SAPIENS in XR: Operationalizing Interaction-Attention in Extended Reality

Cristian Pamparău, Ovidiu-Andrei Schipor, Alexandru Dancu and Radu-Daniel Vatavu

MintViz Lab, Ştefan cel Mare University of Suceava, 13 Universităţii, Suceava, 720229, Suceava, Romania .

*Corresponding author(s). E-mail(s): radu.vatavu@usm.ro; Contributing authors: cristian.pamparau@usm.ro; ovidiu.schipor@usm.ro; alexandru.dancu@gmail.com;

Abstract

We examine peripheral interactions in XR environments, for which we propose a conceptual space with two specialized dimensions, Interaction-Attention and Reality-Virtuality. We also formalize the notion of an "XR display" to expand the application range of ambient displays from physical environments to XR. To operationalize these conceptual contributions for researchers and practitioners, we capitalize on Sapiens, an open-source event-based software architecture for peripheral interactions in smart environments, to propose Sapiens-in-XR, an extended architecture that also covers XR displays. In a simulation study based on a Poisson probabilistic model of notification delivery, we demonstrate the efficiency of the event processing pipeline of Sapiens-in-XR with an average processing time of just 18ms from event creation to delivery. We present simulations of peripheral interaction scenarios enabled by our conceptual space and Sapiens-in-XR, and report empirical results from a controlled experiment implementing one scenario, where users were asked to maintain their focus of attention in the central field of view while notifications were displayed at the attention periphery. Our results show similar user perception and the same level of user performance with understanding and recalling content of notifications in either the virtual and physical environments. Our conceptual space, software architecture, and simulator constitute tools meant to assist researchers and practitioners to explore, design, and implement peripheral interactions in XR.

Keywords: Extended reality, peripheral interaction, conceptual space, event-based software architecture, head-mounted displays

1 Introduction

Our everyday lives represent a never-ceasing game of interaction-attention, where information from the physical environment and the body reaches the state of conscious awareness, as it flows between the center and periphery of our attention. This intrinsic feature to human biology and cognition has been skillfully exploited in the Human-Computer Interaction (HCI) and Ambient Intelligence (AmI) communities for designing and implementing peripheral interactions in the context of calm computing [1], smart environments [2], and ambient media [3,4] delivered by ambient displays [5,6] and, recently, for virtual [7] and augmented reality [8,9] environments in the Virtual Reality (VR) community. The spectrum of possibilities for interaction design addressing various levels of human attention is known as *the interaction-attention continuum* [10]. At one end of this continuum, *focused interactions* capture the center of attention by being intentional, conscious, and under the direct, precise control of the user. At the opposite end, *implicit interactions* take place outside the attentional field, are subconscious and unintentional. Located in-between, *peripheral interactions* [2,10,11] target the periphery of attention by featuring subconscious and intentional, but also direct, yet imprecise control. These characteristics turn peripheral interactions into highly effective personal routines that seamlessly support the majority of everyday life activities [12] without overloading conscious cognitive processes [10].

Although peripheral interactions have been extensively examined in the scientific literature, both in terms of conceptual development [2,10] and practical opportunities for novel interactive systems and applications [13–16], implementations have relied so far on information that users perceive following stimuli from the physical world. However, as augmented (AR), mixed (MR), virtual (VR), and extended reality (XR) worlds become increasingly integrated in our lives, *e.g.*, the upcoming metaverse [17] with profound expected influence on human social functioning [18], and given the increasing availability of devices that support immersive user experiences [19], we believe it is high time to examine peripheral interactions for which the substratum is represented by physical-virtual environments. To the best of our knowledge, such an investigation has not been conducted so far. Nevertheless, the formalism of XR [20] as the umbrella concept for AR, MR, and VR offers the opportunity to examine peripheral interactions for a variety of hybrid environments specified along Milgram *et al.*'s [21] Reality-Virtuality Continuum.

In this context, our specific research questions are:

RQ_1. How can peripheral interaction be formalized for XR environments? We address this research question with a new conceptual space that combines Interaction-Attention [10] and Reality-Virtuality [21].

RQ_2. How do users perceive content delivered in XR at the periphery of their attention? We examine this research question with user experience measures of noticeability, appropriateness, comfortability, and usefulness.

RQ_3. What is the user performance with content delivered at the periphery of attention in XR environments compared to physical ones? We examine this research question with measures of content understanding and recall.

In line with these research questions, we make several theoretical and practical contributions, as follows:

1. We present an examination of peripheral interactions for which the substratum is represented by stimuli originating from XR environments. Specifically, we build on SAPIENS [16], an open-source software architecture for peripheral interaction in smart environments, which features dedicated software components to assist engineering of systems and applications implementing peripheral interactions in physical environments.
2. We introduce a dedicated conceptual space for peripheral interactions in XR on top of Bakker and Niemantsverdriet's [10] Interaction-Attention Continuum and Milgram *et al.*'s [21] Reality-Virtuality Continuum, and formalize the concept of an "XR display" to extend the coverage of ambient media displays [1,6] from physical environments to XR. These conceptual contributions constitute into a new framework for supporting the development of possible applications in the spectrum of XR interaction-attention.
3. To demonstrate and validate our theoretical concepts in the context of previous research in the area of peripheral interactions, we develop a software simulator, SAPIENS-IN-XR, representing an extension of SAPIENS [16]. To this end, we introduce dedicated software components for presenting information using XR displays, and we deliver our simulator online for XR researchers and practitioners; see Figure 1 for a screenshot. We also validate SAPIENS-IN-XR with a technical evaluation, where multiple XR displays are dynamically generated in our software architecture following Poisson distributions to model notification events.
4. We conduct a controlled experiment with fifteen participants using the HoloLens head-mounted display (HMD) to demonstrate one possible application enabled by our framework for interaction-attention in XR, but also to collect user experience measures. We report positive user perceptions of the noticeability, appropriateness, usefulness, and comfortability of content delivered in the form of notifications in XR at the periphery of user attention.

2 Related Work

We relate to prior work on peripheral interactions with a focus on the concept of useful field of view and software applications and tools for supporting the development of such interactions. We also present an overview of SAPIENS [16], the software architecture for peripheral interaction in smart environments on which we build to deliver SAPIENS-IN-XR. Finally, we relate to the scientific literature examining immersive user experiences in XR.

Fig. 1 Snapshot of SAPIENS-IN-XR, our online simulator for interaction-attention in XR environments. The physical devices illustrated in this figure, such as the wall display, sound surround system, and motion tracker, target various points on the Interaction-Attention Continuum [10] in the physical world. XR displays, depicted in red, are dynamically created to support peripheral interaction-attention tasks in XR. SAPIENS-IN-XR is available at http://www.eed.usv.ro/mintviz/resources/SAPIENS/in-XR.html.

2.1 Peripheral Interaction

The growing number of computer devices per user and the interactions they enable in the context of the mobile and wearable computing paradigms have led to an increased load on the users' center of attention. Since focused attention is serial in nature, only one device, application, or system at a time can effectively exploit it [22]. The next subsection introduces relevant human vision mechanisms in relation to this aspect from the perspective of the useful field of view (FOV). By sequentially shifting the center of attention among multiple devices and the interactions they enable, the illusion of task parallelism is achieved with the side effect of increasing cognitive load because of frequent context switching. An alternative solution to expand the information spectrum for the users of a smart environment, while avoiding unwanted increase in cognitive load, is for devices and interactions to target the periphery of user attention and shift to the center only when needed [11,23]. This approach, an instance of calm technology [1], has been formalized as peripheral interaction [2,11]. Examples of systems include Cow-Clock [13], FireFlies [14], Lantern [24], or Audience Silhouettes [15], to name a few, which have employed physical devices of various kinds to deliver information at the periphery of user attention. Also, a few systems have proposed peripheral interactions for AR environments [9,25,26], but such contributions have been scarce compared to the large body of work using physical devices in physical environments. We believe that this state of affairs is a direct consequence of lacking concepts and frameworks for combined attention and hybrid reality concepts, which have developed in isolation in distinct scientific communities. In Section 3, we address this aspect by proposing a conceptual space for Attention-Interaction in XR, which we instantiate and validate with our dedicated software architecture, SAPIENS-IN-XR.

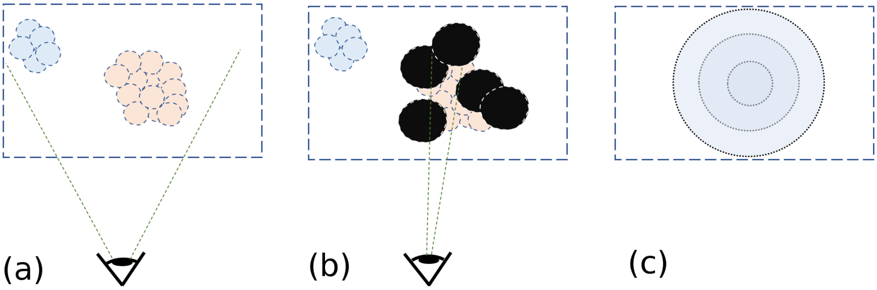

Fig. 2 The decrease in the coverage of the useful field of view, a concept from Mackworth [27], from (a) an initial large angle to (b) a considerably reduced angle because of the presence of noise. *Notes:* light orange circles (a,b) denote the central vision area, blue circles (a,b) depict the peripheral view, and black circles (b) represent noise; figure (c) depicts the central and peripheral field of views as a series of concentric circles.

