Designing Gaze-Aware Attention Feedback for Learning in Mixed Reality

Shi Liu
Karlsruhe Institute of Technology
Karlsruhe, Germany
shi.liu@kit.edu

Peyman Toreini
Karlsruhe Institute of Technology
Karlsruhe, Germany
peyman.toreini@kit.edu

Alexander Maedche
Karlsruhe Institute of Technology
Karlsruhe, Germany
alexander.maedche@kit.edu

ABSTRACT

Mixed Reality (MR) has demonstrated its potential in the application field of education. In particular, in contrast to traditional learning, students using MR get the possibility of learning and exploring the content in a self-directed way. Meanwhile, research in learning technology has revealed the significance of supporting learning activities with feedback. Since such feedback is often missing in MR-based learning environments, we propose a solution of using eye-tracking in MR to provide gaze-aware attention feedback to students and evaluate it with potential users in a preliminary user study.

CCS CONCEPTS

- Human-centered computing → Mixed / augmented reality.

KEYWORDS

mixed reality, eye-tracking, feedback for learning

ACM Reference Format:

1 INTRODUCTION

Mixed Reality (MR) has shown its great potential in the application field of education [15, 22]. Specifically, MR empowers learners to perceive information and discover content in a self-directed way [25]. However, these systems often lack an effective approach to providing feedback and guidance throughout the learning process [28, 35]. Compared to other learning formats such as mobile learning applications, such feedback can be more beneficial for learners due to the spatial distribution of the learning material in MR. Our work aims to deal with this conundrum and proposes a self-directed learning application in MR with gaze-aware attention feedback.

Current research in educational technology and human-computer interaction (HCI) suggests that learning in MR environments can improve the motivation and learning outcomes of students in certain contexts [18, 20]. Examples of such mixed reality integrated learning environments (MILE) can be found in medical training and science education [3, 4]. In these learning contexts, MR contributes to visualizing abstract or intangible content with 3D models. While it can help students gain a greater understanding of the learning material, students don’t always report more confidence after using an MR application, as compared to other learning formats [18, 39]. Several reasons have been identified, including usability issues [29], poor content design [8], and cognitive overload [18, 21]. Therefore, MR may not be superior to traditional learning formats. Instead, some research suggested that it should be implemented as a complement to traditional learning strategies [2, 6].

While most MR applications in the current research either intentionally or unintentionally embrace constructivist learning strategies, such as discovery learning or inquiry learning [13], feedback throughout the learning journey is often missing. Researchers such as Ibáñez et al. [16] have addressed this problem. However, they only treated such a complication from a strategical perspective, by integrating other learning strategies (e.g., scaffolding learning) in MR. Furthermore, these works primarily explored how to assist students by delivering proper instructions. Our work instead attempts to support students by providing feedback based on their visual attention.

By examining the eye-tracking features on Microsoft HoloLens 2, along with finding inspirations from related work in eye-tracking research, our work presents an MR learning application, where students are supported with a gaze-aware attention feedback system. The following sections report the background of existing research, as well as the design and implementation of the system. Lastly, we present early feedback on our work collected in a preliminary user study and insights for future work.

2 RELATED WORK

Using MR to support learning activities has been explored in diverse settings. In medical training, MR applications can not only “make large amounts of information more navigable” [4], but they are also capable of visualizing invisible anatomy structures to support the understanding of tacit knowledge [40]. Other contexts with successful MR learning applications include science education [3], human-robot-interaction [14, 30], language learning [33, 38]. While the benefits of using MR in education have been observed, challenges and limitations also exist.

Students are not guaranteed to learn effectively with MR. First, the learning content in MR may not be effectively designed. Domain experts or instructors often lack the skills to develop an MR application, while MR researchers don’t necessarily have a sufficient understanding of the domain knowledge [1]. As a consequence, effective MR learning experiences are not always replicable. Secondly,
most research in MR education has not incorporated feedback in their systems. To some extent, the commonly used constructivist learning strategies in MR magnify this gap [13, 16]. This potential research gap suggests that not all current challenges in MR research are due to the technical limitations of MR. Instead, researchers should explore learning strategies in MR from a human-centered perspective [7].

