ModularHMD: A Reconfigurable Mobile Head-Mounted Display Enabling Ad-hoc Peripheral Interactions with the Real World

Isamu Endo
Tohoku University
Sendai, Japan
isamu@riec.tohoku.ac.jp

Kazuki Takashima
Tohoku University
Sendai, Japan
takashima@riec.tohoku.ac.jp

Maakito Inoue
Tohoku University
Sendai, Japan
maakito.inoue.p5@dc.tohoku.ac.jp

Kazuyuki Fujita
Tohoku University
Sendai, Japan
k-fujita@riec.tohoku.ac.jp

Kiyoshi Kiyokawa
Nara Institute of Science and
Technology
Nara, Japan
kiyo@is.naist.jp

Yoshifumi Kitamura
Tohoku University
Sendai, Japan
kitamura@riec.tohoku.ac.jp

Figure 1: ModularHMD's main use cases: (a) user is immersed in VR content using ModularHMD with three modules displaying extended VR views; (b) user can remove bottom module and employ it as an input device to interact with VR content; (c) user can remove side module for peripheral awareness and lend removed module to a non-HMD user as an instant interaction device, allowing ad-hoc cross-device collaboration of VR content.

ABSTRACT

We propose ModularHMD, a new mobile head-mounted display concept, which adopts a modular mechanism and allows a user to perform ad-hoc peripheral interaction with real-world devices or people during VR experiences. ModularHMD is comprised of a central HMD and three removable module devices installed in the periphery of the HMD cowl. Each module has four main states: occluding, extended VR view, video see-through (VST), and removed/reused. Among different combinations of module states, a user can quickly setup the necessary HMD forms, functions, and real-world visions for ad-hoc peripheral interactions without removing the headset. For instance, an HMD user can see her surroundings by switching a module into the VST mode. She can also physically remove a module to obtain direct peripheral visions of the real world. The removed module can be reused as an instant interaction device (e.g., touch keyboards) for subsequent peripheral interactions. Users can end the peripheral interaction and revert to a full VR experience by re-mounting the module. We design ModularHMD's configuration and peripheral interactions with real-world objects and people. We

also implement a proof-of-concept prototype of ModularHMD to validate its interactions capabilities through a user study. Results show that ModularHMD is an effective solution that enables both immersive VR and ad-hoc peripheral interactions.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; *Mixed / augmented reality*.

KEYWORDS

Asymmetric interaction, situational awareness, shape-changing interface

ACM Reference Format:

Isamu Endo, Kazuki Takashima, Maakito Inoue, Kazuyuki Fujita, Kiyoshi Kiyokawa, and Yoshifumi Kitamura. 2021. ModularHMD: A Reconfigurable Mobile Head-Mounted Display Enabling Ad-hoc Peripheral Interactions with the Real World. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21)*, October 10–14, 2021, Virtual Event, USA. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3472749.3474738>

This work is licensed under a Creative Commons Attribution International 4.0 License.

UIST '21, October 10–14, 2021, Virtual Event, USA
© 2021 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-8635-7/21/10.
<https://doi.org/10.1145/3472749.3474738>

1 INTRODUCTION

Head-mounted displays (HMDs) have become more affordable, portable, and no longer require external equipment (e.g., [12]). Such mobile HMD setups offer huge potential beyond entertainment uses [80], including enabling mobile working environments [19] in which users can enter a virtual workspace while they are

physically in an office [25, 63], in shared spaces [36, 84], or on transportation like trains and airplanes [4, 53], etc. In such scenarios, HMD users require a certain level of situational awareness of the real world to respond to sudden physical interactions or nearby actions from non-HMD users. However, most current HMDs use occlusive cowls to induce an immersive VR experience that completely isolates the HMD user from the real world [36]. We call such isolation the "HMD boundary," which causes the following two issues. First, HMD users cannot view and control around-hand real objects (e.g., cups, pens, keyboards, etc.) without removing the headset [6, 7, 20, 21, 37, 40, 48, 74]. Second, HMD users have limited situational awareness and interactivity with nearby people [46, 54] while wearing the headset.

Individual solutions have been proposed for each issue. For the first issue of hindered ad-hoc object interaction, the most common approach uses a video see-through (VST) display that visualizes surrounding objects in the VR view (e.g., [7, 33, 46]). Another approach is to redesign the HMD. Closely related to our work, in Multi-channel Dynamic Immersion (MDI) [74] transparency-controllable LCD panels are installed in the side cowls of an HMD to dynamically provide peripheral vision of the real world to HMD users. The second issue of limited ad-hoc asymmetric collaboration is more complicated because bidirectional situational awareness between HMD and non-HMD users should be ensured (e.g., [46, 78]). A viable solution is FaceDisplay [24], which adopts a unique HMD form. It is comprised of three outward-facing touchscreens that visualize the HMD user's face, allowing co-located collaboration between HMD and non-HMD users through touch interactions around the HMD user's head. Nonetheless, even though the two issues have been addressed separately, no coherent solution has addressed them altogether.

In this paper, we propose ModularHMD, a reconfigurable mobile HMD concept that enables HMD users to perform two types of ad-hoc peripheral interactions with real-world devices and non-HMD users while maintaining an immersive VR experience. ModularHMD's basic idea is using detachable modular devices in the HMD's peripheral cowls. Each module has four main states: occluding, extended VR view, video see-through (VST), and removed/reused. Among different combinations of module states, a user can quickly setup the necessary HMD forms, functions, and real-world visions for ad-hoc peripheral interactions without removing the headset. Although the concept of modular devices offers huge customizing opportunities, we focus on the above two main issues arising from the HMD boundary. Fig. 1 illustrates our proof-of-concept prototype of ModularHMD, which is comprised of a conventional HMD with three removable module devices for its side and lower cowls. In the initial configuration of ModularHMD, all the modules are attached and their inward-facing screens can display an extended VR view of the central VR content (Fig. 1(a)). Figs. 1(b) and (c) illustrate the two types of peripheral interactions with peripheral devices (e.g., a hand-held keyboard) and a real-world non-HMD partner. Both interactions, which are initiated by removing a module from the HMD, can be conducted by reusing it as an interaction device and terminated by returning it to the HMD. Fig. 1(b) shows an example of text typing during a VR experience where the HMD user is reusing the removed bottom module as a touch keyboard that is visible through the opened-bottom slot.

Fig. 1(c) is an example of an ad-hoc cross-device collaboration with a real-world partner where the HMD user gives the removed-side module to a partner as an interaction device. They exchange facial expressions and eye contact through the opened-side slot. The cyclic action of removing, reusing, and remounting a module explicitly manages the workflow of ad-hoc peripheral interactions as well as makes the HMD setup fully mobile and self-contained (without requiring extra devices).

ModularHMD is inspired by a series of previously suggested HMD form concepts (e.g., [8, 24, 42, 45, 55, 58, 74]), but to the best of our knowledge, introducing a modular mechanism into a mobile HMD is completely novel. In this paper, we conduct an exploratory design of the modular mechanism of ModularHMD through an iterative process, design four peripheral interactions to address the two HMD boundary issues, implement a proof-of concept prototype, and progressively confirm how our prototype achieve the designed interactions through four user study sessions. The following are the three contributions of this work: 1) we propose ModularHMD, a new HMD design concept with a reconfigurable structure that involves removable and reusable module devices at the periphery of the HMD cowl; 2) we discuss its detailed design process and peripheral interactions with real-world devices and partners; and 3) we implement its proof-of-concept prototype and conduct a set of user studies demonstrating its four interactions that mitigate the HMD boundary issues.

This paper reports significant updates to our previous work [11] with new design, prototyping, and user studies.

2 RELATED WORK

2.1 HMD boundary issues

Mobile HMDs are used in shared spaces [36, 84], homes [78], offices [25, 63], and in transportation like trains and airplanes [4, 53]. For such location-independent VR offices, Grubert et al. [19] argued that high mobility and immersiveness are the core elements of HMDs, but they also indicate the critical issues of the existing HMDs that arise from their occlusive cowls: lower situational awareness of the real world, poor physical interaction, hindered communication among users, and poor text-typing performance.

2.2 HMD user situational awareness for physical interaction

2.2.1 Tracking HMD-user status. The simplest solution for gaining real-world situational awareness is to remove the HMD headset (e.g., flip visor [1]). However, HMD users lose access to the VR world, which does not allow real-time interaction with the ongoing VR content using peripherals (e.g., physical keyboard). An optical see-through HMD (e.g., [49]) and a non-cowl HMD (e.g., [30, 77]) support peripheral awareness and interactions during VR experiences but they always sacrifice immersive experiences.

2.2.2 Visual cues of real-world information. A common solution is overlaying captured real-world objects and people on VE, which includes the Augmented Virtuality (AV) approach [7, 28, 31, 33, 46, 55, 78], displaying real-world people or objects as previously prepared 3D models [6, 16, 20, 21, 37, 40, 47, 48, 65, 68, 69], and their combinations [78]. Although these visual cues significantly

increase the situational awareness of HMD users, they do not help bystanders estimate the status of HMD users or facilitate collaborative interactions.

2.2.3 Transparency-controllable HMD. An approach closely related to ours is a transparency-controllable HMD that allows HMD users to see the real world. CompoundDome [45] is a dome-shaped transparency-controllable display that fully covers a user's head and allows her to look at real-world objects and perform face-to-face communication with real-world users while watching a projected image on its inner surface. Its design goals are close to ours, although it does not support stereoscopic vision and collaborative interaction. MDI [74] is a dynamic immersion-control HMD that uses transparency-controllable LCD panels in the HMD cowl, enabling HMD users to see and access peripheral devices (e.g., smartphones, keyboards, a mouse) without removing the headset by making LCD panels transparent. Our ModularHMD is built upon this idea but significantly expands it in three ways: enabling direct vision of the real world, self-contained peripheral interaction, and cross-device collaboration between HMD and non-HMD users.

2.3 Interaction between HMD and non-HMD users

Due to the HMD boundary, non-HMD users cannot easily estimate such states of the wearer as facial expressions, gaze directions [13], and the presented VR content. To address this issue, many efforts have focused on how to present the HMD-user status and how to share VR content with others.