2.2 Useful Field of View

According to Bakker *et al.* [2], peripheral interaction is grounded by divided attention and multitasking theory. By building on the same conceptual foundations, our approach is to investigate specific human factors involved in peripheral interaction. To this end, we draw from human vision research, which defines the Useful Field of View (Figure 2) as the area around the fixation point from which information is processed during a given visual task [27]. Based on early studies of human vision [28], it has been established that, when noise is introduced in the visual field (Figure 2c), the effective fixation area represented by the useful FOV is limited to only 2°, a phenomenon referred to as tunnel vision [27,28]. This level of performance is in stark contrast to the peripheral vision, where large objects can be detected in a FOV spanning between 50° and 90°. Consequently, as object density increases, cognitive processes must reduce the useful field of view to an area that can be processed effectively. In Section 5, we present a controlled experiment where the main task is maintaining focused attention in the central FOV, while content is displayed at the periphery of user attention with the goal to explore implications of the mechanisms of human vision for the design of peripheral interactions in XR.

2.3 Software Tools for Smart Environments

We use the term “smart environment” to denote “an ecology of interconnected smart devices that impacts users’ perception of, attention to, and interpretation of information, content, and feedback delivered by the smart environment,” following a definition from Schipor *et al.* [16, p. 5] provided in the specific context of applications and systems implementing peripheral interactions. From the wide scientific literature on smart environments [29–31], we are interested in software tools and architectures for engineering interactions and applications thereof. For instance, Caballero *et al.* [32] introduced a software architecture for emotion detection and regulation in smart environments that processes psychological signals, facial expression, and aspects of user behavior, and provides responses to regulate users' emotional states towards positive moods. HyBloSE [33] is a blockchain-based framework for smart environments that runs on low-power devices. Seraj *et al.* [34] proposed BEESM, a block-based end-user programming tool that facilitates rapid prototyping of applications for smart environments. SpaceState [35] is a system for designing spatial user interfaces that use depth cameras to react to changes in the physical layout of a room. For this purpose, SpaceState identifies the state of the physical environment of the room in real-time, which enables adaption to various room states and reaction to transitions between states. The Proximity Toolkit [36] is a framework designed for rapid prototyping of proxemic-aware interactions in smart environments. It offers out-of-the-box support for monitoring entities present in a smart environment, *e.g.*, users, devices, and non-digital things, and reports proximity information about those entities, such as orientation, distance, motion, etc. Schipor *et al.* [37] introduced Euphoria, an event-based software middleware to support development of interactions with heterogeneous devices with applications for smart environments of many kinds, from smart rooms to smart vehicles. Euphoria constituted the foundation for Sapiens [16], the software architecture for peripheral interactions in smart environments that we employ to deliver Sapiens-in-XR, our extension to cover peripheral interactions for XR environments. Due to its direct relevance to our work, we present an overview of Sapiens next.

Schipor *et al.* [16] introduced Sapiens, a specialized event-based software architecture that enables engineering of peripheral interactions in smart environments with a variety of I/O devices, from public interactive surfaces and wall displays to personal mobile and wearable devices. Sapiens capitalizes on the multimodal nature of human attention and implements a probabilistic algorithm for inferring user attention for an ecology of devices that deliver messages and notifications. To support these features, Sapiens implements event-based processing with five software components: (1) Attention-Detection in charge with estimating users' focus of attention in the physical environment, (2) Priority-Management establishes device priority for delivering notifications, (3) Interruptibility-Prediction handles external events and forwards messages to output devices, (4) Context-Awareness collects and interprets proximity-related information for devices registered in the smart environment, and (5) Device-Interchangeability enables notifications to be delivered by output devices via specific modalities. We refer readers to Schipor *et al.* [16] for technical details about the Sapiens software architecture.

2.4 Software Tools for XR environments

Several software tools have been proposed in the scientific community to support development of applications for XR environments. For example, Billinghurst and Nebeling [38] introduced a tool for rapid prototyping of XR experiences meant for a non-technical audience. Nebeling *et al.* [39] implemented XRStudio, a pipeline for delivering lectures in VR that capitalizes on immersive content, live streaming, and virtual production techniques from film making. Speicher *et al.* [40] introduced 360theater, a new method and companion tool for rapid prototyping of AR/VR experiences involving 360° videos. Pohl *et al.* [41] developed Body LayARs, a toolkit that supports development of body-based AR prototypes. Body LayARs features a graphical programming environment with a device-independent runtime abstraction. Rompapas *et al.* [42] implemented Project Esky, an open-source XR framework for high fidelity hand-based interactions with virtual content. Despite an increasing interest in software tools to assist development of XR environments and design of new XR experiences, peripheral interactions have not been addressed so far in relation to stimuli originating from XR environments. Next, we complete our discussion of the scientific literature with an overview of prior work that examined aspects of user attention in XR.

2.5 Attention State in XR environments

Attention "is used to focus the human cognitive capacity on a certain sensory input so that the brain can concentrate on processing information of interest" [43] (p. 166). Biocca *et al.* [44] introduced the "attention funnel," an omnidirectional 3D cursor to guide user attention towards objects completely outside the FOV. Bonanni *et al.* [45] used layered interfaces designed according to cueing and search principles employed by the process of attention focus to reduce user cognitive load and improve user performance with and confidence in the system. Lu *et al.* [46] investigated a subtle cueing method to support visual search in AR, which was found as effective as explicit cueing, but less distracting. Finally, Veas *et al.* [47] examined the effectiveness of directing user attention by imperceptibly modifying video features using a saliency modulation technique. Although this prior work has contributed useful results to inform the design of attention-driven user interfaces in XR environments, none addressed interactions performed at the periphery of user attention.

2.6 Summary

We overviewed prior work on peripheral interactions from the perspective of the useful FOV and software tools for smart environments. We found that, despite considerable scientific contributions in the area of peripheral interactions for physical environments, scientific research examining attention-interaction in XR has been very scarce. In this context, both conceptual developments and practical tools are needed, which we introduce next.

3 Formalization of Peripheral Interaction in XR

XR can enhance the capacity of a smart environment to support peripheral interactions in many ways, such as by compensating limitations of physical displays with flexible form factors in XR. For example, many display characteristics, such as size, location, orientation, output modality, and source of content, are essential for designing interactions across the Interaction-Attention Continuum [10], but are not flexible because of the rigid form factors of physical displays. The alternative of an "XR display" is an appealing concept to address such limitations. XR displays, unlike their physical counterparts, could be instantiated on demand, according to context and specific events, such as events generated by changes in user eye gaze, head orientation, and body behavior as well as other aspects relevant for inferring the user's center of attention, as discussed in Section 2. Moreover, the content presented on XR displays could be dynamically superimposed on and aligned with the physical environment, including physical and ambient displays. For example, virtual replicas of physical devices with which the user is already familiar, *e.g.*, a TV screen [48], are among the many form factors [49] that an XR display could take to match user needs and support more effective interactions in hybrid, physical-virtual worlds. To support such desiderata and address research question RQ_1, we introduce in this section a conceptual space for peripheral interactions in XR and formalize the concept of an "XR display".

3.1 Conceptual Space for Interaction-Attention in XR

Following the lack of a conceptual framework for the formalization and operationalization of interaction-attention in XR, we propose a two-dimensional space that draws on Bakker and Niemantsverdriet's [10] Interaction-Attention Continuum for the physical environment and Milgram *et al.*'s [21] Reality-Virtuality Continuum for mixtures of physical and virtual environments, respectively; see Figure 3. In this space, application scenarios that have been previously proposed for peripheral interaction, *e.g.*, those described in Bakker an Niemantsverdriet [10], refer to physical devices and displays and, thus, are directly contained by our conceptual space in the area delimited by the dashed line corresponding to the "Real" endpoint of the vertical axis. However, XR environments that move away from this endpoint lead to new opportunities for interactions and immersive experiences. To operationalize these opportunities for practical applications, we divide our space in five regions marked with the numbers ① to ⑤ in Figure 3. In region ①, content is delivered to the center of attention by a physical display, while an XR display, rendered for example using a pair of smartglasses, presents additional information. In region ②, an HMD with contextual understanding of the physical surroundings leverages non-digital objects from the physical environment as anchors to deliver notifications to the user's center or periphery of attention, *e.g.*, weather information displayed next to a physical window. Applications from region ③ address multiple users engaged in a collaborative task in XR, but the same message is delivered differently to each user at various points corresponding to the Interaction-Attention Continuum according to their engagement with the task or perceptual-motor abilities. In region ④, messages are displayed on a physical device from the periphery of the user's FOV and enriched with virtual content. When the physical display exits the user's FOV, the virtual content takes over the periphery of attention. A characteristic of this region is that

