



HAL
open science

MarkAR: Exploring the Benefits of Combining Microgestures and Mid-Air Marks to Trigger Commands in Augmented Reality

Charles Bailly, Lucas Pometti, Julien Castet

► **To cite this version:**

Charles Bailly, Lucas Pometti, Julien Castet. MarkAR: Exploring the Benefits of Combining Microgestures and Mid-Air Marks to Trigger Commands in Augmented Reality. Lecture Notes in Computer Science, 2024, Lecture Notes in Computer Science, 15445, pp.163-181. 10.1007/978-3-031-78593-1_11. hal-04826225

HAL Id: hal-04826225

<https://hal.science/hal-04826225v1>

Submitted on 9 Dec 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

MarkAR: Exploring the Benefits of Combining Microgestures and Mid-Air Marks to Trigger Commands in Augmented Reality

Charles Bailly¹[0000-0002-7797-7368], Lucas Pometti¹[0009-0002-0709-0443], and
Julien Castet¹[0000-0002-7017-7288]

Immersion SA, Bordeaux, France `first.last@immersion.fr`

Abstract. Immersive technologies like Augmented Reality (AR) have a promising potential for group activities. Collaborative AR systems allow co-located and/or distant users to create, visualize and manipulate together a large variety of virtual content in order to take actions and decisions. However, contrary to traditional 2D interfaces, there is currently no standard interface or generalized interaction technique for AR. This paper explores an approach for triggering commands linked to AR content. We propose MarkAR, a system combining virtual thumbnail representations of AR content (called *vignettes*), a microgesture and 3D mid-air mark system inspired by marking menus. Experimental results from a qualitative study highlight that MarkAR offers a great usability, subtle interactions and an easy way to trigger both discrete and continuous commands in AR.

Keywords: Augmented Reality, Microgesture, Commands, Gestural Interaction, Collaboration

1 Introduction

There is currently no established practices and tools for collaborating in Augmented Reality (AR) [30]. In fact, many factors are impairing a vast adoption of AR. First of all, AR may not be relevant in many contexts where traditional tools are enough. Secondly, form factors are to be considered. For instance, requiring collaborators to wear additional devices such as Head-Mounted Displays (HMDs) can be troublesome because of their bulkiness. Then, perceptual issues [20], interfaces an interaction design still raise many research questions and explorations, for instance to limit physical arm tiredness linked to prolonged mid-air gestures [10, 28]. Nonetheless, AR offers great benefits for individual and collaborative scenarios by allowing to create, visualize and manipulate virtual content. It can allow each user, co-located or at a remote location, to interact with 2D and 3D virtual content. Many application domains can benefit from collaboration in AR [18], from emergency response to simulation of engineering processes. There is thus on the one hand a potential for AR and on the other hand scientific and technical challenges.

One of these challenges is to investigate how to easily trigger commands upon virtual content in AR contexts. One may argue that a common approach would be to use AR menus [23]. While they offer many benefits, menus can occupy a significant visual space in the already size-limited augmented Field of View of current AR Head Mounted Displays (HMDs), even if they can be transient. Other researchers proposed to perform a specific hand gesture per command [15]. This approach does not occupy any visual space but requires users to learn a set of gestures and the mapping with system commands.

This paper explores a compromise between menus and gestures to take the best of both worlds. We present MarkAR, a technique inspired by 3D marking menus [27] and based on microgestures. On the one hand, marking menus are a variant of radial menus that allow a user to trigger a command by making a mark in the direction of the desired menu item without popping-up the menu [34]. On the other hand, microgestures are short gestures of the hand on itself [3]. Their small amplitude can offer quick and subtle interactions while limiting physical tiredness. Overall, MarkAR allows to remotely trigger discrete and continuous commands over various AR content (textual notes, images, 3D models, and so on). In our system, this content is represented as virtual thumbnails called *vignettes* (see Figure 1).

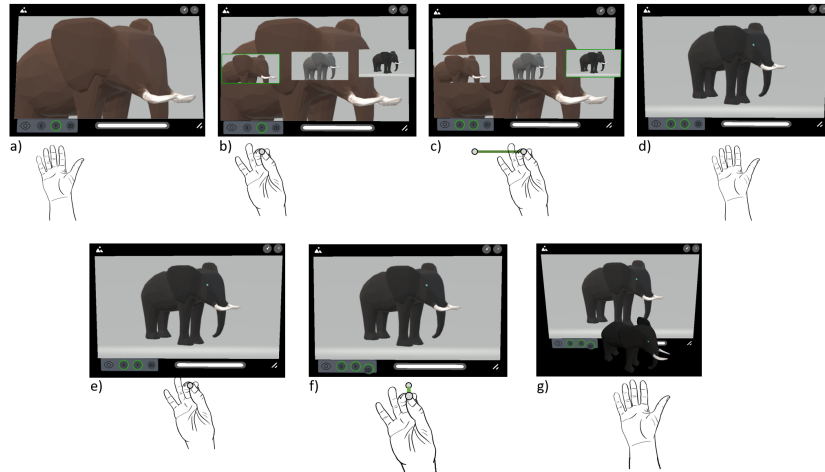


Fig. 1. Examples of MarkAR interactions on a Picture vignette in Augmented Reality. a) to d): Performing a right-oriented mark to trigger the timeline and change the version of the picture. e) to g): Performing a pull mark to extract a 3D model from the current picture vignette.

Our goal is to investigate if this approach can facilitate the trigger of commands in AR, even for novices. We want to evaluate the impact of MarkAR of physical tiredness, mental workload, easiness of learning and subtlety (*i.e.* the

capacity to trigger commands without disturbing other users). MarkAR could be used in many AR context. As a first exploration of its benefits, we nonetheless limit ourselves to a subset of them. In this paper, we focus on contexts with: 1) A small number of users (groups of 2-4 users for instance), 2) A collaboration around a shared AR workspace with a variety of contents and 3) A potential need for punctual, private interaction (for instance, quickly checking another document during a meeting). Crisis management meetings [35], product design sessions or practical work in AR between students are examples of such contexts. The contributions of this paper are the following:

- The design and implementation of MarkAR, a system to visualize and trigger commands on virtual content in AR.
- Results from a first experimental study evaluating the potential of MarkAR and lessons learned from it.