2.3.1 Tracking HMD-user status. To capture the HMD user's facial expressions, tracking systems installed inside the HMD have been proposed, such as eye-trackers [72], IR cameras [10, 61], and photo-reflective sensors [51]. These proposals all focused on facial tracking for face-to-face or avatar-based communication in VR, but not for real-world interaction between HMD and co-located non-HMD users.

2.3.2 Sharing VR content using external devices. To share the VR content of an HMD user with others, recent commercial HMDs have a function that mirrors the content onto other screens (e.g., [1, 12]). Many previous studies also used external devices, such as desktop displays [81], touch displays [17, 26, 31, 73], projections [14, 23, 28, 79], and a CAVE-based display [32] to share VR content with nearby non-HMD users, where non-HMD users may be invited to interact with it. However, they require additional setup in advance and might not be suited for ad-hoc interaction [38] in mobile VR scenarios.

2.3.3 HMDs for collaboration with non-HMD users. To enable asymmetric interaction between HMD and non-HMD users, new HMD compositions have been proposed [8, 24, 42, 57]. HMD Light [79] is a unique mobile HMD with a projector on its top which projects VR content onto the physical floor surface of the environment, where non-HMD users can interact with the projected VR content using their feet (or hands). But, it does not offer the facial expressions of HMD users and is not fully mobile and self-contained because strong physical setup requirements exist

(e.g., a flat, clear surface for projection and foot-based interaction). FaceDisplay [24] is a full mobile HMD with three outward-facing touchscreens and a depth camera at its back, allowing ad-hoc asymmetric collaboration between HMD and non-HMD users through touch or gesture interactions. The idea of a multi-display around HMD is closely related to ours, but its primary motivation facilitated entertainment applications with unique physical formations around the HMD user's head, and did not support interactions with peripherals. TransceiVR [73] is a platform that allows cross-device, asymmetric VR collaboration where an external user can see the VR mirroring view and annotate with a connected tablet. This work offers a practical cross-device interaction platform, although it does not aim to create ad-hoc interaction in a mobile context. Therefore, we expand the previously proposed efforts by providing ad-hoc distributed cross-device workspaces for both users through ModularHMD's quick modular interactions, where natural formations (e.g., side-by-side) and full mobility can be ensured.

2.4 Modular device interactions

To physically mitigate the HMD barrier, we pay particular attention to shape-changing devices that enable dynamic forms of interfaces. Among various types of shape-changing approaches [59], we focus on the modular device concept, which consists of multiple equal-sized parts that can be collected or separated. Although this idea has been applied to some mobile devices (e.g., [18, 44, 67]), examples are scarce for HMD. For instance, Rakkolainen [58] used two smartphone-based modules to extend an HMD's FOV; yet no modular action was explored. Modular [55] allows different types of optical vision augmentations by physically changing the camera-lens systems on the back of an HMD. The Nintendo Switch Labo VR kit [52] offers modular interaction where the modular controller is initially mounted on the HMD. It can also be removed to be reused as an interaction device. Such implementations also change the shape of an HMD with various cowls and attachments to fit game scenarios. These systems provided strong inspiration and demonstrated the potential of modular devices to new HMD designs, even though their primary purpose was not to mitigate HMD boundary.

Another unique feature of the modular device concept is reusing free modules for different purposes to increase functionalities and customizability. "Device lending" is an important example studied with modular smartphones [67]. Here the owner of the main modular smartphone can provide interaction authority to a trusted person (e.g., a family member) by giving them one of the modules. When the instant interaction is over, the module can be returned to its owner. This is more convenient for managing the authentication process between owners and users than ad-hoc uses with their own devices. This work also served as a primary inspiration for quickly facilitating ad-hoc collaboration across the HMD boundary.

In summary, our ModularHMD primarily builds on prior MDI [74], FaceDisplay [24], and Modular smartphones [67]. We extend MDI by employing physical and functional modularities that allow full visual access and control for real-world devices and partners. We were also inspired by FaceDisplay and achieved different ad-hoc asymmetric interaction using cross-device distributed workspaces that are rapidly created by a module-lending mechanism.

Table 1: Issues and solutions

Requirements	Type	Necessary interactions
Immersion to VR	-	<i>In₁</i> Immersion into VR: HMD users can become immersed in VR content.
Interaction with physical objects	-	<i>In₂</i> Ad-hoc around-hand object interaction: HMD users can interact with the peripherals around their hands.
Interaction with non-HMD partners	Conversation	<i>In₃</i> Ad-hoc face-to-face conversation: HMD and non-HMD users can have conversations while seeing each other's facial expressions.
	Collaboration	<i>In₄</i> Ad-hoc cross-device collaboration: Non-HMD users can join the VR content interaction of HMD users.

3 MODULAR HMD

3.1 Requirements and goals

As discussed, our goal is to design a set of interactions to broadly address the two HMD boundary issues: limited awareness to real-world objects and partners, while maintaining an immersive VR experience. Table 1 summarizes the detailed requirements and necessary interactions we plan to design and implement. Three additional interactions, *In₂* – *In₄*, are chosen based on our targeted mobile work contexts. We are specifically interested in supporting active interactions where HMD and/or non-HMD users can join the VR content manipulations. This will significantly motivate physical bidirectional interactions among users and devices. When users only perform passive interactions with unidirectional content transmission, existing solutions (e.g., VST or screen mirroring, etc.) might be sufficient.

We achieved these balanced interactions by establishing the following assumptions derived from ad-hoc mobile work contexts:

- Complete mobile HMD setup: every function must be self-contained without any external devices.
- Immersive VR experience *In₁* is the priority, and *In₂* – *In₄* can be executed only when ad-hoc peripheral interactions occur.
- Peripheral interactions are basically short and related to the ongoing VR content.

Needless to say, *In₁* is the most fundamental interaction of a VR HMD. *In₂* and *In₃*/*In₄* correspond to the two types of peripheral interactions with real-world objects and nearby persons. For *In₂*, we focused on a subtle and active object interaction, where the conventional VST mode may not work correctly. *In₃* and *In₄* are related to asymmetric collaboration, although they are separately considered in this work. *In₃* supports social awareness, which is based on the prior suggestion that non-HMD users might want to know the statuses of the HMD users during the interactions. We focused on eye contact between users, which is critical non-verbal information to enrich the social interaction. *In₄* is a subsequent complex interaction built upon the created social relations. Here we focused on cross-device collaboration because it offers much more flexible workspaces than existing single shared screen approaches (e.g., HMD Light[79]) and easily supports both passive (e.g., viewing) and active (e.g., manipulating) interactions by non-HMD users.

3.2 Approach

Having identified that the underlying problem for both the HMD boundary issues is HMD's physical structural limitation, we tackle this core issue by introducing the physical flexibility of the HMD cowls. We expect that physically removing/remounting the peripheral part of the HMD supports *In₂* and *In₃* and reusing the modules by the HMD users themselves or someone else assists *In₂* as a private device and *In₄* for cross-device interactions.

3.3 Design considerations and parameters

We describe our design considerations for creating a ModularHMD prototype in terms of HMD physical form, FOV, user's visual field, and modular interactions.

3.3.1 HMD peripheral cowls. We first revisit HMD's physical forms. The conventional HMD's physical components are a central screen, two lenses, peripheral cowls, head belts, etc. We focus on cowls that have been designed to comfortably fit a user's face and create a dark-room inside the HMD. In addition to the shapes, our work includes two more critical parameters of HMD cowls: the areas (right, left, bottom, and up from the user's point of view) and transparency (between opaque to transparent). Unfortunately, most existing HMDs have full cowls with opaque materials to induce an immersive experience, and these variables make different functionalities specifically for peripheral interactions. Since interaction above the head is rare, we focus on how to control the transparency of the right, left, and bottom cowls. Potential approaches might use a video-see-through monitor and transparent materials, but we additionally introduce physically removing or remounting the cowl because it can be switched between completely opaque and transparent in a bidirectional way for both users. While the shape-changing interface is getting more attention, such a physical shape-changing concept (like a door opening) has not been studied in VR HMD design. This hypothesis that re-purposes the cowl becomes a starting point for ModularHMD's modular mechanism.

3.3.2 Module device functions. Our biggest design parameter for modular device interaction is a function for each module. To satisfy our goals (*In₁* - *In₄*), we considered four functions for each module: an occluding panel, an extended VR view, a video see-through (VST), and removed/reused. The first function, an occluding panel, is the

Figure 2: Arrangement of central VR and two peripheral side modules relative to user's field of view and eye movement characteristics: (up) naive design: central VR screen supports a wide FOV (blue) that unfortunately creates sides modules around far peripheral areas: (bottom) solution: central VR FOV is reduced and centers of peripheral module's screens are arranged and packed within immediate viewing area.

most fundamental one to create darkness inside the HMD. The extended VR view is an optional function to increase the HMD's FOV. VST quickly and unidirectionally controls the transparency of the HMD barrier from the HMD-user's perspective, while Removed provides a full transparency for both HMD and non-HMD users. Depending on the required situational awareness, each is differently usable. Reused benefits from the modular concept allows quick preparation of a secondary interaction device, which definitely accelerates the ad-hoc cross-device interaction. Other functions (e.g., haptics [9], tangible interface [75]) can be considered for each module to enrich the VR experience, however such efforts are beyond the current scope of our design within the HMD boundary.

Each module's screen face direction can also be considered as a parameter that determines the HMD appearance and functionalities. If the module's screen is set to outer-facing, the HMD might be used to visualize the HMD user's status or the VR content, similar to FaceDisplay [24] or FrontFace [8]. But to simplify the prototyping and primarily support HMD user interactions, we only use inner-facing module screens without a reversing action in the initial stage of our ModularHMD design process.

3.3.3 Module device arrangement adapting human visual field. Our next design challenge is arranging the modular devices into an HMD. Here, even though the human's visual field is the main point, it cannot be an independent design variable. Thus, this section focuses on our iterative design process for adapting multiple modules into comfortable viewing characteristics for humans. Fig. 2 shows the relationship between HMD screens and the human visual field.