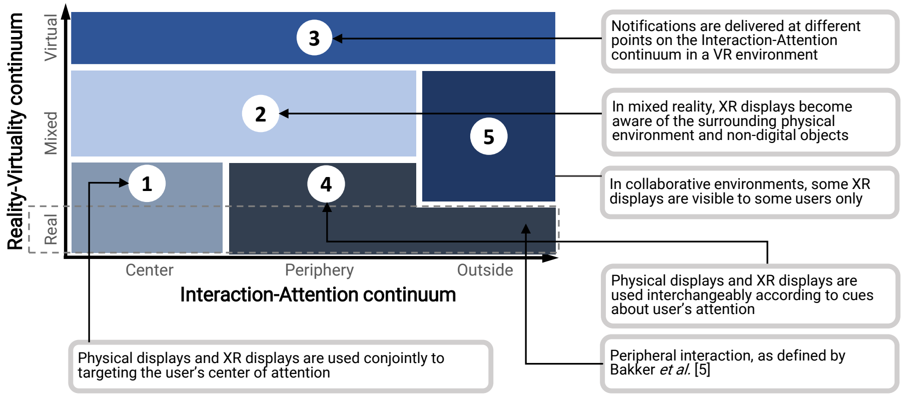

Fig. 3 A two-dimensional conceptual space for peripheral interaction in XR in the conjunction of the Interaction-Attention [10] and Reality-Virtuality [21] continua. Various applications are enabled by different regions of this space and are discussed in the text. Peripheral interactions explored in the scientific literature, *e.g.* [10], reside along the "Real" region of our conceptual space delineated by the dashed line.

physical and XR displays can be interchangeably employed by users or targeted by the application implementing peripheral interactions in XR according to specific cues about users' focused attention. Finally, region ⑤, although falling by definition outside the attentional span of the user, represents a viable option for XR displays to momentarily hide from the FOV. For example, in a multi-user collaborative application, the same content is visible to some of the users, but not to all as a direct effect of users' orientations and FOVs but also of content type, ownership, and privileges to access content. The nature of a mixed world [21], combining both physical and virtual elements, can also be explored to present content using different output modalities, representing different instances of the Reality-Virtuality Continuum. These examples show how our conceptual space enables a variety of scenarios that extend the characteristics of conventional physical displays, including the diverse form factors that ambient displays [5,6] take in physical environments. Next, we formalize the concept of a XR display.

3.2 XR Displays

The notion of an XR display becomes critical when extending Interaction-Attention in the context of the Reality-Virtuality continuum to differentiate against physical displays. To this end, an overview of display taxonomies proposed so far in the scientific literature is useful to highlight the need for and to contextualize the new concept of an XR display for peripheral interactions.

Physical displays have been characterized in the scientific literature in many ways. For example, Itoh *et al.* [19] used a classification based on three types of realism: *spatial realism* emerging from virtual images and the real world, *temporal realism* subsuming latency of many sources (tracking, application, image generation, and display latency), and *visual realism* referring to visual qualities, such as color reproduction, dynamic range, occlusions, depth, and FOV. Lantz *et al.* [50] identified three classes of displays: (1) small-scale, single-user displays (*e.g.*, HMDs or desktop stereoscopic displays), (2) medium-scale displays (*e.g.*, CAVEs and power walls), and (3) large-scale displays for immersion experiences involving groups of users (*e.g.*, IMAX and domes). Regarding large displays, Ardito *et al.* [51] identified specific display characteristics, such as *visualization technology*, *setup*, *interaction modality*, *application*, and *location*, which can be used to differentiate between different display classes. Heller *et al.* [52] examined wearable displays and identified two categories. The first category describes the *placement* of the display and specifies *location* (*e.g.*, head, waist, legs, etc.), *accessories* (*e.g.*, eyewear, rings, etc.), *clothing* (shirts and tops, notions, dresses), and *skin and body* (*e.g.*, skin, face and neck, hands, and fingers). The second category regards the properties of the content rendered on the display with the following subcharacteristics: *audience* (public, intermediate, private), *temporal aspects* (refresh rate, persistent/ephemeral) and *information density* (monochromatic, RGB elements, etc.). By examining "Fused Twins" as a step towards AR media architecture, Grübel *et al.* [53] introduced *augmentable user screens*, *i.e.*, displays defined along Milgram *et al.*'s [21] Reality-Virtuality Continuum [21] that are (i) 3D and (ii) only visible in an immersive world.

Based on this previous work, we introduce XR displays as a specific category of virtual objects rendered in XR environments that feature Itoh *et al.*'s [19] spatial and visual realism, Lantz *et al.*'s [50] multiscale and multiuser characteristics, and build on top of Grübel *et al.*'s [53] criteria for augmentable user screens. To these characteristics, we add two more that operationalize XR displays from the perspective of software architecture for peripheral interactions [16], as follows: (i) the properties (*e.g.*, size, location, orientation, etc.) of an XR display can be dynamically changed during and following events from the physical-virtual environment and user interaction, and (ii) XR displays can be created and destroyed on-the-fly in the XR environment. Property (i) specifies the high flexibility of an XR display in contrast to its physical counterpart, while property (ii) acknowledges the heterogeneity of application needs for smart environments with multiple content sources and I/O devices that demand user attention [16]. According to this definition, XR displays exist at the intersection of the physical and the virtual with convenient characteristics that make them transition fluently, in terms of form factor and behavior, between the center and periphery of user attention. With these characteristics, XR displays can cover all the regions of our conceptual space for peripheral interaction in XR illustrated in Figure 3. Next, we show how XR displays can be implemented in applications supported by SAPIENS-IN-XR, our software architecture and simulation tool for peripheral interactions in XR.

4 Sapiens-in-XR, an Event-based Software Architecture and Simulator for Peripheral Interactions in XR

We introduce SAPIENS-IN-XR, our software architecture for peripheral interactions in XR. Since SAPIENS-IN-XR is open source and freely available online as a Three.js[1] web application, we do not insist on its engineering details. Instead, we present the new, specialized software components dedicated to peripheral interaction in XR according to our conceptual space from Section 3 and report empirical results from a technical evaluation of SAPIENS-IN-XR.

4.1 The Software Architecture of Sapiens-in-XR

Figure 4 illustrates the SAPIENS-IN-XR software architecture built on top of SAPIENS [16] that, at its turn, was built on top of Euphoria [37], an event-based software architecture for implementing interactions with heterogeneous I/O devices in smart environments. On these foundations, SAPIENS-IN-XR introduces new software components specific to XR, highlighted in Figure 4:

- XR-PHYSICAL-REALITY-FUSION is a software middleware that merges information from both physical and XR displays, designed to behave as an abstract interface to I/O devices. This component communicates with SAPIENS [16] to reuse the attention-detection and context-aware features of the latter.
- The ATTENTION-DETECTION component supersedes the one from SAPIENS [16] to address aspects of user attention towards entities from the XR environment. For instance, HMDs that render virtual content integrate eye gaze and head trackers, hand and gesture detectors, and speech recognition. This information is fed into ATTENTION-DETECTION to complement the data collected by the original version from SAPIENS.
- The DEVICES layer specifies physical devices, *e.g.*, a wall display from the physical environment or the user's smartwatch and, as a particular subcategory, XR devices, *e.g.*, a pair of smartglasses for AR, next to XR displays as formalized in Section 3. While physical devices are tightly coupled with a particular adapter (platform, operating system, communication protocol, API, etc.), the latter represent software objects that expose generic interfaces; see the top layer from Figure 4. Nevertheless, the nature of the display, either physical or virtual, is transparent to the business logic of SAPIENS-IN-XR, where all of the displays are ultimately represented with abstract software objects that expose specific capabilities, *e.g.*, a fixed screen size for a physical display or multiple form factors possible for an XR display, to present information to users.
- The XR-DISPLAYS-HANDLER software component is in charge of creation, manipulation, and destruction of the software objects representing displays in the SAPIENS-IN-XR architecture. It receives input from

[1] https://threejs.org/

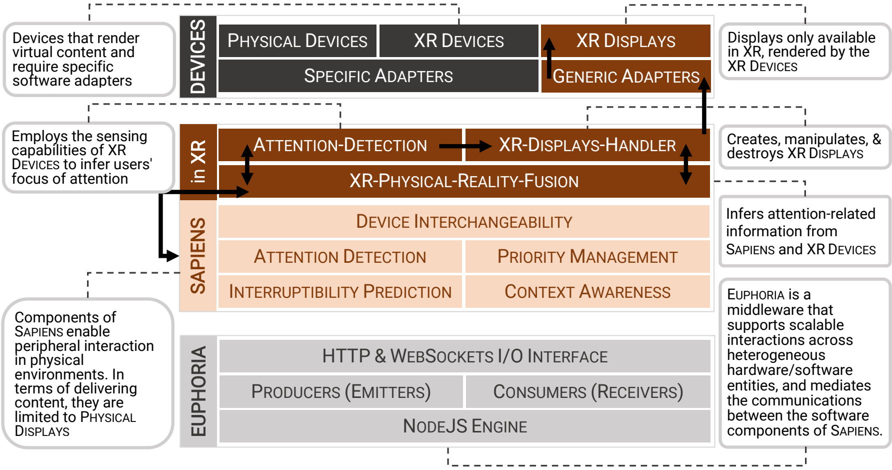

Fig. 4 Visual illustration of the SAPIENS-IN-XR software architecture (top two layers), built on top of SAPIENS [16] (middle layer) that, at its turn, builds on top of Euphoria [37] (bottom layer). The top two layers implement software components that are specific to XR displays and peripheral interactions in XR, such as adapters for devices that render XR content and modules implementing XR displays, according to our formalism. The second and third layer are shown connected since SAPIENS-IN-XR reuses components of SAPIENS.

