



BlendMR: A Computational Method to Create Ambient Mixed Reality Interfaces

VIOLET YINUO HAN, Carnegie Mellon University, USA

HYUNSUNG CHO, Carnegie Mellon University, USA

KIYOSU MAEDA, University of Tokyo, Japan

ALEXANDRA ION, Carnegie Mellon University, USA

DAVID LINDLBAUER, Carnegie Mellon University, USA

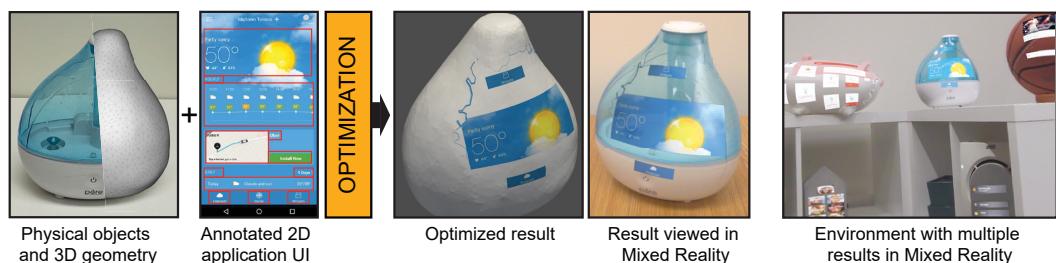


Fig. 1. We present an automatic approach for integrating virtual 2D applications into physical objects. Our approach takes the geometry of a physical object and an annotated 2D application as input, and produces a mapping of important sub-elements of the application to the geometry. The result can then be viewed in Mixed Reality. Right shows live view of the results as shown in an Varjo XR-3 headset.

436

Mixed Reality (MR) systems display content freely in space, and present nearly arbitrary amounts of information, enabling ubiquitous access to digital information. This approach, however, introduces clutter and distraction if too much virtual content is shown. We present BlendMR, an optimization-based MR system that blends virtual content onto the physical objects in users' environments to serve as ambient information displays. Our approach takes existing 2D applications and meshes of physical objects as input. It analyses the geometry of the physical objects and identifies regions that are suitable hosts for virtual elements. Using a novel integer programming formulation, our approach then optimally maps selected contents of the 2D applications onto the object, optimizing for factors such as importance and hierarchy of information, viewing angle, and geometric distortion. We evaluate BlendMR by comparing it to a 2D window baseline. Study results show that BlendMR decreases clutter and distraction, and is preferred by users. We demonstrate the applicability of BlendMR in a series of results and usage scenarios.

CCS Concepts: • Human-centered computing → Mixed / augmented reality; • Computing methodologies → Perception.

Additional Key Words and Phrases: augmented reality, adaptive user interfaces, computational interaction

Authors' addresses: **Violet Yinuo Han**, yinuoh@andrew.cmu.edu, Carnegie Mellon University, Pittsburgh, USA; **Hyunsung Cho**, hyunsung@cs.cmu.edu, Carnegie Mellon University, Pittsburgh, USA; **Kiyosu Maeda**, kiyosu775@g.ecc.u-tokyo.ac.jp, University of Tokyo, Tokyo, Japan; **Alexandra Ion**, alexandraion@cmu.edu, Carnegie Mellon University, Pittsburgh, USA; **David Lindlbauer**, davidlindlbauer@cmu.edu, Carnegie Mellon University, Pittsburgh, USA.



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

© 2023 Copyright held by the owner/author(s).

2573-0142/2023/12-ART436

<https://doi.org/10.1145/3626472>

ACM Reference Format:

Violet Yinuo Han, Hyunsung Cho, Kiyosu Maeda, Alexandra Ion, and David Lindlbauer. 2023. BlendMR: A Computational Method to Create Ambient Mixed Reality Interfaces. *Proc. ACM Hum.-Comput. Interact.* 7, ISS, Article 436 (December 2023), 25 pages. <https://doi.org/10.1145/3626472>

1 INTRODUCTION

Mixed Reality (MR) has the potential to change how users interact with digital information. Currently, user interfaces are presented on the displays of desktop computers, smartphones, or tablets. With MR, and especially with future always-on head-mounted displays (HMDs) that are light-weight and worn everyday, personalized virtual contents can be displayed freely in space, and include nearly arbitrary amounts of information. This enables more flexible access to information, since users can keep digital information directly in their environments, similar to current ambient displays and decorative objects like clocks or digital picture frames.