Even though researchers have adopted different constructivist strategies (e.g., experiential learning [32], collaborative learning [30], inquiry-based learning [43]), the reason for integrating constructivism in MR is similar: disruptive technologies such as Mixed reality allow students to control “the pace of their own learning through hands-on experiences” [25]. Hence, interactive and explorative learning environments in MR accord with constructivist theories [27]. From a learner-centric perspective, applying self-directed and constructivist learning strategies in MR encourages students to learn actively, resulting in “better learning outcomes evidenced by higher retention of learning” [24].

Whilst students can benefit from self-directed learning environments, feedback and guidance in such environments should be considered. Self-directed learning environments without guidance and feedback usually lead to poor outcomes [19, 23] and can be a “pitfall” for e-learning platforms [9]. Reflective feedback based on learners’ attention can support them in comprehending excessive information presented in a short period of time. Besides, providing reflective feedback in MR is especially meaningful due to the increased cognitive overload [18]. While reducing cognitive load using eye-tracking has been explored in MR [41], little research has implemented eye-based support in learning contexts or proposed guidelines for designing reflective feedback to students. A very few examples of related research include the works of Ibáñez et al. [16] and Oh et al. [28], where they embraced the scaffolded learning strategy, aiming to achieve a balance between creating explorative learning experiences and promoting a good understanding of the learning content [42]. In addition, Thoravi Kumaravel et al. [35] highlighted the importance of having feedback in MR learning and presented an MR tutoring environment where instructors remotely give students feedback. Nonetheless, receiving remote feedback from instructors is not always feasible. Therefore, providing feedback for learning in MR, especially in self-directed learning scenarios, remains a research gap. Meanwhile, research in eye-tracking can provide valuable insights regarding this gap.

Using desktop eye-trackers, researchers have successfully developed gaze-aware attentive systems to provide visual attention feedback [36]. In educational contexts, desktop eye-trackers have also been employed to assist programming [26, 37]. Since HoloLens is also equipped with eye-tracking sensors, providing visual attention feedback to students is feasible. Furthermore, exploring content in MR can require more attention from students, because the spatial arrangement of material adds more complexity to the learning environment. Therefore, integrating attention feedback in MR can potentially enhance the learning experience and improve learning performance. Currently, incorporating eye-tracking for this purpose in MR has not been well-explored [8]. Thus, our work aims to investigate this research gap and explore the potential of providing visual attention feedback in MR learning environments.

To summarize, the review of existing research reveals a research gap in providing gaze-aware attention feedback for MR-based self-directed learning. In an effort to fill this gap, the following research question is articulated: \textbf{RQ} \textit{How to design a gaze-aware MR system to provide effective attention feedback in self-directed learning environments?}

\section{METHOD}

In this section, we describe the method for designing the gaze-aware MR system, following a human-centered design approach.

\subsection{Specifying the Context}

Before implementing the system, a specific learning context and a target user group need to be defined. Since we are HCI researchers and have adequate domain knowledge in HCI, we selected an established concept in HCI as the underlying learning material: the Model Human Processor (MHP). The MHP describes human information processing (HIP) in a simplified approach and is considered a pioneering interdisciplinary concept in HCI research [5]. The concept analogizes a human user to a computer, consisting of a perceptual system, a cognitive system, a motor system, and memory storage. Each system has a defined cycle time, which researchers can use as a reference when designing interactive systems. The reason for selecting this concept can also be attributed to the possibility of learning it in a variety of formats, ranging from 3D rendering of CAD models to videos. We believe that learning such an abstract concept with MR requires great attention, and supporting it with gaze-aware feedback can be beneficial for HCI students.

We used Microsoft HoloLens 2 as the MR device in our work. HoloLens is an optical see-through head-mounted display (OST-HMD) with extended features such as speech recognition and spatial awareness. The built-in eye-tracking sensors have a refresh rate of 30Hz, which is adequate for providing gaze-aware attention feedback [34]. Having decided on the learning content and the device, the context of this work is specified as follows: \textit{Exploring the self-directed learning experience of an HCI lecture in HoloLens 2, supported by an integrated gaze-aware attention feedback system.}

\subsection{Iterative Design Process}

The human-centered design process emphasizes understanding users’ needs and requirements. In an effort to achieve this, we reflected on our personal experience of learning HCI concepts during lectures. Besides, preliminary interviews were conducted among seven HCI students. While some students acknowledged that “it (the MHP concept) seems to be a pioneering concept at that point” and “the efforts to bridge cognitive science and computer science” (S1, S3), others suggested that the concept is “too abstract to feel the relevance in real-world use” (S5). The interview results indicate that the intended learning support of attention feedback can be beneficial for such abstract concepts.