ATTENTION-DETECTION to update XR displays. This component has a logical connection with the PRIORITY-MANAGEMENT component of SAPIENS [16], which it uses to obtain information about the notifications to be delivered and their priority.

For other technical aspects of the implementation of SAPIENS-IN-XR, we refer readers to the open-source code of our software architecture available online at http://www.eed.usv.ro/mintviz/resources/SAPIENS/in-XR.html.

4.2 Technical Performance of Sapiens-in-XR

To understand the technical performance of SAPIENS-IN-XR, we conducted a technical evaluation in which XR displays, notifications to be delivered with XR displays, and the corresponding events for creating, transmitting, and rendering notifications were processed in the SAPIENS-IN-XR architecture.

We conducted our evaluation in the form of a simulation experiment by measuring the processing time of an event, in milliseconds, from its creation (*i.e.*, new information becomes available in the XR environment to be displayed to the user) to finalization (*i.e.*, the message encapsulating the information reaches the target display). Processing time is a common measure to characterize the technical performance of an implemented software architecture for processing messages through its various layers and components [16,37] and can be evaluated in different conditions represented by different content types and sizes and environment complexity, respectively [54]. Regarding the latter, we varied the number of XR displays from 1 (corresponding to a lightweight environment with the smallest number of displays competing for the user attention) to 16 (a dense environment with many displays) in a geometric progression with the common ratio 2, *i.e.*, N=1, 2, 4, 8, and 16 displays, respectively. For each simulation condition consisting of N displays, we injected events in the Sapiens-in-XR architecture at random moments in time sampled from Poisson distributions with the rates λ=10, 5, and 1, corresponding to different expected numbers of notifications occurring over a 5-second time interval. These values describe "aggressive," highly attention demanding environments (λ=10), *e.g.*, representative for an action video game, where notifications occur very frequently, to more "calmer" environments with more time elapsing between consecutive notifications (λ=5 and λ=1, respectively). We modeled notification delivery in our simulation experiment as a Poisson process [55] since (i) events are discrete properties, (ii) events are independent, *i.e.*, the occurrence of one event does not affect the probability of a second event to occur, (iii) the average rate of events occurring is independent of any occurrences and under our control with the λ parameter, and (iv) two events cannot occur at exactly the same instant (our software architecture is not intended to be a real-time processing architecture).

Our technical evaluation was a controlled experiment that measured the response time of Sapiens-in-XR function of two factors of environment complexity, the number N of XR displays and the notification rate λ. We ran Sapiens-in-XR as a Node.js[2] server application on a Windows 10 desktop PC (3.60GHz Intel Core i9-9900KF CPU, RAM 64GB) in a 5GHz Wi-Fi network implemented with an ASUS RT-AC87U router. For each combination of $N \times \lambda$, we generated 100 events and averaged the time measured for processing those events by the software components of the Sapiens-in-XR pipeline, from the creation of the events to their delivery on a XR display. The results showed an average of 11ms processing time per event for N=1 and 2 displays, 13ms for N=4 to 8, and 18ms for N=16 displays, which represent fast response times for practical applications under a variety of simulated conditions.

4.3 Sapiens-in-XR as a Simulator for Peripheral Interactions in XR Environments

Our discussion from Section 2 revealed few studies on the user attentional state in XR, of which most [43,44,46,47] related to visual search tasks. In this context, where empirical evaluations are lacking, but the scope of peripheral interaction in XR is wide (Figure 3), tools to assist simulation of new scenarios for interaction-attention in XR are a valuable resource for researchers and practitioners. In the following, we show how our web implementation of Sapiens-in-XR can be used to simulate such scenarios.

We implemented Sapiens-in-XR on the web to enable rapid specification and exploration of diverse layouts of XR displays according to the various regions of our conceptual space from Figure 3 with the goal to inform design and implementation of peripheral interactions in XR. Figure 5 shows a few

[2] https://nodejs.org/en/

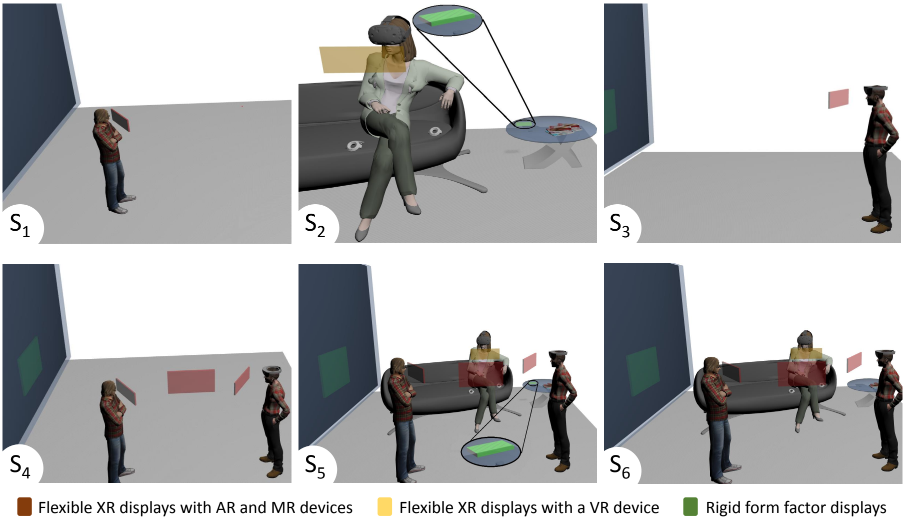

Fig. 5 Examples of application scenarios simulated using SAPIENS-IN-XR for peripheral interaction in XR with various devices and user configurations: (1) single user watching a movie in AR using a pair of smartglasses, while also receiving notifications in the glasses, (2) single user watching a movie in VR using an HMD, receives notifications in VR via the HMD, but also in the physical world via the smartphone; (3) AR movie and notifications delivered via AR glasses, and notifications on the wall display; (4) multiple users, movie and notifications in MR via the HoloLens HMD, one user also receives notifications via the wall display; (5) multiple users watch a movie in AR and VR with different devices, receive notifications in AR and VR, but also from the smartphone and wall display (6); multiple users, AR and VR movie, notifications in AR and VR, one user also receives notifications on the wall display.

examples where diverse types of XR displays are used to render content. We use notifications as the key component for exploring peripheral interaction. The XR displays that deliver notifications are represented with rectangular objects in our simulator, *e.g.*, the red display in the front of the user from Figure 5.1. Notifications posses specific properties, such as level of priority (low, medium, and high, same as in SAPIENS [16]), behavior (*e.g.*, blinking for visual notifications and volume for audio notifications), and can be activated and deactivated at will during the simulation.

Our simulator communicates with and employs the components of the SAPIENS-IN-XR software architecture via standard HTTP and WebSocket calls in the same way a real environment would in a deployed application. In fact, the SAPIENS-IN-XR architecture does not differentiate between calls that come from a simulated or actual device. This level of abstraction enables SAPIENS-IN-XR to be readily used in a variety of scenarios where the components can be software, hardware, or simulated, as in the case of our web application. The SPECIFIC and GENERIC ADAPTERS are embedded as interfaces into the software components of the simulator, which resides on the DEVICES layer of the SAPIENS-IN-XR architecture illustrated in Figure 4.

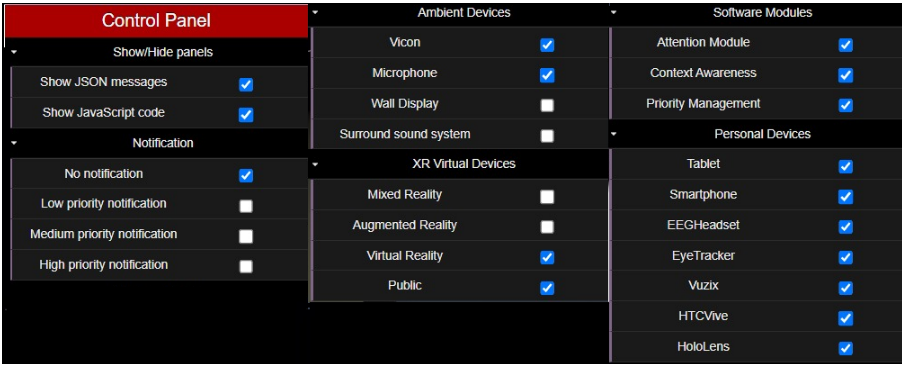

Fig. 6 The control panel of the SAPIENS-IN-XR online simulator for peripheral interactions in XR that enables specification of simulation options for the software architecture components of SAPIENS, simulated display devices, and corresponding notifications. For example, checking the "Wall Display" option enables notifications to also be shown on a physical display from the physical environment, and checking the "Priority Management" option enables the SAPIENS component for prioritizing notifications on specific devices. See Figure 1 from Section 1 for the presentation of the control panel in the SAPIENS-IN-XR simulator.