2 Related work

Our work is built on research about collaborative scenarios, AR and interaction techniques like microgestures to trigger commands.

2.1 AR for collaboration

The most obvious benefit of AR for collaborative contexts is the ability to visualize and interact with virtual content in a common mixed environment. This ability has been investigated for many application domains [19], from emergency management [35] to collaborative AR board games [11]. For instance, Broll et al. explored architectural design and urban planning tasks in a collaborative augmented environment [2]. Regenbrecht et al. proposed a collaborative AR system for product design and reviews [26]. Besides, collaboration aspects are actively investigated in the Immersive Analytics community [6].

From a more general point of view, Constantinides and Quercia underlined three major axes in which technologies could support meetings: execution, psychological safety and physical comfort [5]. Execution imply "getting things done", with a clear goal in mind and a structured approach to create actionable items [5]. In addition to visualizing and interacting with objects directly linked to the meeting topic of interest, AR can support the execution of meetings. Radu et al. identified many existing AR features related to this aspect [24], like coordinating attention and instructions. For instance, Kim et al. studied the combination of visual communication cues in Mixed Reality with HMDs [12]. Piumsomboon et al. proposed a miniature AR avatar to represent the gaze direction and body gesture of a remote VR user [22]. The authors observed that this *Mini-Me* avatar increased the feeling of social presence for users.

AR also has the potential to support the psychological safety. According to Constantinides et al., psychological safety includes ensuring all participants 1) give and receive appropriate levels of attention and 2) feel comfortable [5]. For instance, AR can be used to help gather participant feedback during a meeting [8] and to provide private spaces [25].

2.2 Triggering commands in AR

Many AR interaction metaphors, modalities and techniques have been explored to trigger commands. In the following, we consider two major approaches emerging from the literature: menus and (micro)gestures. Other modalities such as voice commands [1] have been excluded since using voice is in conflict with our focus on contexts with a need for private interaction.

AR menus Similarly to traditional Graphical Interfaces (GUIs), menus are a common approach to regroup set of commands in AR. Researchers have investigated how menus could involve physical and virtual elements, like the *Shake menu* from White et al. [31]. Pourmemar et al. compared several methods to interact with hierarchical menus in AR [23]. The authors observed that mid-air hand gestures were judged more tiring than head pointing to select commands on a radial AR menu. Pfeuffer et al. proposed *PalmGazer*, an AR menu system based on hand and eye input [21]. The menu UI can be positioned according to three possible reference points: on the user's hand, above the hand or facing the user's head. Other researchers explored how marking menus could be used in immersive systems based on HMDs. For instance, Lim et al. investigated hand-drawn loop marks to select items in a 2D virtual menu [16]. Ren et al. evaluated different menu layouts and target directions combined with hand gestures [27]. However, the authors used a setup with a Kinect, a computer monitor and stereoscopic glasses instead of the now available AR headsets. On a different note, Kulshreshth et al. explored 3D gesture menu selection by using finger and hand poses, for instance by putting in evidence three fingers to select the third menu item [13].

Microgestures Other works also considered using gestures to trigger commands in AR. In particular, microgestures [3] are drawing more and more interest since they can limit physical tiredness (avoiding the "Gorilla arms" effect [23]). For instance, Liu et al. proposed *Gunslinger*, a bimanual technique where microgestures are performed in a relaxed arms-down position to enter into interaction modes and trigger commands [17]. Li et al. studied the mapping of different hand microgestures with commands (for instance, forming the *Ok* hand sign to accept a call) [15]. On a more general scope, Lambert et al. explored how to graphically represent the various range of microgestures presents in the literature [14]. Based on the movements performed by fingers, the authors distinguished three types of microgestures: *Context-switching*, *In-Context* and *Stationary* microgestures. *Context-switching* is the most frequent type and include touch and release events like in [7]. It is worth noticing that currently available HMDs like the Hololens 2 from Microsoft ¹ or the recent Visio Pro from Apple ² natively propose this type of microgesture for basic interactions like grabbing virtual objects or confirm

¹ <https://www.microsoft.com/fr-fr/hololens>

² <https://www.apple.com/apple-vision-pro/>

a selection with a thumb-index pinch gesture. *In-Context* microgestures imply that fingers keep the same context, i.e. that they keep contact with the same surface or stay in the air. For instance, Xu et al. proposed a text-entry technique on a virtual keyboard as small as the tip of an index [32]. The authors designed several swipe microgestures to enter into word selection mode or to delete a letter. Finally, *Stationary* microgestures involve factors other than movement, such as pression [29].

2.3 Positioning of this paper

Microgestures are promising to trigger commands during AR sessions as they could limit the impact on users' physical comfort [5]. Subtle interactions are also interesting in this context to interact with the system without disturbing too much the other members of the group. However, having a different microgesture for each available command like in [15] may require a significant learning process and impact users' cognitive load. On the contrary, our aim is to explore a system limiting the number of elements novice users would have to learn and memorize. We also wanted to avoid slow selection mechanisms such as dwell time [23] because they can heavily impact the interaction.

The study by Pfeuffer et al. [21] is closer to our work. However, the authors focus on menu UIs for a single user. On the contrary, we 1) consider collaboration aspects and 2) we want to explore how to preserve the visual attention of users on meeting objects and files instead of navigating through hierarchical menus and UIs.

Overall, we formulated 4 Design Principles (DP) for MarkAR to follow:

- **DP1: Quick and subtle** - Giving the ability to interact without disturbing others.
- **DP2: Comfortable** - Limited physical tiredness, usable with the arm resting on an armchair or table.
- **DP3: Focused on meeting content** - No navigation through additional UIs and menus.
- **DP4: Easy to learn for novices** - Limited amount of learning and memorization.

3 MarkAR

3.1 Displaying virtual content as vignettes

The first question we considered was how to represent in AR traditional 2D content like images, videos and PDF files. According to DP4, we wanted users to quickly get familiar with our system, including AR novices. We thus selected a thumbnail visual representation close to existing 2D systems: *vignettes*. Vignettes are virtual containers with 1) a content and 2) GUI elements to visualize its internal state and interaction purposes, as shown in Figure 2.