Current solutions to MR interfaces, however, largely focus on displaying 2-dimensional windows in specific areas of users' environment (e.g., [7, 8, 12, 31, 35]). Their placement relies on factors such as the furniture in a room [19], or the semantic connection between the physical and virtual objects [7]. While this simplifies access to information, having many virtual interfaces in an environment can lead to perceived clutter and distraction. Current MR interfaces lack the unobtrusive qualities of peripheral objects and ambient displays, that can be integrated with environments without causing clutter, even when the user is not actively interacting with them. This challenge is exacerbated if MR systems aim at enabling continuous access to the dozens of applications an average smartphone user has installed [21]. While providing users with menus to toggle the visibility of individual applications and contents could alleviate this challenge, this approach re-introduces the need for users to search for desired information explicitly. This eliminates the key benefit of simple and unobtrusive access to information offered by MR.

To address this challenge, we present BlendMR, a novel approach for creating ambient MR interfaces. Users can assign existing 2D applications to physical objects by placing them next to physical objects. BlendMR then blends the information of virtual interfaces onto the geometry of the physical objects. Our approach preserves MR system's continuous access to information, while reducing its obtrusiveness by offering an alternative form of information display for peripheral interfaces that users are not currently interacting with.

Scenario. Figure 2 provides a walkthrough of BlendMR's example usage. Situated in a future where MR glasses are lightweight and for daily wear, a user is relaxing in their living room reading a magazine. BlendMR displays glanceable information as ambient information embedded in the environment. Hence, the user is able to perform daily tasks in this always-on MR environment without being exposed to visual clutter and distraction from the MR interface. The user has manually chosen mappings between interfaces and objects that are meaningful and convenient to them: a sporting event interface is displayed on a soccer ball, allowing them to quickly access game scores and schedules; a weather interface is displayed on a humidifier, allowing them to keep track of the temperature. The user becomes interested in a sporting event displayed on a soccer ball. They stand up and approach the interface, and put their hand next to the MR interface. Their proximity triggers the application to go into full display mode, so the user can interact with the application, such as purchasing a ticket to the sporting event of their interest. After performing the interactions, the user goes back to the sofa and continues flipping through the magazine, while BlendMR switches the representation back to the ambient display.

Approach. BlendMR takes as input a 2D smartphone application, and a 3D scan of an arbitrary object in a room. Users can decide what physical object an application should be mapped onto,

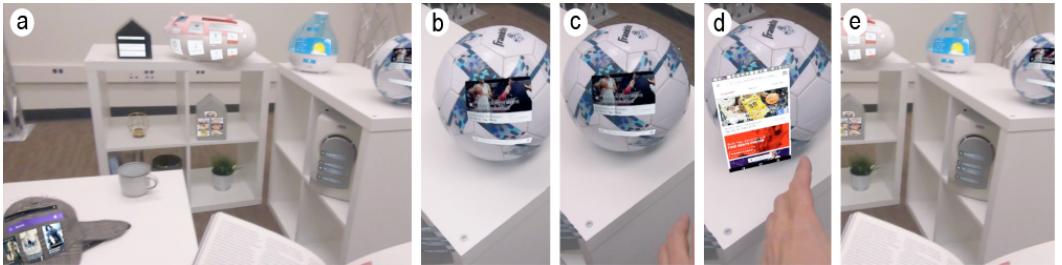


Fig. 2. (a) A user is relaxing in their living room, with BlendMR unobtrusively displaying peripheral information in the environment. (b) The user is interested in the sport event displayed on the soccer ball, and intends to purchase tickets. (c) The user approaches the soccer ball. (d) The sport event application’s full display is triggered by the user’s proximity. (e) After interactions with the application, the user sits back on their sofa and continues flipping through a magazine without distraction from virtual contents. *MR output is recorded live through the Varjo XR-3 headset.*

e.g., a weather app onto a physical air purifier as shown in Figure 1, or leverage other approaches that decide this mapping based on e.g., semantic associations [7]. Our novel approach analyzes the physical object, and finds areas that are suitable to “host” virtual elements based on aspects such as curvature and position in a room with respect to users. Using combinatorial optimization, BlendMR then maps the elements of the application to the geometry of a physical object, and displays the information through a see-through head-mounted MR display. Our approach effectively transforms 2D applications into always-on ambient information displays embedded in everyday objects. BlendMR solves the problem of how to integrate secondary virtual interfaces with physical objects for simple and unobtrusive information access, minimizing potential clutter and distraction. In future always-on MR, such integrated interfaces could live in a user’s home, work environment, etc. the same way physical objects such as a calendar can always be placed on a shelf in these environments. When users want to interact with an application, they just move their hand close to a blended interface to reveal the original representation, as shown in Figure 2. *All MR examples in this paper are recorded live through a Varjo XR-3 video see-through MR headset.*

We evaluate our approach in a user study ($N = 16$), comparing BlendMR with a standard 2D window baseline. Participants performed reading tasks in two different scenes (living room, desk) and simultaneously answered questions regarding the contents of virtual interfaces in the scenes. Results show that BlendMR statistically significantly reduced perceived clutter and distraction, and was preferred over the baseline without decreasing legibility. We believe that this highlights the applicability and usefulness of BlendMR. We demonstrate the potential of our approach in a series of application scenarios, specifically, casual interaction in a living room, social interaction, and an office setup. Finally, we provide insights and qualitative results on the impact of individual sub-objectives of the proposed optimization method, and additional results of 13 individual mappings.