To control the simulation, we implemented a Control Panel with check boxes corresponding to the following categories of elements that can be monitored during a simulation of peripheral interaction in XR: panels, notifications, software modules, personal devices, ambient devices, and XR devices; see Figure 6. The first category, panels, enables visualization of JSON messages exchanged between the different software components of SAPIENS-IN-XR and a visualization of the corresponding JavaScript code, respectively. The second category, notifications, enables simulation of different priority levels (low, medium, and high) for the content delivered to users, *i.e.*, higher priority notifications are delivered by devices located more closely to the center of the FOV and by more devices, respectively, following [16]. The third category, software modules, is used to activate/deactivate specific modules of the SAPIENS-IN-XR architecture, such as ATTENTION-DETECTION, CONTEXT AWARENESS, and PRIORITY MANAGEMENT illustrated in Figure 4, to understand their effect on notification delivery. The remaining categories from the Control Panel present different types of display devices, from personal to ambient to XR displays. Examples include personal device categories already available in SAPIENS [16] (tablet, smartphone, electroencephalography (EEG) headset, and eye tracker), and new categories representative of XR personal displays, according to our definition from Section 3: AR, MR, and VR glasses and HMDs. The ambient device category includes the ambient displays from SAPIENS [16].

4.4 Examples of Simulations with Sapiens-in-XR

We exemplify in the following a diversity of scenarios simulated using the web implementation of SAPIENS-IN-XR for peripheral interaction using XR displays of many kinds; see Figure 5 for snapshots captured using our simulator. From left to right and top to bottom, we present six simulation scenarios (S_1

to S_6), of increasing complexity in terms of the number of involved users and types of display devices:

S_1) *Single user, single display, AR.* In this scenario, Michael watches a movie using a pair of light AR smartglasses, such as Vuzix Blade. Occasionally, notifications are displayed at the periphery of his FOV.

S_2) *Single user, multiple displays, VR and physical world.* In this scenario, Sandra watches a movie using the HTC Vive HMD and occasionally receives notifications rendered in the HMD at the periphery of her FOV. Also, high-priority notifications are enabled on the smartphone located on the table. In this scenario, notifications came from two personal devices and require attention switching between the physical and virtual world.

S_3) *Single user, multiple displays, MR.* Matthew watches a movie using the HoloLens HMD and receives notifications both on the HMD and the wall display at the periphery of his FOV. Just like in the previous example, notifications came from two devices, but there is no switching of attention between the two worlds since the world configuration is mixed.

S_4) *Multiple users, AR and MR devices.* In this scenario, Michael and Matthew watch the same movie, but on two different devices: Michael uses the light AR smartglasses as in scenario S_1 and Matthew the HoloLens HMD as in S_3. They receive notifications on these devices, but Matthew can also observe notifications on the wall display from the room.

S_5) *Multiple users, AR, MR, and VR devices.* Michael, Sandra, and Matthew watch the same movie using different devices: lightweight AR smartglasses, VR HMD, and HoloLens HMD, respectively, as in the scenarios S_1, S_2, and S_3. They receive notifications on these devices. Matthew can also see notifications delivered on the wall display, and Sandra receives notifications on her smartphone.

S_6) *Multiple users, AR, MR, and VR devices (variant).* This scenario is similar to S_5 with the main difference that Sandra can also receive notifications from the wall display in augmented virtuality [21]. To consume a wall display notification, her VR HMD renders a part of the physical world showing the wall display.

These examples illustrate application scenarios of various complexities that bring about various nuances of implementing peripheral interactions in XR. Scenario S_1 explores the interaction between one user and a personal device capable of rendering lightweight AR content in 2D, such as the Vuzix Blade smartglasses. Scenario S_2 moves the interaction and consumption of virtual content in VR, but also allows notifications from the smartphone, a personal device representative of interactions in the physical world. However, Sandra, the protagonist of the second scenario, has to actively switch her attention between the physical and virtual environments to consume notifications delivered independently in the two worlds. Scenario S_3 eliminates this drawback by replacing the VR HMD with the HoloLens headset, a MR device. Moreover, reading messages from the wall display, an ambient device representative of physical smart environments, becomes possible for the MR user. Scenario

S_4 scenario is a combination of S_1 and S_3, S_5 a combination of S_1, S_2, and S_3, and S_6 a variation of S_5 with multiple users. In these scenarios, different users employ different personal display devices with different capabilities for presenting virtual content in AR, MR, and VR, and some of the users can also access notifications delivered by physical displays.

5 Experiment

Simulations represent an useful intermediate step to understand the complexity of implementing peripheral interactions in XR environments enabled by a diversity of display devices and supporting various kinds of XR displays, respectively, but also to inform design requirements of the corresponding applications and systems implementing those interactions. However, to evaluate and understand aspects of end-user perception (according to our research question RQ_2) and user performance (research question RQ_3) with peripheral interactions in XR, user studies are in order. Following our approach based on the concept of the useful FOV (see Section 2), we implemented a user study in the form of a controlled experiment with the main task represented by users maintaining the focus of attention in the central FOV, while notifications are displayed at the periphery of their attention. To this end, we selected one of the scenarios exemplified in Subsection 4.4 to implement in our experiment. We preferred scenario S_3, where the user wears a MR HMD and, thus, has access to notifications rendered both in the HMD and on ambient displays from the physical environment since it can be easily implemented with just one personal device giving access to both the physical and virtual worlds and one ambient display, respectively.

5.1 Participants

Fifteen participants (11 men and 4 women), representing young adults between 18 and 29 years old (mean M=20.5, standard deviation SD=3.0 years), participated in our study. Their self-reported average daily time dedicated to watching TV and video streaming platforms was between one and seven hours (M=3.5, SD=2.0), and the most preferred TV genres were movies (86.7%), comedy/sitcoms (53.3%), science and technology shows (46.7%), documentaries (40%), and game shows (40%), respectively. In terms of their previous experience with XR devices, five of the participants (33.3%) reported having used HMDs before our study.

5.2 Apparatus and Procedure

We developed a HoloLens application presenting a 21-minute long movie,[3] and created a set of sixteen text notifications to be delivered at specific times during the movie. All the notifications contained between eight and sixteen

[3] The movie "40 days" was produced by Imagine8 Series and is available from the web address https://www.youtube.com/watch?v=R4FRlQ1GujU.

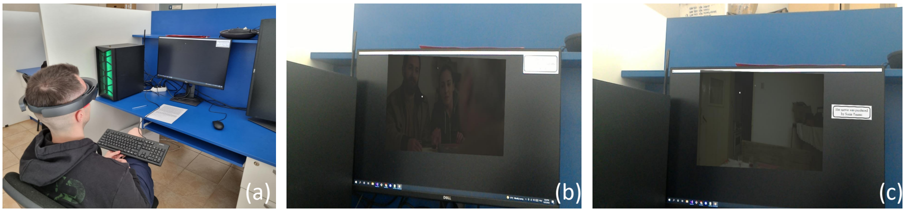

Fig. 7 Apparatus used in our experiment conducted to examine user perception of and performance with peripheral interactions in XR: (a) a participant watches a movie using the HoloLens HMD, while notifications are delivered on a physical display (b) or inside the HMD (c), two conditions for which we examine notifications addressing peripheral attention.

words that were rendered on two or three lines of text using a black font on a white background, and were related to the movie. Half of the notifications was rendered in the HMD, next to the movie, at the attention periphery; see Figure 7c. The other half was displayed on a physical screen connected to a Windows PC, on which another application was running; see Figure 7b. We controlled the environment (position of the chair, the physical screen, and orientation of the user during the experiment) so that the notifications displayed by HoloLens were approximately aligned with those displayed on the physical screen when the participant was facing forward, as illustrated in Figure 7a. Participants were instructed to press the Space key on a keyboard as soon as they observed a notification. The order of the notifications was randomized per participant and type of display, HMD or physical screen. Notifications lasted for eight or sixteen seconds, depending on the length of their text.[4] Before the start of the experiment, participants signed a consent form and filled out a questionnaire with demographic information. A post-experiment questionnaire collected several measures of user experience presented in Subsection 5.4.

5.3 Design

Our experiment was a within-subject design with one independent nominal variable, Display-Type, with two conditions (*virtual* and *physical*) and five dependent variables: Recall-Rate, Noticeability, Appropriateness, Comfortability, and Usefulness, described in Subsection 5.4.