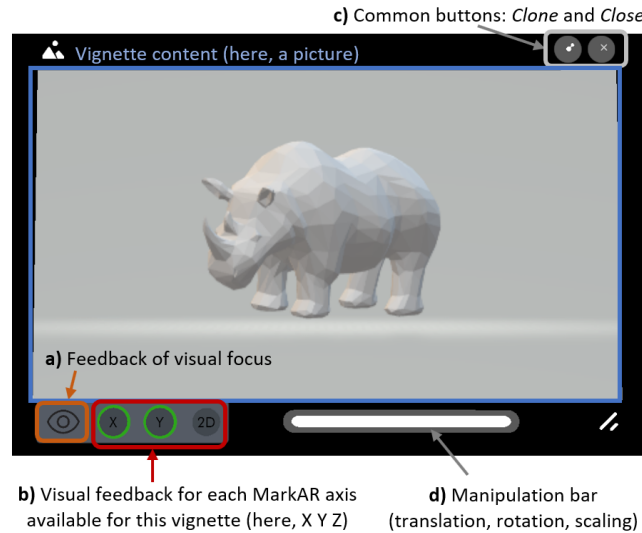


Fig. 2. Details of the common graphical elements of a vignette.

Vignettes are displayed in AR and synchronized between the participants. More precisely, their position, rotation, scale and content are synchronized in real time. Again motivated by our principle DP4, we chose to let users perform basic manipulations (translation, rotation, scaling) with traditional interaction techniques. For instance, a HoloLens user can grab a nearby vignette by targeting the manipulation bar with the default pinch hand gesture (thumb + index) of the HMD. **The goal of MarkAR is therefore not to replace or compete with baseline pointing/manipulation techniques pre-implemented in current AR HMDs.** Instead, we wanted MarkAR to be a complement to easily trigger commands. Our original motivation was to offer an all in-one view like the Common Operational Picture for crisis management meetings, but it is also applicable to other contexts like brainstorming. Overall, we implemented 4 types of vignette in MarkAR: Textual note, Image, Video and PDF vignettes, as illustrated in Figure 3.

3.2 Triggering commands through marks

The main goal of MarkAR is to trigger commands on nearby and distant vignettes. We wanted users to be able to perform common operations for each type of vignette, like navigating through the pages of a PDF or changing the color of a Textual note. We thus needed an interaction technique allowing both continuous and discrete input. Besides, we aimed at creating interaction metaphors that could be shared among all types of vignettes, instead of having to learn a command mapping for each vignette type. To do so, we designed a technique based on a microgesture and mid-air marks, as illustrated in Figure 1.



Fig. 3. Overview of different types of AR vignettes present in MarkAR.

MarkAR mode Users first have to look at the vignette they are interested in (principle DP3). At any time, the system activates only the closest vignette currently hovered by user’s head gaze (Figure 2a). The user can then enter into MarkAR mode by performing a three-finger pinch gesture (thumb + index + middle), illustrated in Figure 4. According to the categories proposed by Lambert et al., it corresponds to a *Context-switching* microgesture [14]. Using a microgesture was motivated by principles DP1 and DP2. Besides, we wanted a comfortable microgesture with an ergonomic hand position to draw mid-air marks (DP2), close to existing HMDs reference gestures (DP4, easy to learn) and robustly detectable without additional hardware on users (optical detection). As long as the microgesture is held, the MarkAR mode stays active. Within this mode, the user can draw a virtual line (a *mark*) in a given direction. If users release the gesture while being too close to their starting point (less than 3cm away from the initial MarkAR microgesture detection), no command is triggered. We added this *neutral zone* mechanism to allow users to cancel a MarkAR interaction without triggering any effect. In a way, drawing a mark could be compared to controlling a virtual slider with a low number of ticks (discrete command) or high number of ticks (continuous command). Releasing the microgesture triggers the command corresponding to that direction.

Mapping of commands Following DP4, we did not want users to memorize a set of microgestures. Instead, we chose to map commands to the direction of marks, following semantic associations on three main axis relative to the user:

- **The X-axis is the timeline axis.** Marks going in the left direction are mapped to backward navigation (for instance, going back in a video or previous pages of a PDF). Marks going right are mapped to future navigation (ex: next pages of the PDF).

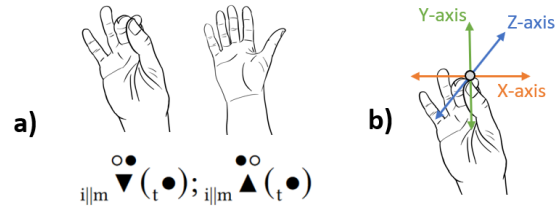


Fig. 4. a) The microgesture to enter MarkAR mode and its μ Glyph notation [3]: a three fingers pinch (thumb, index and middle), palm facing the user. b) Available directions when the microgesture is held.

- **The Y-axis is the vertical slider axis.** It is used to increase (up-mark) or decrease (downward-mark) a given value (for instance, controlling the audio volume of a video).
- **The Z-axis is the depth axis.** It is used to switch from 2D to 3D (for instance, pulling the content of an image to extract a 3D model from it) and vice versa (pushing back a 3D model into it’s 2D image).

This design creates a link between the motor space and the command (for instance, going up and down to control the audio volume of a video). We hypothesized that this may facilitate learning MarkAR and reduce users’ cognitive load, especially since the axis have the same logic for all vignettes (ex: Y-axis is always the vertical slider). A summary of the available commands is presented in the Table 1.

Vignette type		Axis	Command
Textual note		Y	Change the note color (hue)
PDF		X	Switch to previous/next pages
Video		X	Return/advance to timecode
		Y	Decrease/Increase audio volume
Picture		X	Switch to previous/next image
		Y	Decrease/Increase image opacity
		Z	Extract/Fade 3D model of the image

Table 1. Summary of available commands on all vignette types.