A person's full visual field (without head motion) is approximately 190 and 125 degrees horizontally and vertically (50 degrees up, 75 degrees down) [29], shown as black lines. However, the effective visual field is much narrower and changes depending on the tasks; it becomes 60 degrees (horizontal) and 55 degrees (vertical) for correct symbol recognition and 3D depth perceptions [71]. Considering the eye's maximum movement range (reported to be 50 degrees [64]), one can immediately view 100-degree horizontal and vertical viewing areas with both eyes, depicted as the green dotted lines. The areas outside this immediate view are defined as peripheral vision, which cannot identify the content details yet capture the motions in an environment. To reflect such a human visual system, current commercial HMDs offer a wide FOV up to 210 degrees [70]. The average FOV is around 90 degrees, which is sufficiently substantial to induce an immersive experience.

The design challenge in module arrangement is how to achieve both an immersive VR experience and immediate peripheral interactions. The central VR screen's FOV must be kept as much as possible and the peripheral modules must be arranged within the user's immediate visual field. Figs. 2 (a1)-(a2) show a naive design example that uses a commercial HMD with a high FOV of 90 degrees (blue in Figs. 2 (a1) and (a2)). With this approach, a substantial VR experience can be made, but the peripheral modules are arranged around far peripheral areas. The orange and purple areas in these figures indicate the FOV of the side and bottom modules, respectively. Unfortunately, most exceed the green areas (immediate view angles supported by eye movements), which means that these peripheral modules cannot be viewed in the user's center of sight. This makes it difficult for the HMD user to focus on the images

Figure 3: Structural summary of ModularHMD by different states of three modules.

of the peripheral modules or the real world when the modules are removed.

Our solution is reducing the central FOV to around 60 degrees, which is narrower than the current HMD trend, yet still within a range that can provide an immersive experience [39] (blue area in Figs. 2 (b1) and (b2).) Such HMD's FOVs can be adjusted by modifying the lens arrangement for all the screens. With this approach, the center of the peripheral modules can be successfully arranged within the immediate viewing range (green lines). Even though this module arrangement is a compromise, it remains necessary to achieve a new mobile HMD that enables both VR and peripheral interactions. To compensate for the decreased central FOV, we provide additional extended VR views using the side module screens.

3.3.4 Module device action. Next we design the modular device. One requirement is easy module state manipulations for HMD users who cannot directly see their hands. Another requirement is compatibility with subsequent device “reusing” interactions. After a brief survey of modular devices (e.g., [52, 67]), we chose a sliding slot for module removal/remounting because its physical rail guides the user for simple module alignment, and its structure can be easily 3D-printed. Other options include snapping with a magnet, although such a remounting action might be inaccurate. We also include an option to toggle the VST mode by double-tapping the back side of the module, inspired by a commercial HMD's gestural command.

We may consider a set of guidelines to help the HMD user initiate a correct module interaction. For example, in case an approaching person does not intend to talk with the HMD user, removing a side module is too much effort for the HMD user. A current use suggestion is activating VST first after hearing footsteps and then observing the person's intentions to perform the appropriate subsequent actions. We do not implement a system or software to guide the HMD user for now. Nevertheless, an additional situational tracking system to detect the spatial elements of real-world surroundings (e.g., approaching persons) will be highly helpful and this information will be incorporated into the modular interactions in the future.

3.4 ModularHMD interactions

Figure 3 depicts the whole states of the ModularHMD's modules. The central, gray rectangular area in Fig. 3 (a) indicates the basic HMD view that enables a primary high-quality VR experience. Each peripheral module is made with a smart device whose states can be switched among occluding panels, extended VR view (blue in Figs. 3 (b)), VST view (purple (c)), and removed/reused (orange (d1) and green (d2)). If all the modules are occluding panels, that design is identical to a regular HMD. ModularHMD basically supports the four interactions in Table 1, and the following illustrate some derivative interactions for more realistic contexts (Fig. 4).

(a) **In_1 Immersive VR (all modules: VR):** When all the modules are attached and display the extended views of the VR content, the user has an immersive VR experience augmented by the peripheral VR views [22, 35, 58, 82, 83] provided by the screens of the inner-facing modules. Fig. 4 (a) shows an example of this interaction where the user is watching a 360 degree walkthrough VR video.

(b) **In_2 ad-hoc around-hand object interaction (bottom: VST):** When a module is put into the VST mode, the user has desktop awareness to grab peripheral objects on the desk, e.g., a cup, a pen, and notes, while remaining focused on the central VR content (Fig. 4 (b)). Due to the current VST limitations (e.g., delay, depth field mismatches, etc.[62]), it might not be suitable for such subtle delicate interactions as text typing. But we believe it will be more powerful in near-future high-quality VST systems (e.g., Varjo XR-3 [76]).

(c) **In_2 ad-hoc device interaction (bottom: removed)** is a more mobile-oriented interaction. Instead of VST, if the bottom module is removed, the HMD user can directly see the desktop devices as well as her own hands through the opening, allowing subtle around-hand device manipulations. In addition, the removed bottom module immediately becomes available as a hand-held input device (e.g., a keyboard, touchscreen etc.) that can be manipulated with her visible fingers. She can place the module back in the HMD to return to the full VR experience (Fig. 4(c)). This interaction will be useful in full mobile work scenarios (e.g., train or air-plain sheet) specifically where desktop devices, voice inputs or VR-gesture controllers are unavailable or unacceptable.

Figure 4: Interaction patterns of ModularHMD

(d) *In₃* **ad-hoc peripheral situational awareness (side: VST or removed)**: This interaction increases the HMD user's peripheral situational awareness and supports the first stage for face-to-face conversations with real-world partner. The HMD user can use the side-module VST mode to recognize a person who is approaching or sitting nearby while she continues to focus on the central VR view (Fig. 4(d)). We focused on side-by-side formation to adapt our module arrangement. Head rotation might be needed to capture persons in other formations (e.g., front or back). This VST mode only provides unidirectional awareness to the HMD user. If bidirectional facial awareness between HMD and non-HMD users is required in the second stage after identifying a nearby non-HMD user, removing a side module can be a valid choice. The empty space by removing the side module allows both users to understand their facial expressions (e.g., attention or gaze direction) for conversations. Fig. 4(e) shows a ModularHMD whose side opening captures a nearby person.

(e, and f) *In₃* and *In₄* **ad-hoc side-by-side cross-device collaboration (side: removed/reused)**: To achieve cross-device collaboration, HMD users can establish a face-to-face communication channel with a real-world partner (Fig. 4(e)) by removing the side module. They can also immediately start asymmetric cross-device interaction by lending the removed module to a partner as an interaction device. Since the module is still working as part of the ModularHMD system, bidirectional interactions can be ensured between modules by partner and VR view by HMD users. If all the modules are lent, a one HMD user and a three-partner collaboration is enabled. Yet for simplicity, we focus on the simplest HMD user: one non-HMD user collaboration.

Here, we discuss an example scenario of a collaboration between an HMD user with a VR modeling application and a non-HMD user joining the VR interaction with a hand-held device. Fig. 4(e) shows the HMD user view and (f) shows the view of the non-HMD user's hand-held device. As shown in Fig. 4(e) the HMD user can see the partner's face as well as the central VR view that visualizes the

partner's position, orientation, and touch interaction information to share the partner's current focus. Fig. 4(f) is a hand-held device, where the partner can change the view port in the VR world using visual joysticks (Move and Rotate), view the position and orientation of the HMD user, and make touch selections. Compared to an ad-hoc connection using their own smartphones, this design is more convenient because the module device was already connected to the VR system and the collaboration is initiated or ended explicitly by the lending or returning action.

3.5 Prototyping

3.5.1 Hardware components. We built a proof-of-concept prototype comprised of a VRM-100 monitor unit (Japan Display Inc. H: 80 × W: 159 mm, Resolution: 2880 × 1600, FOV: 96 degrees with accompanying lens) as a central base HMD and three display modules, their holders, lenses, and attachments. All the components were fixed on an Oculus Rift S head belt. As a module device, we used a 7th generation iPod Touch because it has adequate functions, including a high-resolution display, a backside camera, long battery life, and wireless data communication. See Fig. 16 in the appendix for the detailed layout of each part. We used two, 41-mm diameter Oculus Rift DK2 lenses for the central display and four 35-mm diameter Oculus Rift DK1 lenses for the peripheral modules. Although these lenses are not official VRM-100 accessories (and not designed for smartphone VR), we confirmed that they sufficiently worked for our prototype. The optical design details will be further updated, but the current design adequately worked to prove our concept.

3.5.2 ModularHMD assembly. Figure 5(a) shows all the components of the side module. Except for the iPod Touch and the DK1 lens, all the attachments were 3D-printed. The module holder has two parallel sliding rails, and the iPod Touch and the DK1 lens are inserted. The lenses were positioned carefully so that the images are

Figure 5: ModularHMD assembly

in focus from the user's eyes and a 60-degree FOV (horizontal and vertical) is provided approximately for each eye. See Fig. 16 in the appendix for more details. Fig. 5(b) shows the components of the bottom module. Similar to the side module, the module holder has two sliding rails. Because the bottom module is seen by both eyes, two lenses are set in it. Fig. 5(c) shows the components of a central VR screen. A VRM-100 screen with two DK2 lenses was used to make a central VR module. The sliding slot was not implemented for the central VR screen. We added a short slider for the two lenses to adjust the interpupillary distance (IPD). Fig. 5(d) shows how to assemble all the components into a single HMD. Our 3D-printed module cowls were designed to connect together (bottom left), and the connected display part was attached to an Oculus head belt (0.8 kg).

3.5.3 Modular actions for removing, and remounting. Figures 6 (a) and (b) show how to remove the side and bottom modules. Both removal processes consist of a two-step action. To remove the side module, the HMD user first slides the lens knob to the back to remove it from view. Fig. 6 (a, left) visualizes this action. The component shown by the dotted lines indicates the parts originally occluded by the module. Then the user can pull it out from its holder by hand (Fig. 6 (a, right)) to obtain a full real-world vision. To remove the bottom module, Fig. 6 (b, left) shows an HMD user changing the bottom vision by sliding the two lenses from each

other, and Fig. 6 (b, right) shows a user grasping the display module and pulling it up.