BlendMR is a new type of ambient MR display that is complementary to existing representations such as 2D windows. It allows users to continuously access personalized information while reducing clutter. In our current implementation, BlendMR maps non-interactive MR interface elements onto physical objects, serving as ambient displays. Our approach would easily extend to interactive components, if those were available as decomposed UI elements. We hope to explore this direction, as well as mapping virtual affordances to physical ones, e.g., mapping a virtual button onto a physical surface that affords pressing, in the future. BlendMR showcases a potential future of MR in which virtual applications are no longer all directly transferred from traditional computing devices such as smartphones, but are tightly integrated into users’ physical environments.

The source code of our approach is available at <https://augmented-perception.org/publications/2023-blendmr>. In summary, we make the following contributions:

- A computational approach to map information from 2D applications onto physical 3D objects to create integrated ambient information displays, called BlendMR. This approach takes standard annotated 2D applications and the geometry of 3D physical objects as input, and uses geometry processing and combinatorial optimization to create ambient interfaces integrating the digital and physical.
- A comparative user study ($N = 16$) showing that BlendMR decreases clutter and distraction, and is preferred when compared to a baseline of traditional 2D MR interfaces.
- A set of results and application scenarios highlighting the versatility of the approach for future MR interfaces.

2 RELATED WORK

We leverage prior work on blending physical and virtual objects, adaptive MR interfaces, ambient and peripheral displays, and work on texture mapping, as our approach effectively maps 2D objects (e.g., images) onto 3D geometry.

2.1 Blending physical and virtual

MR holds the promise of blending the physical and the virtual, with augmented contents overlaid on top of the physical world. For efficient interactions in everyday MR, virtual and physical contents need to be aligned. Prior research leveraged planar physical surfaces to host virtual contents, such as the work by Nuernberger et al. [44] or Ens et al. [15]. Geometric alignment increases the integration level, and is useful for view management and general content adaptation. With HeatSpace, Fender et al. [16] use similar approaches to adapt the placement of physical displays in a space to best suit users' behavior. In OptiSpace [17], this approach is extended to adapt content displayed through projection mapping. They only consider users' viewing angle, whereas a core component of BlendMR is to account for the geometry of the object that hosts the virtual elements, and to decide what parts of 2D applications to display based on their importance. Gal et al. [19] took the geometry of a space into account and used heuristic optimization to adapt virtual contents to conform to a space, e.g., by adapting a virtual race track to the furniture in a room. They did not leverage physical contents as containers for 2D applications. Approaches such as SceneCtrl [63], Oasis [55], Remixed Reality [32], Substitutional Reality (Simeone et al. [54] and Suzuki et al. [56]) and TransforMR [28] overlap virtual contents onto physical objects with varying degree of immersion.

Other research in spatial augmented reality [42, 51] and tangible user interfaces [3, 13, 25] show that a spatial mapping of digital information and interactions onto physical geometries provides advantages such as haptic feedback, ubiquitous access to information, and simplified interaction. These works rely on manually mapping virtual elements to physical objects. Given that content creators cannot anticipate which physical objects are in users' space, or what virtual elements they would like to use, these approaches do not scale outside of a controlled lab environment. Recent works leverage semantic authoring to create a mapping between virtual elements and environments [50], and explored UI transition mechanisms in future everyday MR scenarios [37]. Both these approaches focus on 2D virtual elements. We complement these approaches by providing a computational approach to map virtual applications and physical objects in everyday MR environments with the goal of simple access and decreased distraction. BlendMR moves beyond 2D windows towards a tighter integration of virtual contents with physical objects.

2.2 Adaptive MR interfaces

Existing research explored approaches to adapt virtual interfaces for labeling or annotation placement problems [2, 14, 49]. Azuma and Furmanski [1] evaluated their cluster-based label placement algorithm and found that it was the best in terms of objective and subjective measures among four different algorithms. Gebhardt et al. [20] proposed a system that adapts virtual labels based on eye movement data by employing reinforcement learning. We leverage these insights for creating the connection between virtual applications and physical objects, and for controlling the visibility of applications. Our approach then creates a representation that optimally integrates elements of the virtual applications into physical objects.

Other research has taken environmental factors such as physical objects and geometry into account, such as the work by Lages et al. [29], Nuernberger et al. [44], Fender et al. [17], or Tahara et al. [57]. Ens et al. [15] consider spatial consistency, surface structure and visual saliency of surrounding environments to manage interface layouts in VR. Matulic et al. [43] proposed a method that automatically detects suitable physical surface areas for projecting digital information. SemanticAdapt [7] adapts virtual interface layouts by considering the semantic relationships between physical and virtual objects. They leverage insights from computational interaction (cf. Oulasvirta et al. [45, 47]) and adaptive user interfaces (e.g., Supple [18], ADAM[48]) to generate and adapt MR layouts.

