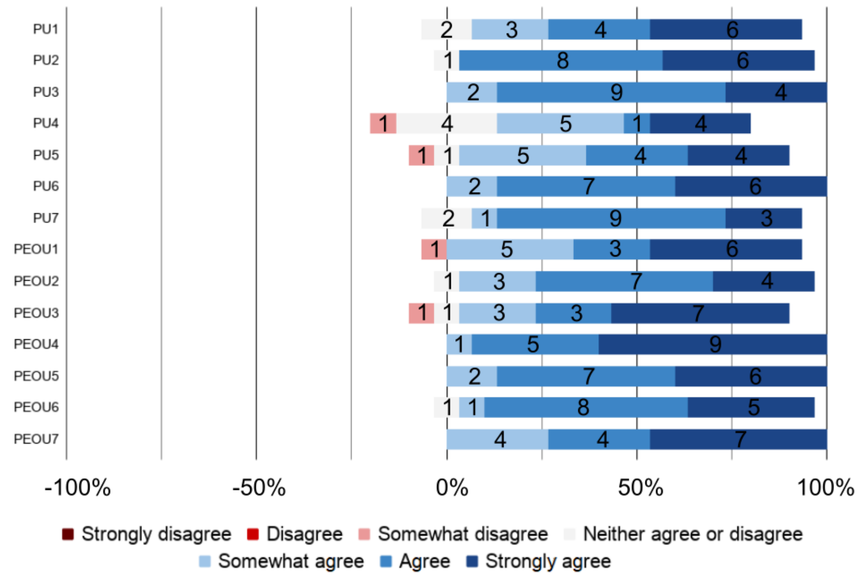


Figure 9: Participants' answers about BumbAR

real applicability of an AR-based user interface for authoring multimedia presentations. In special, it is necessary a more in-depth study explicitly comparing such approach with the WIMP-based authoring tools. Thus, as another future work, we intend to perform a quantitative evaluation comparing the time needed for performing the same tasks with both the BumbAR tool and desktop-based authoring tools.