Guidance feedbacks We implemented several feedback mechanisms to support MarkAR interactions. First, when users perform a mark, a virtual line is drawn to reflect their gesture. The line stays grey as long as the user is within the neutral zone, and becomes green when outside of it. Nonetheless, we wanted users to focus on meeting content (DP3) instead of looking at their hand. We thus added cylinder widgets on the bottom bar of each vignette UI (Figure 2b). These cylinders reflect which axes are available for this vignette. When the

user is drawing a mark, a green circular feedback is continuously updated to show the current progress of the corresponding axis. The cylinder itself will also slightly adjust its shape to reflect the current progress. For instance, extracting a 3D model from an image will also pull and stretch the Z-axis cylinder of the vignette. Moreover, the effect of commands can also be previsualized on the vignette (Figure 5). For instance, users can see the timecode corresponding to their gesture when navigating through a Video vignette.

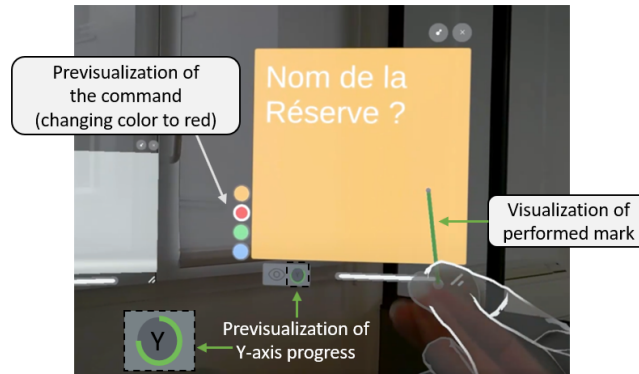


Fig. 5. User performing to downward-mark on a Textual note vignette to change its color.

Finally, we also added subtle audio feedback when the microgesture is initially detected and when it is released.

3.3 Implementation

We developed MarkAR with the Unity framework (version 2021.3.27), using MRTK3 packages for interactions. Custom C# scripts based on MRTK hand detection scripts were designed to detect the 3-fingers microgesture and performed marks. We selected Netcode for GameObjects to implement the multiplayer aspects of MarkAR (client-server architecture, synchronisation of vignettes and 3D models). The 3D models corresponding to Picture vignettes were all loaded and spawned at the beginning of the session. They were simply displayed/hidden and re-positioned when users interacted with the corresponding vignette. We deployed our system on HoloLens 2 HMDs.

3.4 Resulting system

Overall, users only have to use a single microgesture and perform a mark in a given direction to remotely trigger a command on the vignette they are interested in. While doing so, they can keep their visual attention on meeting objects.

Besides, users can combine both 2D and 3D AR content by extracting 3D models from Picture vignettes. To evaluate the potential of the resulting MarkAR system, we conducted an experimental study.

4 Experimental study

Our goal was to evaluate to which extent MarkAR reflects its design principles. We aimed to collect feedback about the usability, physical tiredness, mental workload, easiness of learning and subtlety of our system. We thus conducted a qualitative experimental study. To do so, we took inspiration from the protocol followed by Zhang et al. [33], as detailed below.

4.1 Participants

A total of 16 participants (7 female, 9 male) from 23 to 42 years old ($sd = 4.4$) participated in the study. They were from nearby companies and labs and had different backgrounds including Computer Science, Design, Human Factors and Communication. Participants passed the experiment in pairs (8 groups of two) of comparable AR experience. Demographic details are given in Figure 6. Overall, the experiment lasted for one hour.

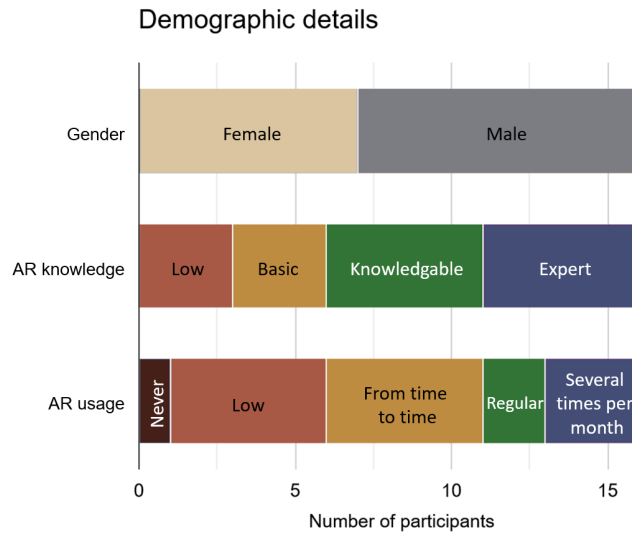


Fig. 6. Demographic details about the participants of the study

4.2 Procedure

The study took place in a calm meeting room with a table, chairs, and two HoloLens 2 headsets. We carefully considered ambient lighting and contrast in this environment in order to ensure both HMD could properly detect the gestures of users. Participants were seated and encouraged to find a comfortable position for their arms. They could choose to use their left or right hand to use MarkAR. The study was split into three phases: an initial phase, an individual phase and a collective phase.

During the initial phase, an experimenter introduced the goal and usage of MarkAR to the pair of participants. Participants could then get more familiar with the system by following a walk-through of the commands associated to the different vignette types. Each time, participants could trigger the command as much as they wanted until feeling at ease with it. Once this training ended, the individual phase started, as detailed in next section. After the individual phase, participants could take a break and remove their headset. They were then encouraged to fill the French version of the SUS questionnaire [9] before starting the collective phase. Finally, at the end of the collective phase, participants could answer a survey based on 5-item Likert scales (from 1: Strongly disagree to 5: Strongly agree) to give their opinion about specific aspects of MarkAR. The experiment ended after a short semi-structured interview.

4.3 Measures

We collected quantitative about the usability and easiness of learning of MarkAR through the aforementioned SUS questionnaire [9]. The final custom questionnaire (see Table 2) and the semi-structured interviews were design to collect subjective feedback about the physical tiredness, cognitive load and perceived subtlety of the system.

4.4 Task

In this section, we detail the tasks performed during the individual and collective phases.

During the individual phase, each participant had to individually complete a set of 12 trials. The AR sessions had the same content (same vignettes on each side) but were not synchronized. Trials were separated into 3 levels of progressive difficulty. *Easy* trials required participants to trigger one basic MarkAR command (for instance: changing a given Textual note color to blue). *Medium* trials required either 1) two MarkAR commands or 2) one MarkAR command and a manipulation task (for instance, moving a vignette to a given location). *Hard* trials required at least two precise MarkAR commands (for instance, setting audio volume to 70% on a Video vignette) and manipulation tasks. Participants always started by the easy trials and ended with the hard ones, but the order of trials within each level of difficulty was randomized.