5.4 Measures

We used a post-experiment questionnaire to collect various measures of user experience and attention representing the dependent variables in our experiment. The measures capture aspects of the benefits subjectively perceived by our participants, *e.g.*, ease of noticing notifications displayed in the virtual and physical worlds, among others, collected with 7-point Likert scales with items ranging from 1 ("strongly disagree") to 7 ("strongly agree"), as follows:

[4] Hsieh *et al.* [56] employed a duration of twelve seconds in their study on messaging notifications in VR. Because of the different lengths of our notifications, short and long, we used eight and sixteen seconds, respectively, with an average that matches the duration of twelve seconds used by Hsieh *et al.* [56].

- NOTICEABILITY, representing the participants' reaction to the statement "Notifications displayed in the virtual world were easy to see."
- APPROPRIATENESS, corresponding to the statement "Displaying notifications in the virtual world was appropriate."
- COMFORTABILITY, corresponding to the statement "Reading notifications in the virtual world was comfortable."
- USEFULNESS, corresponding to the statement "Receiving notifications in the virtual world was useful."

To evaluate the effectiveness of notification delivery, we also measured RECALL-RATE by asking participants questions about the content of the notifications and evaluating the correctness of their answers. As a complementary measure of recall, we evaluated NUM-OBSERVED, where we asked participants to state how many notifications they had observed in both the physical and virtual environments. We also adopted the reaction time measure employed by prior work that evaluated user attention in XR [43,44,46], defined as follows:

- REACTION-TIME represents the average duration, in milliseconds, needed by our participants to confirm, with the press of the Space key, that they observed a notification.

Besides these measures addressing specific aspects of user perception and performance, we also administered the following generic tests to understand the overall user experience of peripheral notifications in XR:

- TASK-LOAD, measured with the NASA TLX test [57][5] to collect participants' subjective ratings of perceived workload on six dimensions (mental demand, physical demand, temporal demand, performance, effort, and frustration). TLX returns a score from 0 (low) to 100 (high workload).
- USABILITY, measured with the System Usability Scale [58]. SUS consists of ten statements that elicit the degree of agreement using 5-point Likert scales with items from 1 ("strongly disagree") to 5 ("strongly agree"). Answers are aggregated into a score between 0 (low) and 100 (perfect).
- PRESENCE, measured with Witmer and Singer's [59] presence questionnaire (PQ v2.0) consisting of 32 questions, of which we used just the 20 questions corresponding to the *involvement/control*, *natural*, *interface quality*, and *auditory* dimensions that were relevant to the scope of our experiment.[6] We normalized PRESENCE by dividing the sum of the ratings, minus 20, to (140−20), where 140 represents the largest possible score obtained with 20 questions and a maximum rating of 7 per question, and 20 the smallest possible score corresponding to all of the ratings being 1. By multiplying the result with 100, we report values of perceived PRESENCE in the interval from 0 (low) to 100 (high), similar to the TASK-LOAD and USABILITY dependent variables, respectively.
- IMMERSION, measured with Rigby *et al.*'s [60] Immersive Experience Questionnaire for Film and TV (Film IEQ) for evaluating video viewing experiences in immersive environments. The test consists of 24 questions,

[5] We implemented the test using www.keithv.com/software/nasatlx/nasatlx.html.
[6] We removed the questions referring to the *haptic* dimension.

e.g., "To what extent did the movie, TV show, or clip hold your attention?," evaluated with 7-point Likert scales with items from 1 ("very little") to 7 ("very much"). We normalized IMMERSION by dividing the sum of the ratings, minus 24, to (168-24), where 168 represents the largest possible score obtained with 24 questions and a maximum rating of 7 per question, and 24 the smallest possible score corresponding to all of the ratings being 1. By multiplying the result with 100, we report IMMERSION values between 0 (low) to 100 (high), respectively.

5.5 Results

5.5.1 User Perception of Notifications

We found positive perceptions (mean ratings above 5.0 on a scale from 1 to 7) of the NOTICEABILITY, APPROPRIATENESS, USEFULNESS, and COMFORTABILITY of notifications displayed in the virtual and physical environment, the two conditions of the DISPLAY-TYPE independent variable. Mean NOTICEABILITY was 5.7 (SD=1.2) for notifications rendered in the virtual environment and 5.8 (SD=1.4) for those delivered on the physical screen, respectively ($Z{=}-0.180$, $p{=}.857$, *n.s.*). Mean APPROPRIATENESS was 5.1 (SD=1.6) for the virtual and 5.3 (SD=1.2) for the physical environment ($Z{=}-0.209$, $p{=}.835$, *n.s.*). Regarding COMFORTABILITY, we found a mean rating of 6.0 (SD=1.1 and 0.9) for each condition of DISPLAY-TYPE ($Z{=}-0.137$, $p{=}.891$, *n.s.*). Finally, USEFULNESS was rated at 5.8 (SD=1.3) and 5.9 (SD=0.8), respectively ($Z{=}-0.142$, $p{=}.887$, *n.s.*). These results indicate similar user perceptions of notifications delivered using the HMD in the virtual environment and the conventional display in the physical world.

5.5.2 User Performance with Notifications

We evaluated RECALL-RATE and NUM-OBSERVED as two measures of recalling the content presented using notifications. On average, participants observed 12.0 notifications (SD=3.3) out of the total number of 16 that were presented during the experiment, with no statistically significant effect of DISPLAY-TYPE ($Z{=}-1.206$, $p{=}.228$, *n.s.*). The mean RECALL-RATE was 86.7% (SD=24.6%) for notifications displayed in the virtual environment and 77.9% (SD=24.2%) for those displayed on the physical screen with no significant effect of DISPLAY-TYPE ($Z{=}-0.640$, $p{=}.522$, *n.s.*). To understand more about user performance, we analyzed REACTION-TIME and found 2.47s (SD=1.52s) and 3.02s (SD=1.56s) for notifications delivered in the virtual and physical environment, respectively, with a significant effect of DISPLAY-TYPE ($t_{(14)}{=}2.456$, $p{<}.05$) and a medium effect size (Cohen's $d{=}.634$). These results indicate the same level of user performance in terms of understanding and recalling content delivered by notifications in either the virtual and physical environments, but our participants were approximately 20% faster at noticing and reacting to notifications delivered in the HMD, where the movie was also presented.

5.5.3 Perceived Usability and Experience

The average SUS was 79.3 (SD=12.8), a high value falling under the forth quartile, the "acceptable" range, and between the "good" and "excellent" levels of usability, according to the interpretation recommendations of Bangor *et al.* [61]. Perceived Task-Load was overall low (M=38.1, SD=15.6) with the lowest score obtained on the *performance* subscale (M=24.3) and the highest on *mental demand* (M=53.3). To put the TLX results into perspective, our scores fall in the second quartile of a distribution of over one thousand TLX scores reported in academic publications [62], in the second quartile for computer activities, the first quartile for video game tasks, and the first quartile for cognitive tasks, respectively, denoting an overall low to medium perceived workload; see Grier's [62] meta-analysis of NASA TLX scores. The average perceived Immersion was 61.7 (SD=8.8), situated in the second part of the measuring scale towards high immersion. The Immersion subscales revealed high *captivation* (M=69.6), representing viewers' enjoyment, how interested they were, and their motivation during the experiment, and high *comprehension* (M=66.7), *i.e.*, how well the concepts and themes of the presented content were understood by the participants. The *real-word dissociation* and *transportation* subscales of the Immersion experience revealed lower scores (M=41.5 and 50.9, respectively), indicating that our participants were aware of the real-world surroundings with moderate feelings of being located in the augmented world. The average Presence was 68.8 (SD=8.2), situated in the second part of the measuring scale towards high levels of perceived presence. We also found a relatively high perception of the *auditory* dimension (M=83.3) of the virtual world, of the *involvement* with the virtual world (M=68.1), and of the extent to which interactions felt *natural* and the virtual world consistent with the physical reality (M=72.2). The *interface quality* subscale, representing the extent to which our participants were able to concentrate on the task of the experiment and were not distracted by controls/displays, indicated a good overall experience (M=69.3).

5.5.4 Summary

Our empirical findings revealed positive user perception of notifications addressing peripheral attention in both the virtual and physical environments as well as similar user performance with respect to understanding and recalling the content of those notifications. The only statistically significant difference observed between the physical and virtual world conditions was represented by our participants having a faster reaction time when notifications were delivered in the virtual world, where the movie was equally presented, a finding that can be explained by lower cognitive demands from not having to manage focused and peripheral attention between the two worlds.

6 Limitations

There are several limitations to our technical contribution represented by the Sapiens-in-XR software architecture. First, when extending Sapiens to Sapiens-in-XR, we implemented specific adapters (Figure 4, top) for just a few devices (Vuzix Blade, HTC Vive, and HoloLens), but interesting future work is the development of a repository for the community to share adaptation modules for other XR devices. Second, the XR-Display-Handler and XR-Physical-Reality-Fusion components of Sapiens-in-XR were instantiated in our simulator on the same platform, but it would be useful to evaluate the impact on the technical performance of Sapiens-in-XR when employing micro-services as containers for software modules and components that run on different platforms. Finally, the current design and implementation of Sapiens-in-XR does not include any module dedicated to the anticipation or prediction of the user experience in specific application scenarios of peripheral interaction in XR environments, although such a module would be very useful to practitioners during simulations.