Lu et al. suggested using the users' periphery to place and access secondary information [35, 36]. Virtual elements are head-anchored 2D windows, which leads to visual clutter if too many elements are displayed. We provide a way to reduce clutter by embedding virtual elements into physical objects.

Lindlbauer et al. [31] used combinatorial optimization (cf. [46]) to control the placement and appearance of virtual applications based on users' cognitive load, their task and environment. We similarly solve an assignment problem, and provide a novel formulation that takes the geometry of physical objects into account. The importance of surrounding environments for placing virtual elements is also supported by work such as by Luo et al. [38], who considered aspects of the physical environment for placement of objects. DiVerdi et al. [12] generated different levels of details for interfaces to fit users' needs. Our work takes a similar approach by taking importance into account for integrating virtual elements into physical objects.

We extend prior work by not only taking the environment and physical objects into account for placement, but by considering their surface geometry for integrating virtual elements. To achieve this, our method leverages geometry analysis and an optimization-based method to detect suitable sub-surface ares and adapt virtual content layouts, respectively, to fit those surfaces.

2.3 Ambient and peripheral displays

Our motivation of integrating virtual interfaces with physical objects to form unobtrusive peripheral MR displays is inspired by the long-standing vision of ambient and peripheral displays. Weiser and Brown [59] proposed the notion of calm technology, which engages the user's periphery instead of their center of attention. Wisneski et al. [60] demonstrated early examples of ambient displays that unobtrusively deliver information. Lyytinen and Yoo [39] proposed a vision of ubiquitous computing where computations are embedded in natural environments. With current MR technologies, more recent works follow this line of vision and propose ideas of pervasive augmented reality, with digital information being constantly available based on the user's physical environment [22], and glanceable augmented reality, where digital information is always available in the users' periphery [34]. ParaGlassMenu [5], as another example, is an unobtrusive MR display around a conversational partner's face to provide rich information in a social setting. Instead of modifying

the placement of virtual elements to achieve this unobtrusive integration, we propose a novel integration mechanism. We manipulate virtual elements to blend onto physical objects, which effectively serve as “hosts” for digital content.

2.4 Texture mapping

Our approach is inspired by general texture mapping methods [24, 41]. Texture mapping refers to the creation of a faithful mapping of 2D image data onto 3D objects. This typically involves parameterization and matching. There exist various approaches in the computer graphics community involving methods such as image warping [53], mesh warping [40], or methods that use neural textures [62]. We leverage standard methods for parameterization, mapping and texture packing [30]. In contrast to work in traditional texture mapping, we do not wish to achieve full coverage, but to sparsely add elements of 2D applications onto the geometry of existing physical objects. We model this as linear programming problem, which can be solved efficiently and in real time. We hope to expand our work and explore other gradient-based methods or learning-based methods for mapping 2D applications onto 3D objects in the future.

3 BLENDMR

The goal of our approach is to reduce clutter by displaying secondary MR applications as integrated ambient displays. To achieve this, we embed 2D applications into existing physical objects. In other words, we take the individual interface elements of a 2D application and map them as textures onto the surface of physical objects in MR. Our approach is designed to serve as a peripheral display showing glanceable information, and an access point to the full applications. For interaction, we envision future approaches to adapt the representation to ones that are well suited for interaction (cf. Cheng et al. [8]). Additionally, information on BlendMR interfaces avoids overlaps, follows established information hierarchies, and is aligned in groups (cf. Dayama et al. [10]). Table 1 summarizes the input parameters and variables for quick reference.

3.1 Input

We leverage two main sources of inputs: annotated 2D applications and scanned 3D meshes of existing physical objects. The mapping between applications and physical objects is created by users by moving an application close to a physical object. This approach could easily be replaced with automatic placement approaches [1, 2, 7] that leverage semantics, physical proximity, or host characteristics. BlendMR then decides *how* applications should be blended into 3D objects based on the design of the 2D applications, the geometry of the physical objects, and the users’ common position in a room (cf. [16]).

3.1.1 2D applications. We use applications from the Rico dataset [11], with examples shown in Figure 3. The dataset consists of mobile phone applications. Each application is divided into individual interface elements, denoted as $e \in E$ in our context, and their bounding boxes. We additionally extract information if pairs of elements are adjacent, denoted as $r_{e_i, e_j} \in \{0, 1\}$. We consider the distance between elements with respect to the overall size of the application, and whether they are horizontally or vertically aligned, which usually encodes grouping relationships.