During the collective phase, participants were seated side by side. They had a

new AR session with synchronized content (4 Picture vignettes, 1 PDF, 1 Textual note and 1 Video vignettes). These vignettes were different from before. The goal of this phase was to simulate a meeting performed with MarkAR. Participants had to collaborate in order to create the best animal reserve. To do so, they had to place the 3D models of selected animals (extracted from Picture vignettes) into a dedicated area. Participants could discuss to chose the animals they wanted, but recommendations were disseminated into the different vignettes. Recommendations were giving advice about the compatibility between animals. Therefore, participants were encouraged to use MarkAR to interact with vignettes, discuss and take decisions.

5 Results

All participants successfully completed individual and collective phases. Besides, they expressed positive opinions about MarkAR. A summary of answers given for second questionnaire (which was given after the collective phase) is available at Table 2. In the following, we detail the results collected about (i) the perceived utility and usability of MarkAR (ii) collaborative aspects. Scores are based on 5-item Likert scales.

5.1 Utility and Usability of MarkAR

Most participants were enthusiastic about MarkAR, expressing that it was *"cool"* (P14) and *"satisfying to use"* (P2 and P5). On average, participants declared that they would like to user MarkAR if they had the opportunity to do a meeting with AR (avg = 4.25, *sd* = 0.56). For instance, P14 explained that MarkAR *"would be useful to work around virtual 3D models of [room] installations and trigger commands on their different parts"*. Using the timeline (X-axis) on Picture vignettes and extracting a 3D model from it was particularly appreciated by participants. These interactions were judged *"well done"* (P13) and *"natural"* (P4). We did not observed any effect or mention of the difficulty of the tasks.

Results from the SUS questionnaire given after the individual phase suggest that MarkAR offers a great usability (avg = 79.1, *sd* = 9.4). Representing common file formats as virtual AR vignettes was appreciated and participants found this representation easy to understand (avg = 4.8, *sd* = 0.3). They reported that commands were also easy to understand (avg = 4.25, *sd* = 0.47) and easy to memorize (avg = 4.3, *sd* = 0.67). P16 reported that *"you memorize quickly, you know what you have to do"*. P9, who was novice in AR, declared that *"it takes a bit of time to get familiar with the Hololens, but once it is done it's quite intuitive and easy to memorize"*. This aspect was confirmed by a low subjective mental demand: participants did not find MarkAR mentally tiring (avg = 2.06, *sd* = 0.96). Besides, they reported that they felt able to know which commands was about to be triggered with their gesture (avg = 4.25, *sd* = 0.66).

Participants did not find MarkAR physically tiring (avg = 2.13, *sd* = 0.78). However, their response was more mitigated about the difficulty to enter into

Question	Average	SD
I found the virtual vignette representation easy to understand	4.8	0.3
I found that the different commands were easy to understand	4.25	0.47
I found that the different commands were easy to memorize	4.3	0.67
I felt able to know which command would be triggered by my gesture	4.25	0.66
I found this technique mentally tiring	2.06	0.96
I found this technique physically tiring	2.13	0.78
I struggled to activate the MarkAR mode (3-fingers microgesture)	2.7	1.35
I could easily switch from moving a vignette to triggering commands	3.8	0.66
I often needed to look at my hands during the interaction	2.3	1.2
I found that MarkAR allowed to easily collaborate with others	3.94	0.85
I would like to use MarkAR during AR meetings	4.25	0.56

Table 2. Overview of answers for the second questionnaire, given after the collective phase. Questions were based on a 5-items Likert scale.

MarkAR mode with the 3-fingers microgesture (2.7 average, $sd = 1.35$). 3 out of 16 participants reported that they were sometimes confused the manipulation gesture (two-fingers pinch) and the MarkAR microgesture. Still, most participants reported that it was easy to switch from vignette manipulations to MarkAR commands (avg = 3.8, $sd = 0.66$). Participants also reported that they did not often needed to look at their hand during the interaction (avg = 2.3, $sd = 1.2$). The head-gaze activation of vignettes was welcomed positively by participants, which found this approach efficient. However, 3 out of 16 participants mentioned that the eye feedback at the bottom left corner of vignettes (Figure 2a) was not visible enough. These participants would have preferred a more visible feedback to know the currently active vignette, especially in dense locations with multiple superimposed vignettes.

5.2 Collaboration aspects

Overall, participants agreed that MarkAR allows to collaborate easily with others: (avg = 3.94, $sd = 0.85$). Verbatims indicate that participants asked questions to their partner about the state of the session ("*Did you put a second elephant?*", P9) and voiced requests ("*Take a grey rhinoceros and put it there*", P9). They also coordinated their working environment (P2: "*Ok, I will move the PDF to let you some space*"). During the collective phase of the experiment, the vignette contents were synchronized: participants could see the result of a command triggered by their partner. However, 2 out of 16 participants also mentioned that they would also like to synchronize the previsualisation of commands (for instance, seeing another user exploring the timeline of a Picture vignette).

MarkAR interactions were perceived as subtle. Nearly all participants declared that the microgesture and marks were not obtrusive and that their attention was not disturbed by the gestures of their partner. Only P16 declared that MarkAR gestures could "*create minor interference, but not really disturb*". P4 suggested that MarkAR could be used to perform punctual personal tasks during meetings.

6 Discussion

6.1 Current limitations

Our work currently presents two limitations. First, the detection of the MarkAR microgesture and marks was fully done optically through the HoloLens sensors. This choice was motivated by our will to avoid burdening users with additional hardware since they already had an HMD. Users were able to easily interact with MarkAR but the robustness of the detection would still need to be improved, especially for less controlled environments. Besides, the field of detection of a headset like the HoloLens is still limited in size, which restrict the body postures users can adopt while interacting with the system.