3.5.4 Software. We developed software to manage all the ModularHMD functions with server-client architecture. Using Unity 2019.4.7f1, we developed Windows server software that drives the base HMD as well as iOS client software that drives each module. The server and the three iPod Touches were connected to exchange data by TCP/IP communication. We used Google Cardboard SDK to display binocular images on the module's screen in the virtual and VST modes (Figs. 4 (a)(b)(d)) and adjusted the IPD of the images using the Viewer Profile Generator¹. For the extended VR mode, additional VR views were automatically generated and presented in our Unity software. Mode switching can be triggered by a sensor program or by a user's explicit interaction (e.g., GUI button), which can be setup in advance for desired interactions. As shown in Figs. 4 (c)(e)(f), we also created software to allow cross-device interaction between the base HMD and the removed module devices. When a module is removed, it displays its VR content for cross-device interaction either automatically or manually, which can be setup in advance. Our user study (described below) employed the manual mode. The touches by the partner on the device module are sent to the server and visualized in the VR view, which allows both users to share their selected information. The bottom module is operated

¹<https://www.google.com/firbaseapp.com/>, last accessed: 2021-03-27.

Figure 6: How to remove each module

Figure 7: ModularHMD user's vision: (a) VR content binocular vision captured by a 360° camera; (b)-(c) a cropped set of 360° images taken from left eye's point of view; (b) VR content displayed by all screens; (c) VST displayed at side and bottom modules while watching VR content, (d) Side and bottom modules are removed while watching VR content.

similarly. As well as keyboard (Fig. 4 (c)), various types of software can be designed that fully leverage the always-connected modules.

3.6 Visual field of ModularHMD user

Figure 7 (a) shows an HMD user's visual field when wearing ModularHMD. The six circles represent a view through each lens. After the iterative design process, ModularHMD's front lens supports 56-degree horizontal and vertical FOVs. Each side module gives an additional peripheral FOV of 41 degrees, which is monocular vision. The bottom module offers an additional 41-degree bottom FOV supporting binocular (stereo) vision. We tried to maximize the continuity across lenses, but we could not avoid a small blind area (2 degrees for each side) at the intersection of the two lenses, whose effect will be tested later.

Figures 7(b)-(c) are a cropped set of 360° images taken from the left eye's point of view. Fig. 7 (b) illustrates how each display hole looks when a VR walkthrough application is presented. A total of six lenses display a wide range of VR content, where the main roles of the side and bottom holes provide users with more peripheral information of the VR world. Fig. 7 (c) illustrates when the user starts using the VST mode to ensure real-world peripheral awareness. The central holes continue to present the VR content, but the side hole enables awareness of the sides, e.g., a person sitting next to the user, and the bottom module provides desktop awareness, in which a cup on the desk is visible. Because the position and orientation of the external camera of each module do not fully align with the user's eye, the VST image looks slightly different from the real-world vision. Fig. 7 (d) illustrates the scenario when the user wants to obtain real-world visions by removing the side and

bottom modules. Different from Fig. 7 (c), the user can directly see and interact with real-world objects (e.g., cup, smartphone etc.) or his partner.

4 USER STUDY

4.1 Aim and overview

We conducted a user study to validate ModularHMD's four interactions ($In_1 - In_4$) using four different sessions (Sessions 1-4). In all the sessions, we avoided unfair comparison with regular HMDs because a wide-FOV commercial HMD is obviously stronger for In_1 and not fully capable of peripheral interactions $In_2 - In_3$. Instead, each session was differently designed to answer a specific core question regarding the main features of each interaction. In each session, participants used ModularHMD in different HMD boundary contexts and provided subjective (e.g., usability, etc.) and behavioral feedback (e.g., text typing, etc.). The order of the sessions was fixed to allow participants to progressively try more complex interactions with ModularHMD. Each participant took about 1.5 hours to complete the four sessions. Our study strictly adhered to governmental and institute COVID-19 prevention protocols. This study's design was officially approved by the local ethics committee board.

4.2 Participants

We recruited 12 participants (males: 7, females: 5, average age: 21.2) who had no knowledge of this study. The majority were experienced beginners of VR HMD. All had normal or corrected visual acuity. We did not invite people who wear glasses to eliminate any concern that

Figure 8: Experimental environment

Figure 9: Left-eye view of three HMD conditions in Session 1

different peripheral visions might depend on glass frame and lens shapes. Each participated in all sessions and got paid approximately US\$30.

4.3 Setup

The experimental setup is shown in Fig. 8. Each participant sat near a desk and wore ModularHMD. The IPD was adjusted to get a clear stereo-vision. Depending on the session, an assistant (acting as a chatting/collaboration partner) sat next to the participant. The main equipment was ModularHMD, a laptop computer for answering questionnaires, a smart device for Session 2, and a gamepad for Session 4. To counterbalance the forward-shifted weight of ModularHMD, we attached a 1.2-kg weight behind the head belt for comfortable wearing and easier head rotation. No participants complained about the weight during their short experiences.

4.4 Session 1: Immersion into VR

Session 1 investigated the immersive experience provided by ModularHMD and how additional peripheral screens and the gap between lenses affected such experiences.

4.4.1 Design. An independent variable was the HMD condition (See below and Fig. 9).

- *Removed* (Fig. 9(a)): Central screen displayed VR and all modules were removed (i.e., cowless HMD).

- *Immersive* (Fig. 9(b)): Central screen displayed VR and all modules were occluding panels (i.e., regular HMD).
- *Extended Immersive* (Fig. 9(c)): Central screen displayed VR and all modules displayed its extended views.

Each participant watched a 1-minute 360-degree walkthrough VR video using one of the HMD conditions. The dependent variable was General Presence; a sub-scale of Igroup Presence Questionnaire (IPQ) [66]. This was measured after the video-watching task. After all the trials, we conducted self-structured interviews that mainly asked about the gap between lenses.

4.4.2 Results. Figure 10 shows the results of the General presence sub-scale of IPQ scores. Depending on the data normality checked by the Kolmogorov–Smirnov test, we used the parametric and non-parametric tests differently. For General Presence, a Freedman test was applied, which found a main effect of HMD condition ($\chi^2(2) = 17.1, p < .001$). A post-hoc analysis using the Scheffe method found that *Immersive* and *Extended Immersive* scores were significantly higher than *Removed* ($p = .05$, and $p < .001$). While no difference was found between *Immersive* and *Extended Immersive*, both exceeded the neutral score, and the latter was the higher of the two. From post-interviews, we obtained many positive reflections for *Extended Immersive*: “The side screens helped increase the FOV when I was looking at the peripheral areas,” and “the gap between the lenses was recognizable but not critical.” However, we

Figure 10: IPQ sub-scale results: (* $p < .05$, * $p < .001$). All bar charts show mean with standard deviation**

also obtained an enlightening comment: “*The Immersive condition was more natural because the view was more converged.*”

4.4.3 Discussion. First, since the *Removed* condition (i.e., the non-cowl HMD) failed to induce an immersive experience, it is not a valid solution. Second, both *Immersive* and *Extended Immersive* provided immersive VR experiences. Last, although *Extended Immersive* might be an upgrade of *Immersive*, its updated benefit was insignificant. From interviews with participants, the small, noticeable physical gaps between the central and peripheral module screens slightly influenced the view continuity. In summary, ModularHMD successfully achieves In_1 , where an extended VR view is a recommended function, but perhaps those who require a more converged view might turn it off.

4.5 Session 2: Ad-hoc around-hand object interaction

Session 2 investigated how ModularHMD helps users perform ad-hoc interactions with real-world peripherals and objects around their hands. To examine this, this session focused on an ad-hoc text-typing task while wearing an HMD. This task has been recognized as challenging [19, 37] in mobile setups as well as a standard benchmark of subtle interaction.

4.5.1 Design. To simulate ad-hoc text-typing during a VR experience, we designed a sequential task. Step 1: an HMD user was immersed in a VR content. Step 2: she received a chat message in the VR and needed to reply using a given keyboard. Step 3: after she replied, she returned to the full VR experience. To roughly calculate the WPM values, in Step 2, participants randomly input ten phrases extracted from [41].

The independent variable was the HMD conditions (VST, Removed). Regular HMDs were not examined because their static cowls are incapable of peripheral real-world interaction. VST is a common solution for observing real-world objects. Removed is our new solution provided by ModularHMD. The detailed method for each condition is described below and shown in Fig. 11. The dependent variables were all recorded actions, including text-typing performance in words per minute (WPM) and the time required by the modular actions as well as the users’ subjective responses to the system usability scale (SUS). The order of the three HMD conditions was counterbalanced.

- **VST (Fig.11(a)):** In Step 1, users wore the ModularHMD and had a smartphone (i.e., an iPod Touch) in their pocket for simulating that they are using their own smartphones. ModularHMD’s side modules were covered by occluding plates. In Step 2, they turned on the VST for the bottom module and took their smartphone from their pocket. The target texts to type were displayed on the central VR screen, and they used the smartphone’s touch keyboard to type the texts as rapidly and accurately as possible. In Step 3, they deactivated the VST mode and returned their smartphone to their pocket.
- **Reused (Fig. 11(b)):** In Step 1, users wore the ModularHMD. In Step 2, they removed the bottom module and reused it as a hand-held keyboard. The target texts were displayed in the central VR view. In Step 3, they remounted the module to the bottom holder. Unlike VST, full real-world vision was provided to support correct eye-body coordination perception without a visual delay.
- **Non-HMD** was prepared as a reference to extract the user’s fundamental text-typing performances. Only Step 2 was tested. Fig.11(c) shows the users performing the same text-typing tasks with the same smartphone touch keyboard. Here, the target texts and the users’ input texts were given at the frontal LCD so that the postures of the users were identical with the other HMD conditions.