7 Conclusion and Future Work

We introduced Sapiens-in-XR, an event-based software architecture for applications implementing peripheral interactions in XR environments. To this end, we formalized the notion of XR displays, and presented a conceptual space for peripheral interactions in XR addressing multiple levels of user attention. The results of a preliminary technical evaluation of Sapiens-in-XR indicated fast processing times of the events generated and consumed in our software architecture. We also discussed possible simulations enabled by Sapiens-in-XR, of which we implemented one in the form of a controlled experiment to evaluate user perception of and performance with peripheral interactions in XR. Our results showed similar perceptions of notifications delivered by the HMD in the virtual environment and a conventional screen in the physical world and the same level of user performance at understanding and recalling the presented content. By adopting the open-source approach of Sapiens [16], we also release our implementation in the scientific community to foster more work in the area of peripheral interactions in XR. Future work will consist in evaluating user performance for other application scenarios informed by our conceptual space involving different types of XR displays and mixtures of displays from the physical and virtual environments. Understanding the user experience of peripheral interactions in various types of XR environments enabled by various XR devices represents another opportunity to increase our understanding and design knowledge of peripheral interactions in XR.

Acknowledgments. This work was supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI-UEFISCDI, project number PN-III-P4-ID-PCE-2020-0434 (PCE29/2021), within PNCDI III. The authors acknowledge support from OSF Global Services, the Mobile Division, Suceava, which kindly provided the HoloLens HMD used for the experiment reported in this work.

Conflict of interest. The authors declare that they have no conflict of interest.

Data availability. The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

[1] Weiser, M., Brown, J.S.: The Coming Age of Calm Technolgy, pp. 75–85. Copernicus, USA (1997). https://doi.org/10.5555/504928.504934

[2] Bakker, S., Hoven, E., Eggen, B.: Peripheral interaction: Characteristics and considerations. Personal Ubiquitous Comput. **19**(1), 239–254 (2015). https://doi.org/10.1007/s00779-014-0775-2

[3] Pogorelc, B., Vatavu, R.-D., Lugmayr, A., Stockleben, B., Risse, T., Kaario, J., Lomonaco, E.C., Gams, M.: Semantic ambient media: From ambient advertising to ambient-assisted living. Multimedia Tools and Applications **58**(2), 399–425 (2012). https://doi.org/10.1007/s11042-011-0917-8

[4] Schmidt, A., Gellersen, H.-W., Beigl, M.: Matching information and ambient media. In: Proceedings of the 2nd International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture. CoBuild '99, pp. 140–149. Springer, Berlin, Heidelberg (1999). https://doi.org/10.1007/10705432_13

[5] Wisneski, C., Ishii, H., Dahley, A., Gorbet, M.G., Brave, S., Ullmer, B., Yarin, P.: Ambient displays: Turning architectural space into an interface between people and digital information. In: Proceedings of the 1st Int. Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture. CoBuild '98, pp. 22–32. Springer, Berlin, Heidelberg (1998). https://doi.org/10.1007/3-540-69706-3_4

[6] Ishii, H., Wisneski, C., Brave, S., Dahley, A., Gorbet, M., Ullmer, B., Yarin, P.: AmbientROOM: Integrating Ambient Media with Architectural Space. In: Proceedings of the CHI 98 Conference Summary on Human Factors in Computing Systems. CHI '98, pp. 173–174. ACM, New York, NY, USA (1998). https://doi.org/10.1145/286498.286652

[7] Endo, I., Takashima, K., Inoue, M., Fujita, K., Kiyokawa, K., Kitamura, Y.: ModularHMD: A Reconfigurable Mobile Head-Mounted Display Enabling Ad-Hoc Peripheral Interactions with the Real World. In: Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology. UIST '21, pp. 100–117. ACM, New York, NY, USA (2021). https://doi.org/10.1145/3472749.3474738

[8] Richards, K., Mahalanobis, N., Kim, K., Schubert, R., Lee, M., Daher, S., Norouzi, N., Hochreiter, J., Bruder, G., Welch, G.: Analysis of peripheral vision and vibrotactile feedback during proximal search tasks with dynamic virtual entities in augmented reality. In: Proceedings of the Symposium on Spatial User Interaction. SUI '19. ACM, New York, NY, USA (2019). https://doi.org/10.1145/3357251.3357585

[9] Schipor, O.-A., Vatavu, R.-D., Wu, W.: Integrating peripheral interaction into augmented reality applications. In: Proceedings of the IEEE Int. Symposium on Mixed and Augmented Reality Adjunct, pp. 358–359 (2019). https://doi.org/10.1109/ISMAR-Adjunct.2019.00-12

[10] Bakker, S., Niemantsverdriet, K.: The Interaction-Attention Continuum: Considering Various Levels of Human Attention in Interaction Design. International Journal of Design **10**(2), 1–14 (2016)

[11] Bakker, S.: Design for peripheral interaction. PhD thesis, Eindhoven University of Technology (2013). https://doi.org/10.6100/IR754544

[12] Tolmie, P., Pycock, J., Diggins, T., MacLean, A., Karsenty, A.: Unremarkable computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '02, pp. 399–406. ACM, New York, NY, USA (2002). https://doi.org/10.1145/503376.503448

[13] Bakker, S., van den Hoven, E., Eggen, B., Overbeeke, K.: Exploring peripheral interaction design for primary school teachers. In: Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction. TEI '12, pp. 245–252. ACM, New York, NY, USA (2012). https://doi.org/10.1145/2148131.2148184

[14] Bakker, S., van den Hoven, E., Eggen, B.: Fireflies: Physical peripheral interaction design for the everyday routine of primary school teachers. In: Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction. TEI '13, pp. 57–64. ACM, New York, NY, USA (2013). https://doi.org/10.1145/2460625.2460634

[15] Vatavu, R.-D.: Audience silhouettes: Peripheral awareness of synchronous audience kinesics for social television. In: Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video. TVX '15, pp. 13–22. ACM, New York, NY, USA (2015). https://doi.org/10.1145/2745197.2745207

[16] Schipor, O.-A., Vatavu, R.-D., Wu, W.: Sapiens: Towards software architecture to support peripheral interaction in smart environments. Proc. ACM Hum.-Comput. Interact. **3**(EICS) (2019). https://doi.org/10.1145/3331153

[17] Dionisio, J.D.N., Burns, W.G.I., Gilbert, R.: 3D Virtual Worlds and the Metaverse: Current Status and Future Possibilities. ACM Comput. Surv. **45**(3) (2013). https://doi.org/10.1145/2480741.2480751

[18] Xu, J., Papangelis, K., Dunham, J., Goncalves, J., LaLone, N.J., Chamberlain, A., Lykourentzou, I., Vinella, F.L., Schwartz, D.I.: Metaverse:

The vision for the future. In: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems. CHI EA '22. ACM, New York, NY, USA (2022). https://doi.org/10.1145/3491101.3516399

[19] Itoh, Y., Langlotz, T., Sutton, J., Plopski, A.: Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays. ACM Comput. Surv. **54**(6) (2021). https://doi.org/10.1145/3453157

[20] Marr, B.: What Is Extended Reality Technology? A Simple Explanation For Anyone. [Last accessed on December 2021] (2019). https://www.forbes.com/sites/bernardmarr/2019/08/12/what-is-extended-reality-technology-a-simple-explanation-for-anyone/#6ffa4ba37249

[21] Milgram, P., Takemura, H., Utsumi, A., Kishino, F.: Augmented reality: A class of displays on the reality-virtuality continuum. In: Telemanipulator and Telepresence Technologies, vol. 2351, pp. 282–292 (1995). International Society for Optics and Photonics. https://doi.org/10.1117/12.197321

[22] Fischer, R., Plessow, F.: Efficient multitasking: parallel versus serial processing of multiple tasks. Frontiers in Psychology **6**, 1366 (2015). https://doi.org/10.3389/fpsyg.2015.01366

[23] Vermeulen, J., Houben, S., Marquardt, N.: Fluent transitions between focused and peripheral interaction in proxemic interactions. In: Peripheral Interaction, pp. 137–163. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-29523-7_7

[24] Alavi, H.S., Dillenbourg, P.: An ambient awareness tool for supporting supervised collaborative problem solving. IEEE Transactions on Learning Technologies **5**(3), 264–274 (2012). https://doi.org/10.1109/TLT.2012.7

[25] Renner, P., Pfeiffer, T.: Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In: Proceedings of the IEEE Symposium on 3D User Interfaces. 3DUI '17, pp. 186–194 (2017). https://doi.org/10.1109/3DUI.2017.7893338

[26] Ku, P.-S., Lin, Y.-C., Peng, Y.-H., Chen, M.Y.: PeriText: Utilizing Peripheral Vision for Reading Text on Augmented Reality Smart Glasses. In: Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces, pp. 630–635 (2019). https://doi.org/10.1109/VR.2019.8798065

[27] Mackworth, N.H.: Stimulus density limits the useful field of view. Eye movements and psychological processes, 307–321 (1976)