We extend each element with information on its importance within the context of an application. Figure 3 illustrates an example where an annotator added importance values to the individual elements. This information is added by an annotator as values between 1 (not important) and 5 (highly important) for each interface element of an application using a custom annotation tool, shown in Figure 3. For our tests and the evaluation, one of the authors labeled 70 interfaces. On average, labeling each interface took about 30 seconds. The information is normalized to the

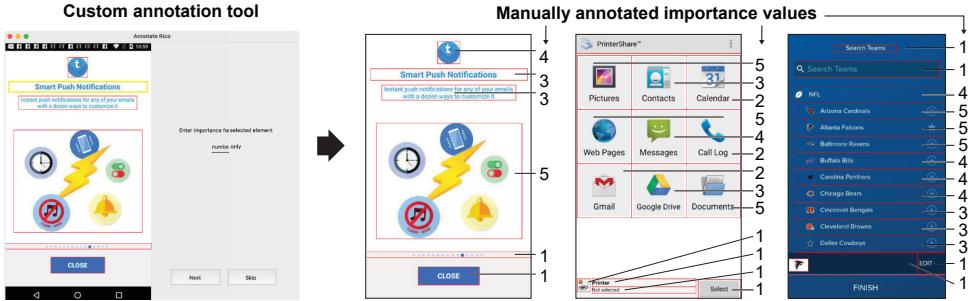


Fig. 3. The custom annotation tool (left) and example applications from Rico dataset annotated with importance scores (right). An annotator selects each element on the interface (highlighted in yellow) and gives it a score between 1 (low importance) and 5 (high importance) by pressing the respective key on the keyboard. The bounds of the individual subelements are retrieved from the RICO dataset.

maximum importance value (5), and stored as importance values $i_e \in \{0, 1\}$. We chose a manual labeling process, as we focus on finding an optimal mapping between a physical object and a 2D application. During labeling, we took various factors into account, such as a UI element's information, size, or hierarchy. For each application, the annotator used an imagined task to break ties between elements with similar importance, such as the middle and right interfaces in Figure 3. We hope to replace this manual annotation procedure with task-dependent automatic models in the future. Extending this with task-specific consideration, or to further automate the process (e.g., Wu et al. [61], Chen et al. [6]) provides an interesting direction for future work.

The size of the elements is decided before the optimization, with a pre-determined scale factor for all elements based on the size of the physical target object, and a minimum size constraint to ensure the readability of the displayed output. The pre-determined scale factor is applied to UI elements before the optimization to adjust their sizes for their physical “host”.

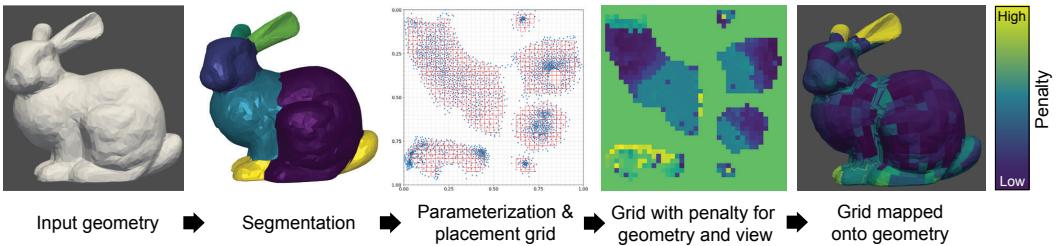


Fig. 4. Overview of the input processing for the geometry of physical objects.

3.1.2 Geometry of physical objects. To integrate virtual applications into physical objects, our system requires mesh representations of physical objects in the user’s environment (see Figure 4). We leverage 3D scans of physical objects, and store the representations as standard meshes with vertices and faces, denoted as $M = \{V, F\}$. We first segment the mesh into individual segments S to simplify parameterization. Each segment can host multiple elements of an application. We use spectral clustering [33] for segmentation, which uses a combination of geodesic and angular distances between segments.

We then parameterize each segment to retrieve the texture maps T using least squares conformal maps [30]. The maps are scaled according to their relative size on the overall mesh M and aligned to