4.5.2 Results. The results are summarized in Fig. 12. Since all the data had a normal distribution, we used a parametric test. Fig. 12 (left) shows the SUS result. A paired t-test found that *Reused* offered a significantly higher SUS than *VST* ($t(11) = 8.69, p < .001, r = .93$). Fig. 12 (right) illustrates the WPM results, with reference data of the *Non-HMD*. A one-way RMANOVA found a main effect ($F_{2,22} = 85.0, p < .001, \eta^2 = .77$) of HMD condition. A post-hoc test with Bonferroni correction found significant differences among all the conditions and *Reused* was the second best: (*VST* vs. Others: $p < 0.001$, *Reused* vs. *Non-HMD*: $p < .001$). We performed a set of video-codings to evaluate the time required for mode switching. *VST* respectively required $M = 11.1, SD = 2.1$ seconds and $M = 8.36, SD = 2.45$ seconds to prepare the hand-held keyboard for text typing and to return the device to the pocket. *Reused* required $M = 5.46, SD = 1.62$ seconds to start the text typing and $M = 7.67, SD = 1.78$ to remount the device. When the module slid smoothly without jamming, the users prepared it within five seconds. Overall, *Reused* was enabled within an acceptable amount of time.

4.5.3 Discussion. The reasons for the lower SUS scores in the *VST* condition are obviously their current technical limitations: visual quality and excessive delays. Some participants felt a considerable delay between the rendered hand motions and their own hand sensations. However, they basically agreed that VST would be efficient for quickly obtaining spatial awareness in simple object control tasks. *Reused* offered a higher SUS score because natural typing experience was provided with a directly visible, hand-held device and their own fingers. The SUS result correlated with the WPM results; a lower WPM of VST was due to lower usability. The difference between *Reused* and *Non-HMD* was unexpected result because both used direct visions. We observed that users sometimes unnecessarily moved their heads when looking at the bottom opening

Figure 11: HMD Conditions in Session 2

Figure 12: SUS and WPM results (* $p < .05$, ** $p < .01$, *** $p < .001$). All bar charts show mean with standard deviation.

slot. Such an unnecessary motion might have affected their typing performances. Yet *Reused* offered an 81 % as good as the performance of the *Non-HMD* condition, which clearly proves benefits of ModularHMD's modular interaction in mobile work contexts.

4.6 Session 3: Ad-hoc face-to-face conversation

Session 3 investigated how ModularHMD helps ad-hoc face-to-face conversation between HMD and non-HMD users. We primarily explored how they read each other's non-verbal facial information (e.g., expression, eye contact etc.) and how they had natural conversations.

4.6.1 Design. To simulate ad-hoc chat between HMD and non-HMD users, we designed a sequential task. Step 1: the HMD user was immersed in VR content. Step 2: a non-HMD (e.g., a colleague) approached her and they had a short, 1-minute chat. Step 3: when the chat was over, both users returned to their original tasks.

The independent variables were the HMD mode (VST, Removed) and the user role (HMD user, non-HMD user), summarized below with detailed instructions and task sequences. Similar to Session 2, regular HMDs were not included due to their limited ability, and we tested VST because it is a current practical solution to mitigate the boundary problem. The dependent variables were SUS, two 7-Likert scale questions about facial expression readability, "I could readily tell when my partner was listening to me", and conversation naturalness, "This felt like a natural conversation", extracted from a previous work [15]. These questions were answered after each independent variable, and SUS was measured after each HMD trials. To simulate a short chat, an experimental assistant acted with the

participant and together freely decided the topic of the conversation. The presenting order of the HMD mode and the roles were counterbalanced.

- *VST* (Fig. 13(a)): *HMD user role*: In Step 1, the user wore ModularHMD to watch a VR video. In Step 2, he activated the VST of one of the side modules and chatted with a non-HMD partner. As shown in Fig. 13(a, left), the HMD user was able to see his partner through the side VST. In Step 3, the user returned to watching the VR by disabling the VST. *Non-HMD user role*: In Step 1, the user was away from the HMD user. In Step 2, he approached the HMD user for a short chat. As shown in Fig. 13 (a, right), the non-HMD user saw that the HMD user continued to wear the HMD. In Step 3, the user said good bye and walked away.
- *Removed* (Fig. 13(b)): The process of the *Removed* HMD condition was basically identical with the VST condition. Two differences were set; the HMD user was instructed to remove the side module in Step 2 and remount it in Step 3. This allowed both users to directly read the facial expressions of their partner.

4.6.2 Results. Since no normality was observed, a Wilcoxon Signed-Rank test was applied. *Removed* significantly helped the non-HMD read the HMD user's status more than VST ($z = 2.19$, $p < .05$) in the non-HMD condition. Here the average VST score was below a neutral score of 4. Since VST did not provide any information to the non-HMD users, this result is obvious and demonstrates the fundamental benefit of a physical opening for both users. In contrast, we found no main effect of naturalness for both variables: HMD-VST: 5.92, $SD = 1.08$, HMD-Removed: 5.92, $SD = 1.44$, Non-HMD-VST: 5.08, $SD = 1.88$, Non-HMD-Removed: 5.83, $SD = 1.40$. Fig. 14(b) shows the SUS results, and VST was significantly more useful than *Removed*. Our video-coding shows the average mode switching times were 1.46($SD = 0.47$) seconds for VST, 5.65($SD = 1.247$) seconds for removing, and 11.3($SD = 2.90$) seconds for remounting.

4.6.3 Discussion. Most non-HMD users experienced difficulty with VST because it provided no facial awareness of the HMD user. In contrast, *Removed* explicitly allowed non-HMD users to read how the HMD user behaves. However, this HMD condition did not make a significant difference in the conversation naturalness. This result is unexpected, which we thoroughly explored in our post-interviews. First, our participants felt that the verbal information was sufficient for smooth turn-taking. Some accepted such asymmetric communication (i.e., conversant faces are unavailable) based

Figure 13: HMD conditions in Session 3: Left image for (a) and (b) is HMD user's view (taken from a 360° camera), and right is HMD user appearance from non-HMD user perspective.

on previous similar online meeting experiences without video. Second, some thought that other types of non-verbal information like nodding by the HMD user conveyed explicit feedback. Third, others complained that the side aperture was too small to see the HMD user's entire face. Thus, the best achievement of the current ModularHMD might be supporting eye contact for fundamental social etiquette. Regarding SUS, VST was more positive, due to its simpler double-tapping gesture command. In summary, these results suggest that VST is efficient for quick chats, and Removed is more recommended for face-to-face interaction. We also received an interesting comment: "Without reusing the interaction, the module's removing and restoring was a bit cumbersome." The "reusing" action that is missing in this session might increase the benefit of physically re-configuring the HMD form.

4.7 Session 4: Ad-hoc cross-device collaboration

Session 4 assessed how ModularHMD supports ad-hoc cross-device collaboration between HMD and non-HMD users. Specifically, we are interested in how well the HMD users can share VR content with non-HMD users and how well the non-HMD users can join the VR content interaction. To understand the effect of this interaction that involves various benefits, we focused on the participants' absolute subjective responses that reflect their total user experiences rather than examining individual benefits. Similar to other sessions, we did not test regular HMDs to avoid unfair interactions. For example, Oculus software only supports unidirectional mirroring to a connected device. Moreover, formal comparisons cannot be formulated even with other functional mobile HMD (e.g., FaceDisplay) due to the differences in setups and benefits.

4.7.1 Design. To simulate ad-hoc cross-device collaboration, we designed a sequential scenario with an HMD user (designer) and a non-HMD user (manager). The initial step immersed the designer

Figure 14: Questionnaire results: (* $p < .05$) All bar charts show mean with standard deviation.

into a VR content editor. The sequence unfolded as follows (Fig. 15). Step 1: a manager stopped by the designer's workplace to discuss his progress. Step 2: he removed a side module to initiate a face-to-face conversation. Step 3: he gave it to his manager to automatically share the HMD's VR content with the module screen. Both users exchange facial expressions through the opened-side module slot. Step 4: the manager joined the VR interaction using the borrowed module device and instructed the designer to change the colors of a couple of the 3D models. The pointing information from both users was visualized on both the VR and module screens (Figs. 4(e) and (f)). Step 5: the manager returned the module to the designer after the collaboration was over. Step 6: the designer remounted the module and returned to a full VR experience.

Each participant experienced both users' perspectives: HMD user (designer) and non-HMD user (manager). Our assistant played the role of one of the users. The independent variables were SUS and a 7 Likert scale question asking "How smoothly was the collaboration initiated and terminated?" The participants also ranked the three possible benefits in order of importance to identify the most powerful one. (A) The HMD user can recognize the behaviors of the non-HMD user; (B) The non-HMD user can recognize the HMD user's facial behaviors; (C) The non-HMD user can smoothly join the HMD user's VR interaction.

4.7.2 Results and discussion. The SUS result was $M = 76.9, SD = 16.4$, which falls within good/excellent levels [3]. The average rating for the smoothness of initiation and the termination of collaboration was $M = 5.33, SD = 1.50$, 1.3 points greater than a neutral score. These two results are early pieces of evidence that suggests that ModularHMD helps sequential ad-hoc cross-device collaboration. Regarding the ranking of the three benefits, 9 of 12 participants rated (C) as the best, 8 rated (A) as second, and 11 rated (B) as third. Thus, users primarily recognized the benefit of (C), which quickly allowed the non-HMD user to join the VR interaction. Merits (A) and (B) are socially critical, but once collaboration starts on the

Figure 15: Workflow of Session 4: ad-hoc cross-device collaboration between HMD and non-HMD users.

distributed workspaces, verbal channels and visualized interaction across devices more directly contribute to the efficient collaboration. Our recommendation is to keep the side slot open for exchanging facial expressions, especially for the start and the end of a collaboration. From the interviews, many participants were positive about the entire interaction process using our cyclic modular actions. We believe that the benefit of ModularHMD will peak in collaboration contexts because physically distributed workspaces, necessary face disclosures, interaction with real-world objects and partners, and workflow between two users are fully managed by a single set of modular actions.