[28] Mackworth, N.H.: Visual noise causes tunnel vision. Psychonomic science **3**(1), 67–68 (1965)

[29] Cook, D.J., Das, S.K.: How Smart Are Our Environments? An Updated Look at the State of the Art. Pervasive Mob. Comput. **3**(2), 53–73 (2007). https://doi.org/10.1016/j.pmcj.2006.12.001

[30] Dunne, R., Morris, T., Harper, S.: A survey of ambient intelligence. ACM Comput. Surv. **54**(4) (2021). https://doi.org/10.1145/3447242

[31] Preuveneers, D., Novais, P.: A survey of software engineering best practices for the development of smart applications in ambient intelligence. J. Ambient Intell. Smart Environ. **4**(3), 149–162 (2012). https://doi.org/10.3233/AIS-2012-0150

[32] Fernndez-Caballero, A., Martnez-Rodrigo, A., Pastor, J.M., Castillo, J.C., Lozano-Monasor, E., Lpez, M.T., Zangrniz, R., Latorre, J.M., Fernndez-Sotos, A.: Smart environment architecture for emotion detection and regulation. J. of Biomedical Informatics **64**(C), 55–73 (2016). https://doi.org/10.1016/j.jbi.2016.09.015

[33] Maselli, G., Piva, M., Restuccia, F.: Hyblose: Hybrid blockchain for secure-by-design smart environments. In: Proceedings of the 3rd Workshop on Cryptocurrencies and Blockchains for Distributed Systems. CryBlock '20, pp. 23–28. ACM, New York, NY, USA (2020). https://doi.org/10.1145/3410699.3413793

[34] Seraj, M., Autexier, S., Janssen, J.: BEESM, a Block-Based Educational Programming Tool for End Users. In: Proceedings of the 10th Nordic Conference on Human-Computer Interaction. NordiCHI '18, pp. 886–891. ACM, New York, NY, USA (2018). https://doi.org/10.1145/3240167.3240239

[35] Fender, A., Müller, J.: SpaceState: Ad-Hoc Definition and Recognition of Hierarchical Room States for Smart Environments. In: Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces. ISS '19, pp. 303–314. ACM, New York, NY, USA (2019). https://doi.org/10.1145/3343055.3359715

[36] Marquardt, N., Diaz-Marino, R., Boring, S., Greenberg, S.: The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11, pp. 315–326. ACM, New York, NY, USA (2011). https://doi.org/10.1145/2047196.2047238

[37] Schipor, O.-A., Vatavu, R.-D., Vanderdonckt, J.: Euphoria: A scalable, event-driven architecture for designing interactions across heterogeneous devices in smart environments. Inf. and Software Technology **109**, 43–59 (2019). https://doi.org/10.1016/j.infsof.2019.01.006

[38] Billinghurst, M., Nebeling, M.: Rapid Prototyping of XR Experiences. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA (2021). https://doi.org/10.1145/3411763.3445002

[39] Nebeling, M., Rajaram, S., Wu, L., Cheng, Y., Herskovitz, J.: Xrstudio: A virtual production and live streaming system for immersive instructional experiences. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA (2021).

https://doi.org/10.1145/3411764.3445323

[40] Speicher, M., Lewis, K., Nebeling, M.: Designers, the Stage Is Yours! Medium-Fidelity Prototyping of Augmented & Virtual Reality Interfaces with 360theater. Proc. ACM Hum.-Comput. Interact. **5**(EICS) (2021). https://doi.org/10.1145/3461727

[41] Pohl, H., Dalsgaard, T.-S., Krasniqi, V., Hornbæk, K.: Body LayARs: A Toolkit for Body-Based Augmented Reality. In: Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology. VRST '20. ACM, New York, NY, USA (2020). https://doi.org/10.1145/3385956.3418946

[42] Rompapas, D., Rodda, C., Brown, B.C., Zerkin, N.B., Cassinelli, A.: Project Esky: Enabling High Fidelity Extended Reality on an Open Source Platform. In: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA (2021). https://doi.org/10.1145/3411763.3451535

[43] Biocca, F., Owen, C., Tang, A., Bohil, C.: Attention issues in spatial information systems: Directing mobile users' visual attention using augmented reality. Journal of Management Information Systems **23**(4), 163–184 (2007). https://doi.org/10.2753/MIS0742-1222230408

[44] Biocca, F., Tang, A., Owen, C., Xiao, F.: Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '06, pp. 1115–1122. ACM, New York, NY, USA (2006). https://doi.org/10.1145/1124772.1124939

[45] Bonanni, L., Lee, C.-H., Selker, T.: Attention-based design of augmented reality interfaces. In: CHI '05 Extended Abstracts on Human Factors in Computing Systems. CHI EA '05, pp. 1228–1231. ACM, New York, NY, USA (2005). https://doi.org/10.1145/1056808.1056883

[46] Lu, W., Duh, B.-L.H., Feiner, S.: Subtle cueing for visual search in augmented reality. In: Proceedings of the IEEE International Symposium on Mixed and Augmented Reality. ISMAR '12, pp. 161–166 (2012). https://doi.org/10.1109/ISMAR.2012.6402553

[47] Veas, E.E., Mendez, E., Feiner, S.K., Schmalstieg, D.: Directing attention and influencing memory with visual saliency modulation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '11, pp. 1471–1480. ACM, New York, NY, USA (2011). https://doi.org/10.1145/1978942.1979158

[48] Vatavu, R.-D., Saeghe, P., Chambel, T., Vinayagamoorthy, V., Ursu, M.F.: Conceptualizing augmented reality television for the living room. In: Proceedings of the ACM International Conference on Interactive Media Experiences. IMX '20, pp. 1–12. ACM, New York, NY, USA (2020). https://doi.org/10.1145/3391614.3393660

[49] Pamparău, C., Vatavu, R.-D.: The user experience of journeys in the

realm of augmented reality television. In: ACM International Conference on Interactive Media Experiences. IMX '22, pp. 161–174. ACM, New York, NY, USA (2022). https://doi.org/10.1145/3505284.3529969

[50] Lantz, E.: A survey of large-scale immersive displays. In: Proceedings of the 2007 Workshop on Emerging Displays Technologies: Images and Beyond: The Future of Displays and Interacton. EDT '07, p. 1. ACM, New York, NY, USA (2007). https://doi.org/10.1145/1278240.1278241

[51] Ardito, C., Buono, P., Costabile, M.F., Desolda, G.: Interaction with large displays: A survey. ACM Comput. Surv. **47**(3) (2015). https://doi.org/10.1145/2682623

[52] Heller, F., Todi, K., Luyten, K.: An interactive design space for wearable displays. In: Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction. ACM, New York, NY, USA (2021). https://doi.org/10.1145/3447526.3472034

[53] Grübel, J., Gath-Morad, M., Aguilar, L., Thrash, T., Sumner, R.W., Hölscher, C., Schinazi, V.: Fused Twins: A Cognitive Approach to Augmented Reality Media Architecture. In: Media Architecture Biennale 20. MAB20, pp. 215–220. ACM, New York, NY, USA (2021). https://doi.org/10.1145/3469410.3469435

[54] Schipor, O.-A., Vatavu, R.-D.: Empirical results for high-definition video and augmented reality content delivery in hyper-connected cars. Interacting with Computers **33**(1), 3–16 (2021). https://doi.org/10.1093/iwcomp/iwaa025

[55] Katti, S.K., Rao, A.V.: Handbook of the Poisson Distribution. Technometrics **10**(2), 412–412 (1968). https://doi.org/10.1080/00401706.1968.10490580

[56] Hsieh, C.-Y., Chiang, Y.-S., Chiu, H.-Y., Chang, Y.-J.: Bridging the Virtual and Real Worlds: A Preliminary Study of Messaging Notifications in Virtual Reality, pp. 1–14. ACM, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376228

[57] Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. Advances in Psychology **52**, 139–183 (1988). https://doi.org/10.1016/S0166-4115(08)62386-9

[58] Brooke, J.: SUS: A Quick and Dirty Usability Scale. In: Jordan, P.W., Thomas, B., McClelland, I.L., Weerdmeester, B. (eds.) Usability Evaluation in Industry, pp. 189–194. CRC Press, London, UK (1996). https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke

[59] Witmer, B.G., Singer, M.J.: Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators and Virtual Environments **7**(3), 225–240 (1998). https://doi.org/10.1162/105474698565686

[60] Rigby, J.M., Brumby, D.P., Gould, S.J.J., Cox, A.L.: Development of a Questionnaire to Measure Immersion in Video Media: The Film IEQ. In: Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video. TVX '19, pp. 35–46. ACM, New York, NY, USA (2019). https://doi.org/10.1145/3317697.3323361

[61] Bangor, A., Kortum, P., Miller, J.: Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. J. Usability Studies **4**(3), 114–123 (2009)

[62] Grier, R.A.: How High is High? A Meta-Analysis of NASA-TLX Global Workload Scores. Proceedings of the 59th Annual Meeting of the Human Factors and Ergonomics Society **59**(1), 1727–1731 (2016). https://doi.org/10.1177/1541931215591373