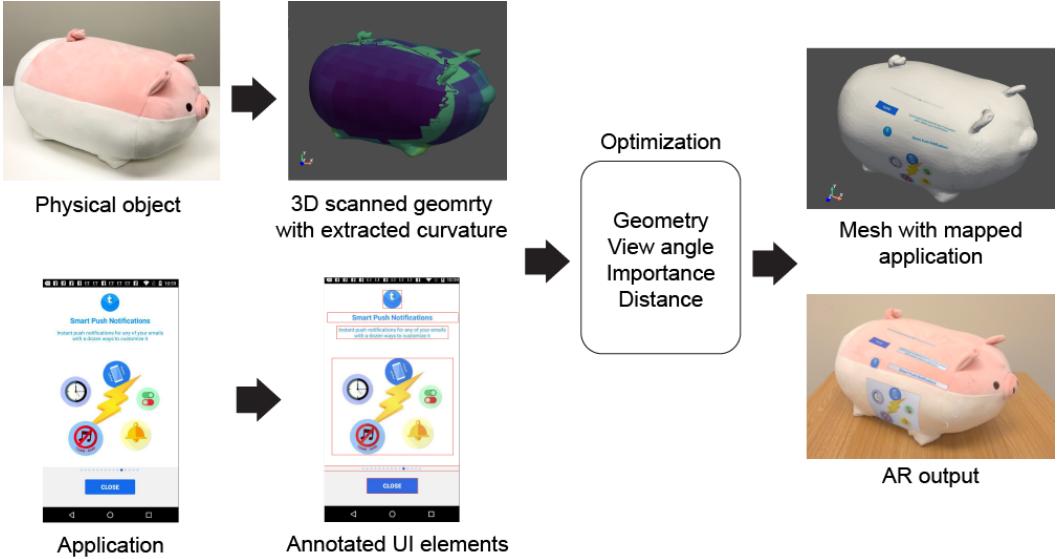


Fig. 5. Overview of the system workflow, from input (geometry of physical object and 2D application) to the output of the optimization and the MR version of the output. *AR output* is recorded live through the Varjo XR-3 headset.

the global horizontal and vertical axis by orienting them so that a randomly chosen 2D horizontal line on the parameterization aligns with a 3D horizontal line on the mesh. Finally, we create a grid of placement candidate slots for each texture, denoted as G , by overlaying a grid on top of the texture maps. We place all individual textures within a single rectangle using a rectangle-packing solver (cf. Sander et al. [52]), and overlay it with an equally spaced grid (see Figure 4, center). Grid cells are oriented correctly as the texture maps are axis aligned. Each grid cell contains multiple vertices and faces, depending on the mesh resolution.

3.2 Blending virtual objects onto physical objects

In the following, we describe how we map the individual interface elements of one application onto a physical object. A system overview is illustrated in Figure 5.

Our system analyzes the geometry of the physical objects, specifically its segments, curvature, distortion in the parameterization, and position with respect to users. This information is used to generate a set of placement candidate positions on the mesh. We then utilize integer programming to decide which elements of an application to place, and where on the mesh. To achieve this, our approach considers the object's geometry, the importance of virtual elements, and their relation and hierarchy in the 2D application.

We map each element e of a 2D application onto the geometry of a physical object, specifically the grid cells g on the texture map as retrieved from the mesh segmentation and parameterization. Each element can cover multiple grid cells. This decision is formulated as common in assignment problems as

$$x_{e,g} = \begin{cases} 1 & (\text{if } e \text{ is assigned to } g) \\ 0 & (\text{otherwise}) \end{cases} \quad (1)$$

We formulate the overall objective as a weighted sum of multiple individual sub-objectives. The sub-objectives are informed by general usability guidelines, and prior work on adaptive user

interfaces [7, 10, 31, 48]. Specifically, we prioritize the presentation of important objects (F_{imp}), ensure that the hierarchies of the 2D applications remain intact, and that related elements are grouped together (F_{dis}). Additionally, we ensure that the areas on the physical objects are well suited for "hosting" elements, i. e., they are well visible, have low curvature, and do not contain features such as wrinkles or tightly packed triangles which could lead to distortion (F_{geo}). Lastly, we optimize for users' primary viewing direction (F_{view}). We formulate the main objective with the considerations as

$$\underset{e,g}{\operatorname{argmin}} w_{\text{geo}} \cdot F_{\text{geo}} + w_{\text{view}} \cdot F_{\text{view}} + w_{\text{imp}} \cdot F_{\text{imp}} + w_{\text{dis}} \cdot F_{\text{dis}}. \quad (2)$$

In the following, we detail the individual sub-objectives. In addition to those, we leverage the overlap-avoidance from Dayman et al. [10].

3.2.1 Geometry. Each grid slot g hosts multiple vertices and faces. For each grid slot, we compute the average Gaussian curvature $K(g)$ of the containing vertices. Additionally, we count the number of vertices that fall within a cell $\gamma(g)$, normalized for the minimum and maximum number of vertices in all cells per segment. Our approach penalizes high curvature and high number of vertices in a cell compared to other cells in the same mesh, as this leads to visual distortion, formulated as

$$F_{\text{geo}} = \sum_e \sum_g x_{e,g} \cdot \frac{1}{2}(K(g) + \gamma(g)). \quad (3)$$

3.2.2 Viewing direction. BlendMR penalizes if grid cells are not facing users, as this decreases the visibility of visual elements. The viewing direction is *only* updated when a user triggers a new integration of an interface-object pair, as we envision BlendMR interfaces to be ambient information displays similar to current decorative objects sitting on a shelf, which are repositioned infrequently. We calculate the dot product between all normals in a cell (\mathbf{n}_g), the vector from users to the cell ($\mathbf{d}_{u,g}$), and average the results, denoted as $\alpha(\mathbf{n}_g, \mathbf{d}_{u,g})$. The sub-objective is formulated as

$$F_{\text{view}} = \sum_e \sum_g x_{e,g} \cdot \alpha(\mathbf{n}_g, \mathbf{d}_{u,g}). \quad (4)$$