5 GENERAL DISCUSSION: DESIGN INSIGHTS AND FUTURE DIRECTION

5.1 Insights and use recommendations

The following summarizes the answers from our prototype and user study about our goals:

- In_1 (Immersion into VR): ModularHMD users successfully became immersed in the VR content using occluding panels or extended VR views. Although extended VR views are recommended, in general, either options are usable.
- In_2 (Ad-hoc around-hand device interaction): ModularHMD removed/reused was more effective in completing delicate text typing compared to VST. However, VST could also be an option for quick or unsubtle object interactions.
- In_3 (Ad-hoc face-to-face conversation): ModularHMD removed function mainly supports eye-contact beyond the HMD boundary and improved the facial expression readability between HMD and non-HMD users.
- In_4 (Ad-hoc cross-device collaboration): ModularHMD removed/reused function successfully offered smooth and convenient cross-device interaction between HMD and non-HMD users. The function allowing non-HMD users to easily join VR interactions was the most critical aspect.

In summary, all of our target interactions were achieved by the current prototype, suggesting that a set of ModularHMD interactions will be a powerful solution to mitigating HMD boundary issues in the mobile VR context.

Possible scenarios with ModularHMD include combined use of HMD and physical touchscreens or keyboards (e.g., [5, 34, 50]). ModularHMD will definitely expand such hybrid interactions. For example, when temporal sketching or dedicate touch input is required in room-scale VR, our ModularHMD's ad-hoc device interaction by the removed module would be highly efficient. For collaborative VR interactions, ModularHMD can be used as a base station for

inviting up to three real-world players. This would be helpful in various situations beyond the office scenario we argued, such as airplane seats, VR demo exhibitions, VR games, etc.

Our design insights can be transferred to a commercial HMD with a narrower FOV (e.g., a smartphone VR). The challenge is adapting the module arrangements of the user's FOV; our design considerations are an initial guideline. If this challenge is solved, the modular mechanics can be feasibly installed into the cowl or head belt as extra HMD accessories. Furthermore, ModularHMD can be converted to FaceDisplay[24] or FrontFace [8] by changing each modular unit's view angle. This could drastically increase interaction opportunities as well as making the formation of ModularHMD and non-HMD users more flexible.

5.2 Limitations and future work

A weakness of our prototype is the gap between lenses, especially when displaying the extended VR view. To optically improve the continuity, perhaps Fresnel lenses [83] can be tried. For example, maybe the gaps can be eliminated through bent lens [56] or a lens array [60], which are being developed for wide-FOV HMDs.

Although miniaturizing ModularHMD is desired, it remains challenging and slightly conflicts with stable and physical modular interactions. Each module size could be variable, but our recommendation is to use smartphone-sized modules for successful peripheral subtle and cross-device interaction. Such miniaturization is an essential limitation of ModularHMD. One possible solution is future small, lightweight HMDs [2, 43] that enable a lighter and more compact ModularHMD with the same conceptual design. The second possible solution is making each module with the currently available flexible/bendable/foldable touchscreen devices, which could minimize the total size of ModularHMD and increase its fitness and design flexibility. If larger and lightweight touchscreens can be used as module devices, larger screen estates (i.e., larger VR content) and opening slots (i.e., larger face disclosure) can be ensured.

Some users struggled to remove and remount the modules. However, we do not believe that these issues are critical because those who just practiced these actions ten times drastically reduced their required times. Other ideas include modifying the structure of the sliding slot (e.g., more physical guides), automating the lens slider using motors, using the VST to make the fingers visible when removing the module, etc.

VST was useful for simple peripheral interactions. Our next version of ModularHMD might employ VST's collaborative and remove actions for complex scenarios (e.g., more people, more complex formations [27]). To do this, VST's performance is improving (e.g., [76]). Eventually, such improved VST will automatically and

progressively provide HMD users with situational awareness based on the proxemics of those around the HMD user as a clear intermediate state between a full VR and a partially opened HMD.

Our study has two limitations: first being the short-duration, controlled scenarios that were provided to each participant. Since we were able to obtain honest user feedback (see appendix) from our participants without social and experiential bias, we believe our study was successful in capturing the initial reflections for the idea and benefits of ModularHMD. However, this study design did not capture how the user could use the ModularHMD for long-term office scenarios. Further studies are necessary to understand how ModularHMD can manage complex and daily HMD boundary cases. The second limitation is lack of formal comparisons with other unique HMDs, such as MDI [74], FaceDisplay [24], or visualization approaches. We encountered difficulties in setting up these closely related HMDs as baselines because various elements coexist in the HMD boundary issues. While the current study went through the four interactions briefly, the next study design would deepen individual interactions and include more scenarios. Examples include single user's hybrid interaction with VR and touch surface (e.g. VR typing), and multi-user collaboration with more than three users, as extensions of Session 2 and 4, respectively.

6 CONCLUSION

We proposed ModularHMD, a new HMD concept using a modular mechanism to support ad-hoc peripheral interactions with real-world devices and partners while maintaining an immersive VR experience for HMD users. We validated our concept through a proof-concept prototype and a qualitative user study with common mobile work tasks. Results confirmed that ModularHMD successfully supported HMD user's real-world awareness and device control as well as ad-hoc cross-device collaboration with real-world partners. We discussed the technical challenges and insights for our next ModularHMD. Future work will include lens structure and opening slot modifications to improve both the VR experiences and real-world awareness.

ACKNOWLEDGMENTS

This work in part is supported by JSPS Kakenhi (20K21799) and the Cooperative Research Project Program of the Research Institute of Electrical Communication, Tohoku University.

REFERENCES

- [1] Acer. 2017. Windows Mixed Reality Headset. <https://www.acer.com/ac/en/US/content/series/wmr>. Retrieved September 14, 2020 from <https://www.acer.com/ac/en/US/content/series/wmr> Visited the 14/09/2020.
- [2] K. Bang, Y. Jo, M. Chae, and B. Lee. 2021. Lenslet VR: Thin, Flat and Wide-FOV Virtual Reality Display Using Fresnel Lens and Lenslet Array. *IEEE Transactions on Visualization and Computer Graphics* 27, 5 (2021), 2545–2554. <https://doi.org/10.1109/TVCG.2021.3067758>
- [3] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies* 4, 3 (May 2009), 114–123.
- [4] Verena Biener, Daniel Schneider, Travis Gesslein, Alexander Otte, Bastian Kuth, Per Ola Kristensson, Eyal Ofek, Michel Pahud, and Jens Grubert. 2020. Breaking the Screen: Interaction Across Touchscreen Boundaries in Virtual Reality for Mobile Knowledge Workers. *arXiv preprint arXiv:2008.04559* (2020).
- [5] S. Boustila, T. Guégan, K. Takashima, and Y. Kitamura. 2019. Text Typing in VR Using Smartphones Touchscreen and HMD. In *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces*. 860–861.
- [6] Sidney Bovet, Aidan Kehoe, Katie Crowley, Noirin Curran, Mario Gutierrez, Mathieu Meisser, Damien O. Sullivan, and Thomas Rouvinez. 2018. Using Traditional Keyboards in VR: SteamVR Developer Kit and Pilot Game User Study. In *Proceedings of the 2018 IEEE Games, Entertainment, Media Conference*. 132–135.
- [7] Pulkit Budhiraja, Rajinder Sodhi, Brett Jones, Kevin Karsch, Brian Bailey, and David Forsyth. 2015. Where's My Drink? Enabling Peripheral Real World Interactions While Using HMDs. *arXiv preprint arXiv:1502.04744* (2015).
- [8] Liwei Chan and Kouta Minamizawa. 2017. FrontFace: Facilitating Communication between HMD Users and Outsiders Using Front-Facing-Screen HMDs. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. Association for Computing Machinery, New York, NY, USA, Article 22, 5 pages. <https://doi.org/10.1145/3098279.3098548>
- [9] Hong-Yu Chang, Wen-Jie Tseng, Chia-En Tsai, Hsin-Yu Chen, Roshan Lalintha Peiris, and Liwei Chan. 2018. FacePush: Introducing Normal Force on Face with Head-Mounted Displays. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 927–935. <https://doi.org/10.1145/3242587.3242588>
- [10] Mariko Chiba, Wataru Yamada, and Hiroyuki Manabe. 2018. Transparent Mask: Face-Capturing Head-Mounted Display with IR Pass Filters. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings (UIST '18 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 149–151. <https://doi.org/10.1145/3266037.3271632>
- [11] Isamu Endo, Kazuki Takashima, Maakito Inoue, Kazuyuki Fujita, Kiyoshi Kiyokawa, and Yoshifumi Kitamura. 2021. A Reconfigurable Mobile Head-Mounted Display Supporting Real World Interactions. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, Article 418, 7 pages. <https://doi.org/10.1145/3411763.3451765>
- [12] Facebook. 2019. Oculus Quest: All-in-One VR Headset. <https://www.oculus.com/quest/>. Retrieved September 14, 2020 from <https://www.oculus.com/quest/> Visited the 14/09/2020.
- [13] Christian Frueh, Avneesh Sud, and Vivek Kwatra. 2017. Headset Removal for Virtual and Mixed Reality. In *Proceedings of the ACM SIGGRAPH 2017 Talks (SIGGRAPH '17)*. Association for Computing Machinery, New York, NY, USA, Article 80, 2 pages. <https://doi.org/10.1145/3084363.3085083>
- [14] Taichi Furukawa, Daisuke Yamamoto, Moe Sugawa, Roshan Peiris, and Kouta Minamizawa. 2019. TeleSight: Enabling Asymmetric Collaboration in VR between HMD User and Non-HMD Users. In *ACM SIGGRAPH 2019 Emerging Technologies (SIGGRAPH '19)*. Association for Computing Machinery, New York, NY, USA, Article 26, 2 pages. <https://doi.org/10.1145/3305367.3335040>
- [15] Maia Garau, Mel Slater, Simon Bee, and Martina Angela Sasse. 2001. The Impact of Eye Gaze on Communication Using Humanoid Avatars. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. Association for Computing Machinery, New York, NY, USA, 309–316. <https://doi.org/10.1145/365024.365121>
- [16] Sarthak Ghosh, Lauren Winston, Nishant Panchal, Philippe Kimura-Thollander, Jeff Hotnog, Douglas Cheong, Gabriel Reyes, and Gregory D. Abowd. 2018. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 24, 4 (2018), 1447–1456.
- [17] Jerônimo G. Grandi, Henrique G. Debarba, and Anderson Maciel. 2019. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces*. 127–135.
- [18] Jens Grubert and Matthias Kranz. 2017. HeadPhones: Ad Hoc Mobile Multi-Display Environments through Head Tracking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3966–3971. <https://doi.org/10.1145/3025453.3025533>
- [19] J. Grubert, E. Ofek, M. Pahud, and P. O. Kristensson. 2018. The Office of the Future: Virtual, Portable, and Global. *IEEE Computer Graphics and Applications* 38, 6 (2018), 125–133.
- [20] Jens Grubert, Lukas Witzani, Eyal Ofek, Michel Pahud, Matthias Kranz, and Per O. Kristensson. 2018. Effects of Hand Representations for Typing in Virtual Reality. In *Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces*. 151–158.
- [21] Jens Grubert, Lukas Witzani, Eyal Ofek, Michel Pahud, Matthias Kranz, and Per O. Kristensson. 2018. Text Entry in Immersive Head-Mounted Display-Based Virtual Reality Using Standard Keyboards. In *Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces*. 159–166.
- [22] Uwe Gruenefeld, Tim Claudius Stratmann, Wilko Heuten, and Susanne Boll. 2017. PeriMR: A Prototyping Tool for Head-Mounted Peripheral Light Displays in Mixed Reality. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17)*. Association for Computing Machinery, New York, NY, USA, Article 51, 6 pages. <https://doi.org/10.1145/3098279.3125439>
- [23] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4021–4033. <https://doi.org/10.1145/3025453.3025683>