Based on the interview results and related work, low-fidelity prototypes were created to visualize the MR application. The system consists of three scenes: a) a brief tutorial of the system; b) self-directed explorative learning; c) presentation of attention feedback. First, the brief interactive tutorial aims to reduce users’ cognitive workload in MR onboarding. In specific, the tutorial introduces...
gesture and eye-based interactions in HoloLens, along with a short description of the learning content. Second, the exploration scene consists of the lecture content with videos, images, texts, and 3D models. In this stage, users are allowed to use gaze-based interactions to explore different parts of several models to trigger corresponding content. Thereby, the system can detect the gaze and visual attention of users and provide individual attention feedback afterward.

The explorative learning scene presents several 3D models and the corresponding learning material. The exploration starts with the central model of a human using a computer. Users can look at different parts of the human body and trigger the corresponding content. The learning content is divided into four sections: visual perception, auditory perception, motor system, and cognition. Each section also comprises one secondary 3D model. For example, a model of eyes is displayed in the visual perception section. Furthermore, each secondary model contains multiple learning items. These items are explanatory content including textbook excerpts, videos, and images. To activate these items, users need to again look at different subparts of the secondary models (e.g., the retina of the eye model) for some time. In summary, the system contains three layers of information: a) the central human-computer model; b) secondary models in four learning sections; c) learning items activated by gazing at secondary models. Figure 1 illustrates this structure and the learning content.

The last scene provides users with gaze-aware attention feedback on their learning experience from the previous scene. Based on eye-tracking data, an overview of all learning items and a visualization of users’ visual attention are presented here. Three colors (blue, yellow, and red) are used to signify three levels of visual attention. Concretely, the content to which users have previously paid great attention is marked in blue, while the content with little attention is marked in yellow. Lastly, completed ignored content is marked in red. Figure 2 is a low-fidelity prototype, showing the design of this scene. Distinguishing different levels of gaze attention and visualizing them is the fundamental design goal. The following section discusses how we implement our design and achieve this goal.

4 IMPLEMENTATION

The implementation started with creating 3D models in Blender [11], a CAD software used commonly for 3D animation. For creating the central model, i.e. a human interacting with a computer, free resources were utilized [12]. The CAD models along with the Mixed Reality Toolkit (MRTK) were imported into Unity [31].

Unity is a game engine that supports development across various platforms, including MR devices such as HoloLens. For HoloLens development, the MRTK packages integrate an “input system and building blocks for spatial interactions and UI” into Unity projects [10]. After setting up the Unity project, three scenes were built based on the design: an introduction scene, an exploration scene, and a feedback scene. In each scene, gaze-based interactions are enabled. The exploration scene contains interactive elements, learning items in videos, images or texts, and rendered CAD models. C# scripts are attached to the learning items and 3D models, collecting eye-tracking data in the exploration scene. In particular, when a user’s gaze focuses on a Unity GameObject, the backend scripts log the name of the GameObject with a timestamp locally on HoloLens.

The attention feedback was implemented based on these scripts. Each learning item has a predefined threshold for enough visual attention. For instance, students are expected to focus their gaze on a learning item with text for 15 seconds. If they do so, the script attached to the corresponding GameObject marks it as “focused” in the log file and keeps the color of its title bar unchanged. Otherwise, the script changes the title bar of the GameObject to a different color (yellow or red), depending on the duration of users’ gaze attention. Such a threshold of 15 seconds and the color-coding design are defined based on proposed suggestions from the preliminary user interviews. After the exploration scene, an overview of all learning items with attentive color-coding is displayed. As previously described, three levels of attention are visualized by three colors of the title bars, creating a rough heatmap of visual attention (See Figure 4). Besides, users are also prompted to see a summary of their learning progress. The summary presents three numbers: the number of learning items a user has completely missed (marked in red), items that a user has only paid little attention to (marked in yellow), and items that a user has explored with great attention (marked in blue).