REFERENCES

[1] ABNT. 2011. 15606-2, 2011. Digital Terrestrial Television-Data Coding and Transmission Specification for Digital Broadcasting-Part 2: Ginga-NCL for fixed and mobile receivers-XML application language for application coding.
[2] Adobe. 2018. Adobe Animate CC Website. https://www.adobe.com/products/animate.html Accessed: February, 01 2018.
[3] Khaled M Alraimi, Hangjung Zo, and Andrew P Ciganek. 2015. Understanding the MOOCs continuance: The role of openness and reputation. *Computers & Education* 80 (2015), 28–38.
[4] R. G. de A. Azevedo, E. C. Araújo, Bruno Lima, L. F. G. Soares, and M. F. Moreno. 2014. Composer: meeting non-functional aspects of hypermedia authoring environment. *Multimedia Tools and Applications* 70, 2 (May 2014), 1199–1228. https://doi.org/10.1007/s11042-012-1216-8
[5] Alexandru Balog and Costin Pribeanu. 2010. The role of perceived enjoyment in the students' acceptance of an augmented reality teaching platform: A structural equation modelling approach. *Studies in Informatics and Control* 19, 3 (2010), 319–330.
[6] Dick C.A. Bulterman, Lynda Hardman, Jack Jansen, K.Sjoerd Mullender, and Lloyd Rutledge. 1998. GRiNS: A GRaphical INterface for creating and playing SMIL documents. *Computer Networks and ISDN Systems* 30, 1-7 (April 1998), 519–529. https://doi.org/10.1016/S0169-7552(98)00128-7
[7] Julie Carmigniani, Borko Furht, Marco Anisetti, Paolo Ceravolo, Ernesto Damiani, and Misa Ivkovic. 2011. Augmented reality technologies, systems and applications. *Multimedia Tools and Applications* 51, 1 (Jan. 2011), 341–377. https://doi.org/10.1007/s11042-010-0660-6
[8] Fred D Davis. 1985. *A technology acceptance model for empirically testing new end-user information systems: Theory and results.* Ph.D. Dissertation. Massachusetts Institute of Technology.
[9] Fred D. Davis. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly* 13, 3 (Sept. 1989), 319. https://doi.org/10.2307/249008
[10] Fred D Davis, Richard P Bagozzi, and Paul R Warshaw. 1989. User acceptance of computer technology: a comparison of two theoretical models. *Management science* 35, 8 (1989), 982–1003.
[11] Romain Deltour and Cécile Roisin. 2006. The Limsee3 Multimedia Authoring Model. In *Proceedings of the 2006 ACM Symposium on Document Engineering (DocEng '06)*. ACM, New York, NY, USA, 173–175. https://doi.org/10.1145/1166160.1166203
[12] Luiz Gomes Soares, Marcio Moreno, Carlos Salles Soares Neto, and Marcelo Moreno. 2010. Ginga-NCL: Declarative middleware for multimedia IPTV services. *IEEE Communications Magazine* 48, 6 (June 2010), 74–81. https://doi.org/10.1109/MCOM.2010.5473867
[13] Ian S Graham. 1995. *The HTML sourcebook.* John Wiley & Sons, Inc.
[14] Taejin Ha, Woontack Woo, Youngho Lee, Junhun Lee, Jeha Ryu, Hankyun Choi, and Kwanheng Lee. 2010. ARtalet: Tangible User Interface Based Immersive Augmented Reality Authoring Tool for Digilog Book. IEEE, 40–43. https://doi.org/10.1109/ISUVR.2010.20
[15] Anne-Cecilie Haugstvedt and John Krogstie. 2012. Mobile augmented reality for cultural heritage: A technology acceptance study. In *Mixed and Augmented Reality (ISMAR), 2012 IEEE International Symposium on.* IEEE, 247–255.
[16] PTC Inc. 2018. Vuforia Augmented Reality SDK. https://www.vuforia.com/
[17] Hiroshi Ishii. 2008. The tangible user interface and its evolution. *Commun. ACM* 51, 6 (June 2008), 32. https://doi.org/10.1145/1349026.1349034
[18] ITU-T. 2009. H. 761, Nested Context Language (NCL) and Ginga-NCL for IPTV Services, Geneva, Apr. 2009.
[19] Hyung-Keun Jee, Sukhyun Lim, Jinyoung Youn, and Junsuk Lee. 2014. An augmented reality-based authoring tool for E-learning applications. *Multimedia Tools and Applications* 68, 2 (Jan. 2014), 225–235. https://doi.org/10.1007/s11042-011-0880-4
[20] Rodrigo Laiola Guimarães, Carlos de Salles Soares Neto, and Luiz Fernando Gomes Soares. 2008. A visual approach for modeling spatiotemporal relations. In *Proceedings of the eighth ACM symposium on Document engineering*. ACM, 285–288.
[21] Tobias Langlotz, Stefan Mooslechner, Stefanie Zollmann, Claus Degendorfer, Gerhard Reitmayr, and Dieter Schmalstieg. 2012. Sketching up the world: in situ authoring for mobile Augmented Reality. *Personal and Ubiquitous Computing* 16, 6 (Aug. 2012), 623–630. https://doi.org/10.1007/s00779-011-0430-0
[22] Moralejo Lucrecia, Sanz Cecilia, Pesado Patricia, and Baldassarri Sandra. 2013. AuthorAR: Authoring tool for building educational activities based on Augmented Reality. IEEE, 503–507. https://doi.org/10.1109/CTS.2013.6567277
[23] Blair MacIntyre, Maribeth Gandy, Steven Dow, and Jay David Bolter. 2004. DART: a toolkit for rapid design exploration of augmented reality experiences. ACM Press, 197. https://doi.org/10.1145/1029632.1029669
[24] Luís Fernando Maia, Carleandro Nolêto, Messias Lima, Cristiane Ferreira, Cláudia Marinho, Windson Viana, and Fernando Trinta. 2017. LAGARTO: A LocAtion based Games AuthoRing TOol enhanced with augmented reality features. *Entertainment Computing* 22 (July 2017), 3–13. https://doi.org/10.1016/j.entcom.2017.05.001