- [24] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. 2018. FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 54:1–54:13. <https://doi.org/10.1145/3173574.3173628>
- [25] Jie Guo, Dongdong Weng, Zhenliang Zhang, Haiyan Jiang, Yue Liu, Yongtian Wang, and Henry B. Duh. 2019. Mixed Reality Office System Based on Maslow's Hierarchy of Needs: Towards the Long-Term Immersion in Virtual Environments. In *Proceedings of the 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 224–235.
- [26] Juergen Hagler, Michael Lankes, and Andrea Aschauer. 2018. The Virtual House of Medusa: Guiding Museum Visitors Through a Co-located Mixed Reality Installation. In *Joint International Conference on Serious Games*. 182–188.
- [27] Edward Twitchell Hall. 1966. *The hidden dimension*. Vol. 609. Garden City, NY: Doubleday.
- [28] Jeremy Hartmann, Christian Holz, Eyal Ofek, and Andrew D. Wilson. 2019. RealityCheck: Blending Virtual Environments with Situated Physical Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3290605.3300577>
- [29] Ian P Howard, Brian J Rogers, et al. 1995. *Binocular Vision and Stereopsis*. Oxford University Press, USA.
- [30] I Am Cardboard. 2017. Pocket 360. <https://www.imcardboard.com/pocket-360.html>. Retrieved September 14, 2020 from <https://www.imcardboard.com/pocket-360.html> Visited the 14/09/2020.
- [31] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. 2015. Dollhouse VR: A Multi-View, Multi-User Collaborative Design Workspace with VR Technology. In *SIGGRAPH Asia 2015 Emerging Technologies (SA '15)*. Association for Computing Machinery, New York, NY, USA, Article 8, 2 pages. <https://doi.org/10.1145/2818466.2818480>
- [32] Akira Ishii, Masaya Tsuruta, Ippei Suzuki, Shuta Nakamae, Junichi Suzuki, and Yoichi Ochiai. 2019. Let Your World Open: CAVE-Based Visualization Methods of Public Virtual Reality towards a Shareable VR Experience. In *Proceedings of the 10th Augmented Human International Conference 2019 (AH2019)*. Association for Computing Machinery, New York, NY, USA, Article 33, 8 pages. <https://doi.org/10.1145/3311823.3311860>
- [33] Kohei Kanamori, Nobuchika Sakata, Tomu Tominaga, Yoshinori Hijikata, Kensuke Harada, and Kiyoshi Kiyokawa. 2018. Obstacle Avoidance Method in Real Space for Virtual Reality Immersion. In *Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 80–89.
- [34] Youngwon R. Kim and Gerard J. Kim. 2016. HoVR-Type: Smartphone as a Typing Interface in VR Using Hovering. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST '16)*. Association for Computing Machinery, New York, NY, USA, 333–334. <https://doi.org/10.1145/2993369.2996330>
- [35] Naoki Kimura, Michinari Kono, and Jun Rekimoto. 2019. Deep Dive: Deep-Neural-Network-Based Video Extension for Immersive Head-Mounted Display Experiences. In *Proceedings of the 8th ACM International Symposium on Pervasive Displays (PerDis '19)*. Association for Computing Machinery, New York, NY, USA, Article 22, 7 pages. <https://doi.org/10.1145/3321335.3324932>
- [36] Jarrod Knibbe, Jonas Schjerlund, Mathias Petraeus, and Kasper Hornbæk. 2018. The Dream is Collapsing: The Experience of Exiting VR. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 483:1–483:13. <https://doi.org/10.1145/3173574.3174057>
- [37] Pascal Knierim, Valentin Schwind, Anna Maria Feit, Florian Nieuwenhuizen, and Niels Henze. 2018. Physical Keyboards in Virtual Reality: Analysis of Typing Performance and Effects of Avatar Hands. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 345:1–345:9. <https://doi.org/10.1145/3173574.3173919>
- [38] Robert E Kraut, Robert S Fish, Robert W Root, and Barbara L Chalfonte. 1990. Informal Communication in Organizations: Form, Function, and Technology. In *Human Reactions to Technology: Claremont Symposium on Applied Social Psychology*. 145–199.
- [39] Kyungmin Lim, Jaesung Lee, Kwanghyun Won, Nupur Kala, and Tammy Lee. 2020. A Novel Method for VR Sickness Reduction Based on Dynamic Field of View Processing. *Virtual Reality* 25 (2020), 331–340. <https://doi.org/10.1007/s10055-020-00457-3>
- [40] Jia-Wei Lin, Ping-Hsuan Han, Jiun-Yu Lee, Yang-Sheng Chen, Ting-Wei Chang, Kuan-Wen Chen, and Yi-Ping Hung. 2017. Visualizing the Keyboard in Virtual Reality for Enhancing Immersive Experience. In *ACM SIGGRAPH 2017 Posters (SIGGRAPH '17)*. Association for Computing Machinery, New York, NY, USA, Article 35, 2 pages. <https://doi.org/10.1145/3102163.3102175>
- [41] I. Scott MacKenzie and R. William Soukoreff. 2003. Phrase Sets for Evaluating Text Entry Techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems (CHI EA '03)*. Association for Computing Machinery, New York, NY, USA, 754–755. <https://doi.org/10.1145/765891.765971>
- [42] Christian Mai, Lukas Rambold, and Mohamed Khamis. 2017. TransparentHMD: Revealing the HMD User's Face to Bystanders. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM '17)*. Association for Computing Machinery, New York, NY, USA, 515–520. <https://doi.org/10.1145/3152832.3157813>
- [43] Andrew Maimone and Junren Wang. 2020. Holographic Optics for Thin and Lightweight Virtual Reality. *ACM Transaction on Graphics* 39, 4, Article 67 (July 2020), 14 pages. <https://doi.org/10.1145/3386569.3392416>
- [44] Nicolai Marquardt, Frederik Brudy, Can Liu, Ben Bengler, and Christian Holz. 2018. SurfaceConstellations: A Modular Hardware Platform for Ad-Hoc Reconfigurable Cross-Device Workspaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 354:1–354:14. <https://doi.org/10.1145/3173574.3173928>
- [45] Eriko Maruyama and Jun Rekimoto. 2019. CompoundDome: A Wearable Dome Device That Enables Interaction with the Real World by Controlling the Transparency of the Screen. In *Proceedings of the 10th Augmented Human International Conference 2019 (AH2019)*. Association for Computing Machinery, New York, NY, USA, Article 40, 4 pages. <https://doi.org/10.1145/3311823.3311827>
- [46] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2143–2152. <https://doi.org/10.1145/2702123.2702382>
- [47] Daniel Medeiros, Rafael dos Anjos, Nadia Pantidi, Kun Huang, Mauricio Sousa, Craig Anslow, and Joaquim Jorge. 2021. Promoting Reality Awareness in Virtual Reality through Proxemics (*IEEE VR '21*). *IEEE*, 21–30. <https://doi.org/10.1109/VR50410.2021.00022>
- [48] Tim Menzner, Alexander Otte, Travis Gesslein, Jens Grubert, Philipp Gagel, and Daniel Schneider. 2019. A Capacitive-sensing Physical Keyboard for VR Text Entry. In *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces*. 1080–1081.
- [49] Microsoft. 2016. Microsoft HoloLens | Mixed Reality Technology for Business. <https://www.microsoft.com/en-us/hololens>. Retrieved September 14, 2020 from <https://www.microsoft.com/en-us/hololens> Visited the 14/09/2020.
- [50] Mark Mine, Arun Yoganandan, and Dane Coffey. 2014. Making VR Work: Building a Real-World Immersive Modeling Application in the Virtual World. In *Proceedings of the 2nd ACM Symposium on Spatial User Interaction (SUI '14)*. Association for Computing Machinery, New York, NY, USA, 80–89. <https://doi.org/10.1145/2659766.2659780>
- [51] Masaaki Murakami, Kosuke Kikui, Katsuhiro Suzuki, Fumihiko Nakamura, Masaaki Fukuoka, Katsutoshi Masai, Yuta Sugiura, and Maki Sugimoto. 2019. AffectiveHMD: Facial Expression Recognition in Head Mounted Display Using Embedded Photo Reflective Sensors. In *Proceedings of the ACM SIGGRAPH 2019 Emerging Technologies (SIGGRAPH '19)*. Association for Computing Machinery, New York, NY, USA, Article 7, 2 pages. <https://doi.org/10.1145/3305367.3335039>
- [52] nintendo-labo. 2018. Nintendo Switch Lab VR kit. <https://labo.nintendo.com/kits/vr-kit/>. Retrieved September 14, 2020 from <https://labo.nintendo.com/kits/vr-kit/> Visited the 14/09/2020.
- [53] Eyal Ofek, Jens Grubert, Michel Pahud, Mark Phillips, and Per Ola Kristensson. 2020. Towards a Practical Virtual Office for Mobile Knowledge Workers. *arXiv preprint arXiv:2009.02947* (2020).
- [54] Joseph O'Hagan and Julie R. Williamson. 2020. Reality Aware VR Headsets. In *Proceedings of the 9th ACM International Symposium on Pervasive Displays (PerDis '20)*. Association for Computing Machinery, New York, NY, USA, 9–17. <https://doi.org/10.1145/3393712.3395334>
- [55] J. Orlosky, T. Toyama, K. Kiyokawa, and D. Sonntag. 2015. ModulAR: Eye-Controlled Vision Augmentations for Head Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics* 21, 11 (2015), 1259–1268.
- [56] Panasonic. 2017. CES 2017 | Press Kits | Panasonic Newsroom Global - Now Get Double the Display with VR Headset. <https://news.panasonic.com/global/presskits/ces2017/>. Retrieved March 22, 2021 from <https://news.panasonic.com/global/presskits/ces2017/> Visited the 22/03/2021.
- [57] Daniel Pohl and Carlos F. de Tejada Quemada. 2016. See What I See: Concepts to Improve the Social Acceptance of HMDs. In *Proceedings of the 2016 IEEE Virtual Reality*. 267–268. <https://doi.org/10.1109/VR.2016.7504756>
- [58] Ismo Rakkolainen, Roope Raisamo, Matthew Turk, and Tobias Höllerer. 2017. Field-of-View Extension for VR Viewers. In *Proceedings of the 21st International Academic Mindtrek Conference (AcademicMindtrek '17)*. Association for Computing Machinery, New York, NY, USA, 227–230. <https://doi.org/10.1145/3131085.3131088>
- [59] Majken K. Rasmussen, Esben W. Pedersen, Marianne G. Petersen, and Kasper Hornbæk. 2012. Shape-Changing Interfaces: A Review of the Design Space and Open Research Questions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. Association for Computing Machinery, New York, NY, USA, 735–744. <https://doi.org/10.1145/2207676.2207781>
- [60] J. Ratcliff, A. Supikov, S. Alfaro, and R. Azuma. 2020. ThinVR: Heterogeneous microlens arrays for compact, 180 degree FOV VR near-eye displays. *IEEE Transactions on Visualization and Computer Graphics* 26, 5 (2020), 1981–1990.