3.2.3 Importance. This sub-objective penalizes if important elements are not displayed, formulated as

$$F_{\text{imp}} = \sum_e \sum_g (1.0 - x_{e,g} i_e). \quad (5)$$

3.2.4 Distance. We aim to keep elements that are adjacent in the 2D application close together when placed on the geometry of the physical objects. This is reflected in the distance sub-objective as

$$F_{\text{dis}} = \sum_{e_i} \sum_{e_j} \sum_g x_{e,g} \cdot r_{e_i,e_j} \cdot d_{e_i,e_j}. \quad (6)$$

r_{e_i,e_j} specifies if two elements are directly adjacent in the 2D application. We then minimize the distance for these elements. We chose this approach to minimize computational complexity by avoiding a full-factorial quadratic formulation. For elements that are directly adjacent, we minimize their distance d_{e_i,e_j} in the grid, calculated as follows. Let c_i, c_j be centers of elements e_i and e_j . We then calculate the angle θ between the vector connecting the centers of the pair and the x -axis. In addition to encouraging pairs in close proximity, we further penalize pairs that are not along any

of the four cardinal directions (i.e., N, S, E, W), illustrated in Figure 6, as such pairs are mostly less related to each other on a given interface. We calculate the distance as

$$d_{e_i, e_j} = \|c_i - c_j\|_2 \cdot (2 - \cos(4\theta)).$$

We chose this formulation as this corresponds to typical 2D layouts with a certain regularity.

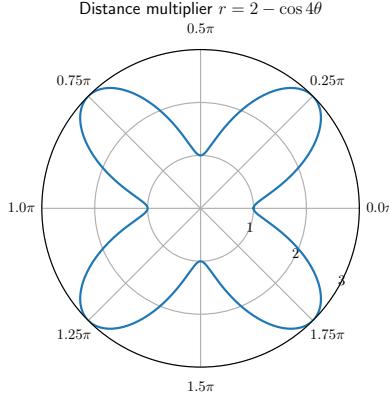


Fig. 6. Illustration of the distance multiplier in a polar plot.

3.2.5 Constraints. While the above objective produces potentially meaningful results, we introduce additional constraints to avoid trivial solutions and preserve information from the original 2D application.

As a first constraint, similar to prior work on adaptive interfaces [7, 31, 48], we ensure to avoid duplicating virtual elements in the layout. Additionally, we enforce a certain level of layout consistency. If an element is above another one, for example, it should not be placed below it when placed onto the physical object, but remain on top, or to its left or right. The constraints enforce similar consistency for the other three directions. To achieve this, we integrate a prior formulation introduced by Dayama et al. [10]. We introduce the same binary variables indicating the relative position between two elements e_i and e_j , specifically if e_j is above e_i (Γ_{e_i, e_j}) or left of e_i (Π_{e_i, e_j}). We refer readers to Dayama et al. [10] for details on how those two variables are constructed. We then construct four constraints, one to constrain the vertical placement

$$\Gamma_{e_i, e_j} = 0 \quad (\text{if } e_i \text{ is above of } e_j) \quad \forall e_i, e_j \quad (7)$$

and one to constrain the horizontal placement

$$\Pi_{e_i, e_j} = 0 \quad (\text{if } e_i \text{ is left of } e_j) \quad \forall e_i, e_j. \quad (8)$$

Note that in this formulation, for any element pair e_i, e_j , we prohibit that they are above each other if e_j is below e_i . Finally, we use the approach of Dayama et al. [10] to avoid overlapping elements.

3.3 Implementation

The optimization is implemented in Python 3.8 using Gurobi 9.5 [23] for solving the constraint optimization problem. We use libIGL[26] for mesh processing, including generation of textures and calculation of Gaussian curvature. The processing takes between 30 and 120 seconds on a Macbook Air M1 2020. The geometry of the physical objects used in this paper was scanned using an EinScan Pro 2X Plus handheld industrial scanner. For displaying objects in AR, we use the Varjo XR-3 headset connected to a commodity gaming computer (Alienware R12). The device is a high-resolution video see-through display that enables overlaying virtual contents on physical

objects without having to consider radiometric interference. The 2D applications are displayed as textures onto manually aligned virtual representations of the physical objects. In the future, we hope to extend our approach with automatic digitization and tracking of 3D objects in an environment. We chose static 3D scans to evaluate the feasibility of our approach.