[25] Erick Lazaro Melo, Caio César Viel, Cesar Augusto Camillo Teixeira, Alexandre Coelho Rondon, Daniel de Paula Silva, Danilo Gasques Rodrigues, and Endril Capelli Silva. 2012. WebNCL: A Web-based Presentation Machine for Multimedia Documents. In *Proceedings of the 18th Brazilian Symposium on Multimedia and the Web (WebMedia '12)*. ACM, New York, NY, USA, 403–410. https://doi.org/10.1145/2382636.2382719

[26] Douglas Paulo de Mattos, Júlia Varanda da Silva, and Débora Christina Muchaluat-Saade. 2013. NEXT: graphical editor for authoring NCL documents supporting composite templates. ACM Press, 89. https://doi.org/10.1145/2465958.2465964

[27] Paul A Pavlou. 2003. Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International journal of electronic commerce* 7, 3 (2003), 101–134.

[28] Benoît Pellan and Cyril Concolato. 2009. Authoring of scalable multimedia documents. *Multimedia Tools and Applications* 43, 3 (01 Jul 2009), 225–252. https://doi.org/10.1007/s11042-009-0268-x

[29] Lloyd Rutledge. 2001. SMIL 2.0: XML for Web multimedia. *IEEE Internet Computing* 5, 5 (2001), 78–84.

[30] Hartmut Seichter, Julian Looser, and Mark Billinghurst. 2008. ComposAR: An intuitive tool for authoring AR applications. In *Proceedings of the 7th IEEE/ACM international symposium on mixed and augmented reality*. IEEE Computer Society, 177–178.

[31] Midieum Shin, Byung-soo Kim, and Jun Park. 2005. AR storyboard: an augmented reality based interactive storyboard authoring tool. In *Mixed and Augmented Reality, 2005. Proceedings. Fourth IEEE and ACM International Symposium on*. IEEE, 198–199.

[32] Esdras Caleb Oliveira Silva, Joel A. F. dos Santos, and Débora C. Muchaluat-Saade. 2013. NCL4WEB: Translating NCL Applications to HTML5 Web Pages. In *Proceedings of the 2013 ACM Symposium on Document Engineering (DocEng '13)*. ACM, New York, NY, USA, 253–262. https://doi.org/10.1145/2494266.2494273

[33] G. Sinem and Steven Feiner. 2003. Authoring 3D hypermedia for wearable augmented and virtual reality. In *2012 16th International Symposium on Wearable Computers*. IEEE Computer Society, 118–118. http://www.computer.org/csdl/proceedings/iswc/2003/2034/00/20340118.pdf

[34] Luiz Fernando Gomes Soares and Simone Diniz Junqueira Barbosa. 2009. *Programando em NCL 3.0: desenvolvimento de aplicaçoes para middleware Ginga: TV digital e Web*. Elsevier.

[35] Luiz Fernando Gomes Soares, Rogério Ferreira Rodrigues, and Romualdo Monteiro de Resende Costa. 2005. Nested Context Model 3.0. *Monographs in Computer Sciences* 18/05 (2005), 35.

[36] Unity Technologies. 2018. Unity 3D. https://unity3d.com/

[37] Wade M Vagias. 2006. Likert-type Scale Response Anchors. Clemson International Institute for Tourism. *& Research Development, Department of Parks, Recreation and Tourism Management, Clemson University* (2006).

[38] Viswanath Venkatesh and Fred D Davis. 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science* 46, 2 (2000), 186–204.

[39] Fernando Vera and J. Alfredo Sánchez. 2016. SITUAR: A platform for in-situ augmented reality content creation. *Avances en Interacción Humano-Computadora* 1, 1 (2016), 90–92.

[40] Fernando Vera, J. Alfredo Sánchez, and Ofelia Cervantes. 2017. A Platform for Creating Augmented Reality Content by End Users. In *Applications for Future Internet*, Enrique Sucar, Oscar Mayora, and Enrique Munoz de Cote (Eds.). Vol. 179. Springer International Publishing, Cham, 167–171. http://link.springer.com/10.1007/978-3-319-49622-1_18 DOI: 10.1007/978-3-319-49622-1_18.

[41] Rafał Wojciechowski and Wojciech Cellary. 2013. Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Computers & Education* 68 (2013), 570–585.