Secondly, the current version of MarkAR offers limited support for collaboration. While the AR vignettes are synchronized between users, the system did not include mechanisms to support factors like group awareness, concurrency or history management. In this work, we focused on evaluating first the utility and usability of our approach. We plan to investigate further those collaboration aspects with a bigger number of meeting participants in a next study.

6.2 Lessons learned

Despite the limitations mentioned above, our work suggest several insights about triggering commands in AR context.

MarkAR allows to subtly and easily trigger commands on AR vignettes The feedback collected in this study about MarkAR is very positive. First, using a microgesture and small mid-air marks seems to have limited the physical tiredness felt by participants. This choice also allowed subtle interaction without disturbing others (Design Principle 1). Being able to interact with the system without impacting the activity or attention of the group is desirable in collaborative contexts. Besides, subtle interaction could also reinforce privacy for personal, punctual interactions (for instance, quickly browsing through an email during an AR meeting). The quick and subtle aspect of MarkAR is coherent with previous studies about microgestures [3, 4]. Moreover, the observed complementarity between the MarkAR mode to trigger commands and the baseline HoloLens technique for spatial manipulations is also a promising insight. Besides, MarkAR seems to offer several benefits compared to traditional AR menus. Drawing marks allows to control both discrete commands (for instance: switching from 2D to 3D with the Z-axis) and continuous commands (for instance: controlling audio volume for a video) instead of simply selecting a menu item. Moreover, it allows users to stay visually focused on the targeted AR content (instead of navigating through menu UIs and hierarchies) while still receiving appropriate feedback. The present study focuses on common file formats and vignettes. This AR vignette representation was found easy to understand and allowed to work with both 2D content and 3D models. Nonetheless, the MarkAR approach could be adapted to any virtual or mixed contents.

Overall, combining a microgesture to enter into a mid-air marking mode and trigger commands thus seems a relevant and usable approach to interact with an AR system that deserves further research.

Mid-air marks require a compromise between accuracy and gesture amplitude In our setup, the MarkAR mid-air marks allowed both discrete and continuous commands. While the former was easy to perform, participants reported that the latter required a high accuracy. For instance, 6/16 participants explained that it could be hard to go exactly to a given PDF page or video timecode this a single MarkAR interaction. In this study, we had made the choice to privilege low amplitude gestures to limit physical tiredness. Independently of the robustness of the tracking system, the motor space linked to the interaction was thus reduced. Besides, we had implemented a linear gain between the size of mid-air marks and the effect of the command. This choice favoured a quick navigation instead of an accurate one. Further research is required to determine the optimal compromise between accuracy and gesture amplitude for this kind of interaction. In the meantime, we recommend designers to dedicate specific attention to commands that could require both accuracy (like going to a specific page) and speed (like browsing quickly towards the end of a PDF) depending on the user’s current goal.

Commands mapping can benefit from user customization The available MarkAR commands were judged easy to understand and memorize, which matches with our design principle DP4. Overall, mapping commands on the X, Y, and Z axis relative to the user was a successful approach in our study. It offered a unified method to interact with different types of AR content that can be customized depending on the context of use. Interestingly, we observed that 5/16 participants express the desire to adapt the X-axis commands. Instead of a timeline metaphor (going from left to right), these participants would have preferred a page-turning metaphor (from right to left). The other participants did not express this want or declared that they appreciated the timeline as is. We thus encourage designers to consider giving freedom to users who would like to customize the available interaction metaphors.

7 Conclusion

In this work, we introduced MarkAR, a system based on a microgesture and mid-air marks to remotely trigger commands in AR. MarkAR proposes to display traditional meeting files as virtual thumbnails that can be synchronized among several users. Available commands are mapped to the direction of mid-air marks in order to obtain a unified gestural language for all types of vignettes. Experimental results highlight the potential of our approach. MarkAR allowed subtle and easy to learn AR interactions and we believe its generalization to other types of content and contexts deserves to be explored.

Our next step consists in adding new vignette types to MarkAR to reflect a larger content diversity. We then plan to study further the collaborative aspects of our system, including group awareness mechanisms.

Acknowledgments. This work was partly supported by the French National Research Agency (ANR) project MIC (ANR-22-CE33-0017).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bonilla Fominaya, A.M., Chew, R.K.R., Komar, M.L., Lo, J., Slabakis, A., Sun, N.A., Zhang, Y.J., Lindlbauer, D.: MoonBuddy: A Voice-based Augmented Reality User Interface That Supports Astronauts During Extravehicular Activities. In: Adjunct Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology. pp. 1–4. UIST '22 Adjunct, Association for Computing Machinery, New York, NY, USA (Oct 2022). <https://doi.org/10.1145/3526114.3558690>, <https://dl.acm.org/doi/10.1145/3526114.3558690>
2. Broll, W., Lindt, I., Ohlenburg, J., Wittkämper, M., Yuan, C., Novotny, T., Fatah gen. Schieck, A., Mottram, C., Strothmann, A.: ARTHUR: A Collaborative Augmented Environment for Architectural Design and Urban Planning. *JVRB - Journal of Virtual Reality and Broadcasting* **1(2004)**(1) (Dec 2004). <https://doi.org/10.20385/1860-2037/1.2004.1>, <http://www.jvr.org/past-issues/1.2004/34>
3. Chaffangeon Caillet, A., Goguy, A., Nigay, L.: μ Glyph: a Microgesture Notation. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–28. CHI '23, Association for Computing Machinery, New York, NY, USA (Apr 2023). <https://doi.org/10.1145/3544548.3580693>, <https://dl.acm.org/doi/10.1145/3544548.3580693>
4. Chan, E., Seyed, T., Stuerzlinger, W., Yang, X.D., Maurer, F.: User elicitation on single-hand microgestures. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. p. 3403–3414. CHI '16, Association for Computing Machinery, New York, NY, USA (2016). <https://doi.org/10.1145/2858036.2858589>, <https://doi.org/10.1145/2858036.2858589>
5. Constantinides, M., Quercia, D.: The Future of Hybrid Meetings. In: Proceedings of the 1st Annual Meeting of the Symposium on Human-Computer Interaction for Work. pp. 1–6. CHIWORK '22, Association for Computing Machinery, New York, NY, USA (Jun 2022). <https://doi.org/10.1145/3533406.3533415>, <https://dl.acm.org/doi/10.1145/3533406.3533415>
6. Ens, B., Bach, B., Cordeil, M., Engelke, U., Serrano, M., Willett, W., Prouzeau, A., Anthes, C., Büschel, W., Dunne, C., Dwyer, T., Grubert, J., Haga, J.H., Kirshenbaum, N., Kobayashi, D., Lin, T., Olaosebikan, M., Pointecker, F., Saffo, D., Saquib, N., Schmalstieg, D., Szafir, D.A., Whitlock, M., Yang, Y.: Grand Challenges in Immersive Analytics. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1–17. ACM, Yokohama Japan (May 2021). <https://doi.org/10.1145/3411764.3446866>, <https://dl.acm.org/doi/10.1145/3411764.3446866>