5 PRELIMINARY USER STUDY

To evaluate the current system and collect early user feedback, a preliminary user study with eight participants (3F, 5M) was conducted. The study focused primarily on collecting qualitative data from interviews and think-aloud protocols. First, participants wore HoloLens to use the application. Meanwhile, they were encouraged to express their feelings and opinions by reporting verbally, i.e. using the think-aloud method. After using the system, semi-structured interviews consisting of three parts were conducted. First, we asked the participants about their overall impressions of the system. Second, we encouraged them to reflect on their learning experience and how they had perceived the attention feedback. Third, we made a free discussion on improving the system for further research.

Results from think-aloud protocols and interviews suggest that users enjoyed the learning experience and the attention feedback. Most participants gave MR the credit for supporting learning and suggested that “it (MR) creates an interactive learning experience” (P3). Participants also explained the reasons for a positive overall impression in detail.

First, the visualization of the learning content was fairly effective, as five out of eight participants reported that the 3D rendering was helpful and innovative, especially “for visual learners” (P2). Second, the variety of learning formats may have helped users focus on the content and made the learning experience more engaging. Seven out of eight participants explored more than ten learning items in the exploration scene. Particularly, one participant suggested that different formats “make the learning experience more fun than a lecture” (P4).

Furthermore, the personalized attention color-coding feedback was well-received. Interviews reveal that a majority of participants (six out of eight) perceived the individual feedback as “helpful” or
“meaningful”. However, participants also reported cognitive overload when receiving the feedback. Two participants (P6, P7) reported that the transition between self-directed exploration and receiving attention feedback is “somewhat abrupt”.

6 DISCUSSION

Based on the results of the preliminary user study, we assume that the positive perceptions somewhat confirm that providing gaze-aware attention feedback based on eye-tracking data can potentially benefit self-regulated learners.

However, the work presented in this paper comes with several limitations and implications for further research. Specifically, there is a need to improve the current attention feedback system in the future. Achieving a balance between providing adequate feedback for learning and keeping the cognitive load of users at a reasonable level requires further research. Furthermore, the current attention feedback only provides limited information. Further research can explore generating an attention heatmap based on the gaze position coordinates and study whether such a more detailed feedback system contributes to better supporting students.

Several design decisions for the system were made based on insights collected from the preliminary user interviews, including the visualization of the attention feedback with three indicative colors. Nonetheless, minor usability challenges were identified during the user study. For further research, integrating participatory design methods with more user involvement can potentially improve the quality of design decisions.

Moreover, the human-centered design process for MR can benefit greatly from additional methods. For instance, the body storming technique mentioned in the work of Kerawala et al. [17] represents an interesting approach for MR prototyping. The technique requires designers to move around in a room and use low-fidelity prototypes to investigate the spatial relationship between the physical environment and the virtual elements. The method used in the preliminary user study is also reflected: quantitative analysis can be integrated for a more in-depth evaluation with a larger sample size. Further research into adapting the system to other learning contexts and studies on the universality of attentive feedback for learning in MR are also needed.

7 CONCLUSION

In this work, we address a research gap in supporting self-directed learning with attention feedback identified by reviewing existing MR research in the field of learning. As a response, we present a gaze-aware learning support system using an MR head-mounted display. The system is capable of detecting users’ visual attention based on eye-tracking data and providing individual feedback on an HCI lecture. Early user study results suggest overall user satisfaction and denote that the system can support students in self-directed learning scenarios. We acknowledged that the system requires further development and a comprehensive evaluation. In conclusion, our work provides a foundation for delivering individual feedback based on eye-tracking data in MR education, and we believe that further investigations can be made to explore its potential in diverse educational contexts.

REFERENCES


Designing Gaze-Aware Attention Feedback for Learning in Mixed Reality


[12] Andy Goralszuk. 206. Vincent Character Rig. https://studio.blender.org/characters/57186e7c7c97f0942a977f42a471v1


[22] Zihan Yan, Yufei Wu, Yiying Li, Yifei Shan, Xiangdong Li, and Preben Hansen. 2022. Design Eye-Tracking: Augmented Reality Headset to Reduce Cognitive