- <https://doi.org/10.1109/TVCG.2020.2973064>
- [61] Jun Rekimoto, Keishiro Uragaki, and Kenjiro Yamada. 2018. Behind-the-Mask: A Face-through Head-Mounted Display. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces (AVI '18)*. Association for Computing Machinery, New York, NY, USA, Article 32, 5 pages. <https://doi.org/10.1145/3206505.3206544>
- [62] Jannick P. Rolland and Henry Fuchs. 2000. Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization. *Presence* 9, 3 (2000), 287–309.
- [63] Anastasia Ruvimova, Junhyeok Kim, Thomas Fritz, Mark Hancock, and David C. Shepherd. 2020. "Transport Me Away": Fostering Flow in Open Offices through Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 595:1–595:14. <https://doi.org/10.1145/3313831.3376724>
- [64] Dieter Schmalstieg and Tobias Höllerer. 2016. *Augmented Reality: Principles and Practice*. Addison-Wesley Professional.
- [65] Daniel Schneider, Alexander Otte, Travis Gesslein, Philipp Gagel, Bastian Kuth, Mohamad S. Damlakhi, Oliver Dietz, Eyal Ofek, Michel Pahud, Per O. Kristensson, Jörg Müller, and Jens Grubert. 2019. ReconViguation: Reconfiguring Physical Keyboards in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 25, 11 (2019), 3190–3201.
- [66] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The Experience of Presence: Factor Analytic Insights. *Presence: Teleoper. Virtual Environ.* 10, 3 (June 2001), 266–281. <https://doi.org/10.1162/105474601300343603>
- [67] Teddy Seyed, Xing-Dong Yang, and Daniel Vogel. 2017. A Modular Smartphone for Lending. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. Association for Computing Machinery, New York, NY, USA, 205–215. <https://doi.org/10.1145/3126594.3126633>
- [68] Lior Shapira and Daniel Freedman. 2016. Reality Skins: Creating Immersive and Tactile Virtual Environments. In *Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 115–124.
- [69] Adalberto L. Simeone, Eduardo Velloso, and Hans Gellersen. 2015. Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3307–3316. <https://doi.org/10.1145/2702123.2702389>
- [70] StarVR Corp. 2018. StarVR. <https://www.starvr.com/>. Retrieved April 2, 2021 from <https://www.starvr.com/> Visited the 02/04/2021.
- [71] Ata Tara, Gillian Lawson, and Alayna Renata. 2021. Measuring magnitude of change by high-rise buildings in visual amenity conflicts in Brisbane. *Landscape and Urban Planning* 205 (2021), 103930. <https://doi.org/10.1016/j.landurbplan.2020.103930>
- [72] Justus Thies, Michael Zollhöfer, Marc Stamminger, Christian Theobalt, and Matthias Nießner. 2018. FaceVR: Real-Time Gaze-Aware Facial Reenactment in Virtual Reality. *ACM Transactions on Graphics*. 37, 2, Article 25 (June 2018), 15 pages.
- [73] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. 2020. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery, New York, NY, USA, 182–195. <https://doi.org/10.1145/3379337.3415827>
- [74] Kien T. P. Tran, Sungchul Jung, Simon Hoermann, and Robert W. Lindeman. 2019. MDI: A Multi-channel Dynamic Immersion Headset for Seamless Switching between Virtual and Real World Activities. In *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces*. 350–358.
- [75] Wen-Jie Tseng, Li-Yang Wang, and Liwei Chan. 2019. FaceWidgets: Exploring Tangible Interaction on Face with Head-Mounted Displays. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 417–427. <https://doi.org/10.1145/3332165.3347946>
- [76] Varjo. 2021. Varjo XR-3 - The industry's highest resolution XR headset | Varjo. <https://varjo.com/products/xr-3/>. Retrieved July 25, 2021 from <https://varjo.com/products/xr-3/> Visited the 25/07/2021.
- [77] VeeR. 2016. VeeR MINI VR Goggle. <https://store.veer.tv/products/veer-mini-vr-goggles>. Retrieved September 14, 2020 from <https://store.veer.tv/products/veer-mini-vr-goggles> Visited the 14/09/2020.
- [78] Julius von Willich, Markus Funk, Florian Müller, Karola Marky, Jan Riemann, and Max Mühlhäuser. 2019. You Invaded My Tracking Space! Using Augmented Virtuality for Spotting Passersby in Room-Scale Virtual Reality. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*. Association for Computing Machinery, New York, NY, USA, 487–496. <https://doi.org/10.1145/3322276.3322334>
- [79] Chiu-Hsuan Wang, Seraphina Yong, Hsin-Yu Chen, Yuan-Syun Ye, and Liwei Chan. 2020. HMD Light: Sharing In-VR Experience via Head-Mounted Projector for Asymmetric Interaction. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20)*. Association for Computing Machinery, New York, NY, USA, 472–486. <https://doi.org/10.1145/3379337.3415847>
- [80] Julie R. Williamson, Mark McGill, and Khari Outram. 2019. PlaneVR: Social Acceptability of Virtual Reality for Aeroplane Passengers. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 80:1–80:14. <https://doi.org/10.1145/3290605.3300310>
- [81] Fei Wu, Jerald Thomas, Shreyas Chinnola, and Evan Suma Rosenberg. 2020. Comparison of Audio and Visual Cues to Support Remote Guidance in Immersive Environments. In *ICAT-EGVE 2020 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, Ferran Argelaguet, Ryan McMahan, and Maki Sugimoto (Eds.). The Eurographics Association. <https://doi.org/10.2312/egve.20201266>
- [82] Robert Xiao and Hrvoje Benko. 2016. Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1221–1232. <https://doi.org/10.1145/2858036.2858212>
- [83] Wataru Yamada and Hiroyuki Manabe. 2016. Expanding the Field-of-View of Head-Mounted Displays with Peripheral Blurred Images. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 141–142. <https://doi.org/10.1145/2984751.2985735>
- [84] Keng-Ta Yang, Chiu-Hsuan Wang, and Liwei Chan. 2018. ShareSpace: Facilitating Shared Use of the Physical Space by Both VR Head-Mounted Display and External Users. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST '18)*. Association for Computing Machinery, New York, NY, USA, 499–509. <https://doi.org/10.1145/3242587.3242630>

A PROTOTYPING

Figure 16: Detailed Layout of ModularHMD prototype

B USER FEEDBACK

Table 2: User feedback

	Comments
Session 1	"In the Removed condition, when I turned my gaze to the edge I could see the real world, this made me feel less immersed.", "In the VR condition, I felt like I had a wider field of view than with a regular HMD, especially when I tried to turn around.", "In the VR condition, I felt like the scale of the front and side images were slightly off."
Session 2	"In the Reused condition, it was slightly troublesome to remove the module.", "In the VST conditions, I could read the characters on the keyboard.", "VST condition takes less time to switch to, so if I'm in a hurry I would use VST whenever possible."
Session 3	"In both VST and Removed conditions, the non-HMD users were sufficiently visible to where I did not notice any difference.", "In the VST condition, I was worried because I didn't know if HMD users were looking at me(non-HMD users)", "I couldn't see the HMD user's mouth, but I didn't feel any discomfort."
Session 4	"HMD users can share their work easily and instantly, which is very useful.", "When collaborating, it is important to be able to see the non-HMD user, because the HMD user will behave differently depending on who the non-HMD user is.", "In this situation, as a non-HMD user I was more interested in the content, so I didn't look at the HMD user's eyes."