7. Ens, B., Quigley, A., Yeo, H.S., Irani, P., Piumsomboon, T., Billinghurst, M.: Counterpoint: Exploring mixed-scale gesture interaction for ar applications. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. p. 1–6. CHI EA '18, Association for Computing Machinery, New York, NY, USA (2018). <https://doi.org/10.1145/3170427.3188513>, <https://doi.org/10.1145/3170427.3188513>
8. Francese, R., Passero, I., Zarranandia, T.: An Augmented Reality Application to Gather Participant Feedback During a Meeting. In: De Marco, M., Te'eni, D., Albano, V., Za, S. (eds.) *Information Systems: Crossroads for Organization, Management, Accounting and Engineering*. pp. 173–180. Physica-Verlag HD, Heidelberg (2012). https://doi.org/10.1007/978-3-7908-2789-7_20
9. Gronier, G., Baudet, A.: Psychometric Evaluation of the F-SUS: Creation and Validation of the French Version of the System Usability Scale. *International Journal of Human-Computer Interaction* **37**(16), 1571–1582 (Oct 2021). <https://doi.org/10.1080/10447318.2021.1898828>, <https://doi.org/10.1080/10447318.2021.1898828>, publisher: Taylor & Francis _eprint: <https://doi.org/10.1080/10447318.2021.1898828>
10. Hansberger, J.T., Peng, C., Mathis, S.L., Areyur Shanthakumar, V., Meacham, S.C., Cao, L., Blakely, V.R.: Dispelling the Gorilla Arm Syndrome: The Viability of Prolonged Gesture Interactions. In: Lackey, S., Chen, J. (eds.) *Virtual, Augmented and Mixed Reality*, vol. 10280, pp. 505–520. Springer International Publishing, Cham (2017). <https://doi.org/10.1007/978-3-319-57987-0-41>, series Title: *Lecture Notes in Computer Science*
11. Huynh, D.N.T., Raveendran, K., Xu, Y., Spreen, K., MacIntyre, B.: Art of defense: a collaborative handheld augmented reality board game. In: *Proceedings of the 2009 ACM SIGGRAPH Symposium on Video Games*. pp. 135–142. Sandbox '09, Association for Computing Machinery, New York, NY, USA (Aug 2009). <https://doi.org/10.1145/1581073.1581095>, <https://dl.acm.org/doi/10.1145/1581073.1581095>
12. Kim, S., Lee, G., Huang, W., Kim, H., Woo, W., Billinghurst, M.: Evaluating the Combination of Visual Communication Cues for HMD-based Mixed Reality Remote Collaboration. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. pp. 1–13. CHI '19, Association for Computing Machinery, New York, NY, USA (May 2019). <https://doi.org/10.1145/3290605.3300403>, <https://dl.acm.org/doi/10.1145/3290605.3300403>
13. Kulshreshth, A., LaViola, J.J.: Exploring the usefulness of finger-based 3D gesture menu selection. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 1093–1102. ACM, Toronto Ontario Canada (Apr 2014). <https://doi.org/10.1145/2556288.2557122>, <https://dl.acm.org/doi/10.1145/2556288.2557122>
14. Lambert, V., Chaffangeon Caillet, A., Goguey, A., Malacria, S., Nigay, L.: Studying the Visual Representation of Microgestures. *Proceedings of the ACM on Human-Computer Interaction* **7**(MHCI), 1–36 (Sep 2023). <https://doi.org/10.1145/3604272>, <https://dl.acm.org/doi/10.1145/3604272>
15. Li, G., Rempel, D., Liu, Y., Song, W., Adamson, C.H.: Design of 3D Microgestures for Commands in Virtual Reality or Augmented Reality. *Applied Sciences* **11**(14), 6375 (Jan 2021). <https://doi.org/10.3390/app11146375>, <https://www.mdpi.com/2076-3417/11/14/6375>, number: 14 Publisher: Multidisciplinary Digital Publishing Institute

16. Lim, Z.H., Kristensson, P.O.: An Evaluation of Discrete and Continuous Mid-Air Loop and Marking Menu Selection in Optical See-Through HMDs. In: Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services. pp. 1–10. ACM, Taipei Taiwan (Oct 2019). <https://doi.org/10.1145/3338286.3340127>, <https://dl.acm.org/doi/10.1145/3338286.3340127>
17. Liu, M., Nancel, M., Vogel, D.: Gunslinger: Subtle Arms-down Mid-air Interaction. In: Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology. pp. 63–71. ACM, Charlotte NC USA (Nov 2015). <https://doi.org/10.1145/2807442.2807489>, <https://dl.acm.org/doi/10.1145/2807442.2807489>
18. Lu, F., Xu, Y.: Exploring Spatial UI Transition Mechanisms with Head-Worn Augmented Reality. In: CHI Conference on Human Factors in Computing Systems. pp. 1–16. ACM, New Orleans LA USA (Apr 2022). <https://doi.org/10.1145/3491102.3517723>, <https://dl.acm.org/doi/10.1145/3491102.3517723>
19. Lukosch, S., Billingham, M., Alem, L., Kiyokawa, K.: Collaboration in Augmented Reality. Computer Supported Cooperative Work (CSCW) **24**(6), 515–525 (Dec 2015). <https://doi.org/10.1007/s10606-015-9239-0>, <http://link.springer.com/10.1007/s10606-015-9239-0>
20. Patterson, R., Winterbottom, M.D., Pierce, B.J.: Perceptual Issues in the Use of Head-Mounted Visual Displays. Human Factors: The Journal of the Human Factors and Ergonomics Society **48**(3), 555–573 (Sep 2006). <https://doi.org/10.1518/001872006778606877>, <http://journals.sagepub.com/doi/10.1518/001872006778606877>, number: 3
21. Pfeuffer, K., Obernolte, J., Dietz, F., Mäkelä, V., Sidenmark, L., Manakhov, P., Pakanen, M., Alt, F.: PalmGazer: Unimanual Eye-hand Menus in Augmented Reality. In: Proceedings of the 2023 ACM Symposium on Spatial User Interaction. pp. 1–12. SUI '23, Association for Computing Machinery, New York, NY, USA (Oct 2023). <https://doi.org/10.1145/3607822.3614523>, <https://dl.acm.org/doi/10.1145/3607822.3614523>
22. Piumsomboon, T., Lee, G.A., Hart, J.D., Ens, B., Lindeman, R.W., Thomas, B.H., Billingham, M.: Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp. 1–13. ACM, Montreal QC Canada (Apr 2018). <https://doi.org/10.1145/3173574.3173620>, <https://dl.acm.org/doi/10.1145/3173574.3173620>
23. Pourmemar, M., Poullis, C.: Visualizing and Interacting with Hierarchical Menus in Immersive Augmented Reality. In: Proceedings of the 17th International Conference on Virtual-Reality Continuum and its Applications in Industry. pp. 1–9. VRCAI '19, Association for Computing Machinery, New York, NY, USA (Nov 2019). <https://doi.org/10.1145/3359997.3365693>, <https://dl.acm.org/doi/10.1145/3359997.3365693>
24. Radu, I., Joy, T., Bowman, Y., Bott, I., Schneider, B.: A Survey of Needs and Features for Augmented Reality Collaborations in Collocated Spaces. Proceedings of the ACM on Human-Computer Interaction **5**(CSCW1), 1–21 (Apr 2021). <https://doi.org/10.1145/3449243>, <https://dl.acm.org/doi/10.1145/3449243>, number: CSCW1
25. Rajaram, S., Chen, C., Roesner, F., Nebeling, M.: Eliciting Security & Privacy-Informed Sharing Techniques for Multi-User Augmented Reality. In:

- Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–17. CHI '23, Association for Computing Machinery, New York, NY, USA (Apr 2023). <https://doi.org/10.1145/3544548.3581089>, <https://dl.acm.org/doi/10.1145/3544548.3581089>
26. Regenbrecht, H., Wagner, M., Baratoff, G.: MagicMeeting: A Collaborative Tangible Augmented Reality System. *Virtual Reality* **6**(3), 151–166 (Oct 2002). <https://doi.org/10.1007/s100550200016>, <https://doi.org/10.1007/s100550200016>
 27. Ren, G., O'Neill, E.: 3D Marking menu selection with freehand gestures. In: 2012 IEEE Symposium on 3D User Interfaces (3DUI). pp. 61–68 (Mar 2012). <https://doi.org/10.1109/3DUI.2012.6184185>, <https://ieeexplore.ieee.org/document/6184185>
 28. Schön, D., Kosch, T., Müller, F., Schmitz, M., Günther, S., Bommhardt, L., Mühlhäuser, M.: Tailor Twist: Assessing Rotational Mid-Air Interactions for Augmented Reality. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–14. CHI '23, Association for Computing Machinery, New York, NY, USA (Apr 2023). <https://doi.org/10.1145/3544548.3581461>, <https://dl.acm.org/doi/10.1145/3544548.3581461>
 29. Sharma, A., Roo, J.S., Steimle, J.: Grasping Microgestures: Eliciting Single-hand Microgestures for Handheld Objects. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–13. CHI '19, Association for Computing Machinery, New York, NY, USA (May 2019). <https://doi.org/10.1145/3290605.3300632>, <https://doi.org/10.1145/3290605.3300632>
 30. Speicher, M., Hall, B.D., Nebeling, M.: What is Mixed Reality? In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–15. ACM, Glasgow Scotland Uk (May 2019). <https://doi.org/10.1145/3290605.3300767>, <https://dl.acm.org/doi/10.1145/3290605.3300767>
 31. White, S., Feng, D., Feiner, S.: Interaction and presentation techniques for shake menus in tangible augmented reality. In: Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality. pp. 39–48. ISMAR '09, IEEE Computer Society, USA (Oct 2009). <https://doi.org/10.1109/ISMAR.2009.5336500>, <https://doi.org/10.1109/ISMAR.2009.5336500>
 32. Xu, Z., Wong, P.C., Gong, J., Wu, T.Y., Nittala, A.S., Bi, X., Steimle, J., Fu, H., Zhu, K., Yang, X.D.: TipText: Eyes-Free Text Entry on a Fingertip Keyboard. In: Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology. pp. 883–899. ACM, New Orleans LA USA (Oct 2019). <https://doi.org/10.1145/3332165.3347865>, <https://dl.acm.org/doi/10.1145/3332165.3347865>
 33. Zhang, L., Agrawal, A., Oney, S., Guo, A.: VRGit: A Version Control System for Collaborative Content Creation in Virtual Reality. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. pp. 1–14. CHI '23, Association for Computing Machinery, New York, NY, USA (Apr 2023). <https://doi.org/10.1145/3544548.3581136>, <https://dl.acm.org/doi/10.1145/3544548.3581136>
 34. Zhao, S., Balakrishnan, R.: Simple vs. compound mark hierarchical marking menus. In: Proceedings of the 17th annual ACM symposium on User interface software and technology. pp. 33–42. UIST '04, Association for Computing Machin-

- ery, New York, NY, USA (Oct 2004). <https://doi.org/10.1145/1029632.1029639>, <https://dl.acm.org/doi/10.1145/1029632.1029639>
35. Zhu, Y., Li, N.: Virtual and augmented reality technologies for emergency management in the built environments: A state-of-the-art review. *Journal of Safety Science and Resilience* **2**(1), 1–10 (Mar 2021). <https://doi.org/10.1016/j.jnlssr.2020.11.004>, <https://www.sciencedirect.com/science/article/pii/S266644962030030X>