3.3.1 Computational considerations. Our approach contains quadratic terms for the relation between elements, negatively impacting the run-time. To address this challenge, we only take distance terms into account for elements that are directly adjacent. Additionally, we limit the number of possible grid slots to typically between 100 and 200, depending on the size of the mesh. Lastly, we decide on the size of the elements based on their importance before the optimization, to avoid this quadratic term. In practice, this decreases the computational complexity considerably, allowing for a fast convergence time on a commodity laptop. We believe this is sufficient, as we do not anticipate our approach to run at interactive rates. Mappings between 2D applications and physical interfaces should update infrequently to enable users to leverage their spatial memory and familiarize themselves with the integrated representation.

3.4 Influence of individual optimization parameters

We explored the influence of the main sub-objectives for geometry F_{geo} and distance F_{dis} on performance and quality of the integration. The different weights are shown in Table 2. We chose to always include the sub-objective for importance F_{imp} , as its removal would lead to a trivial (empty) solution. Additionally, we do not vary the sub-objective for viewing angle F_{view} , as this would lead

Table 1. Input Parameters and Decision Variables

Parameter	Description
Geometry parameters	
$M = \{V, F\}$	Geometry of physical objects with vertices V and faces F
$S = \{s_1, s_2, \dots, s_n\}$	Segments of mesh M
$N_s \in \mathbb{Z}^+$	Number of segments
$T = \{t_1, t_2, \dots, t_n\}$	UV maps for segments with width w and height h
$G = \{g_1, g_2, \dots, g_n\}$	Set of grids containing slots for segments
$p_g^s \in \mathbb{R}^3$	Position of grid cell on mesh segment s
$p_g^t \in \mathbb{R}^3$	Position of grid cell in UV map t
$K(g)$	Average Gaussian curvature of vertices in grid cell g
Element parameters	
$E = (e_1, e_2, \dots, e_n)$	Set of virtual elements, subparts of application
$N_e \in \mathbb{Z}^+$	Number of elements.
$\mathbf{p}_e \in \mathbb{R}^2$	Position of element e in 2D application
$\mathbf{s}_e \in \mathbb{R}^2$	Original size of element e in 2D application
$\mathbf{s}_e^t \in \mathbb{R}^2$	Size of element e in resulting texture map
$i^e \in \mathbb{R}^3$	Importance of element e
$r_{e_i, e_j} \in [0, 1]$	Relation between element e_i and element e_j
$d_{e_i, e_j} \in \mathbb{Z}$	L1 distance between element e_i and element e_j on grid
General parameters	
u	User
$\mathbf{p}_u \in \mathbb{R}^3$	Position of user u
Decision variables	
$x_{e,g} \in \{0, 1\}$	Whether virtual element e is assigned to grid cell g

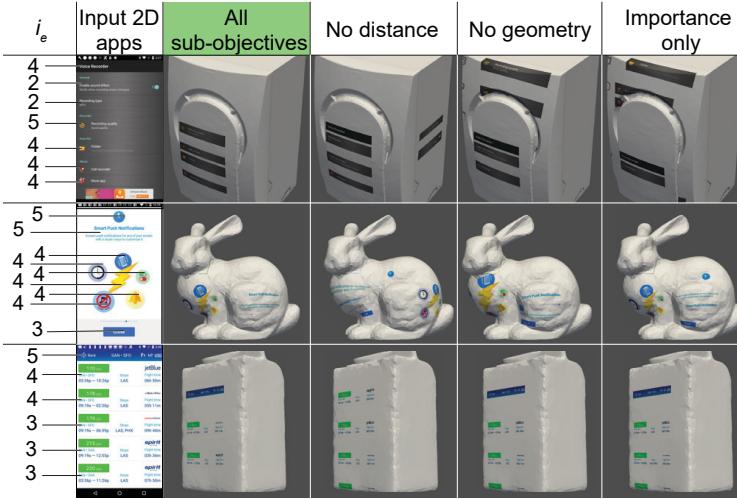


Fig. 7. Results of varying the different sub-objectives of the optimization.

to random placement of elements. We compute the results with these weights for three different 2D applications, each mapped onto the geometry of a different physical object, and add results with all sub-objectives enabled. All experiments are performed on a MacBook Air M1 2020 (8 cores, 8GB Ram), and take between 0.5 and 20 seconds to compute, depending on the mesh resolution and optimization parameters, with higher speed for fewer parameters.

Table 2. Weights of the sub-objectives used parameter exploration.

ID	w_{dis}	w_{geo}	w_{imp}	w_{view}
All sub-objectives	0.25	0.25	0.25	0.25
No distance	0.0	0.33	0.33	0.33
No geometry	0.33	0.0	0.33	0.33
Importance only	0	0.0	0.5	0.5

Figure 7 shows the results of this exploration. Without the distance term, elements are generally further apart, as expected, breaking the hierarchy of the 2D application. Without the geometry term, elements are placed in areas with high distortion (e.g., front part of bunny; bottom edge for box). Using only the sub-objective for importance exacerbates those effects, as particularly apparent for the speaker. We believe that these results show that all constraints are valuable components of the formulation. Fine-tuning weights is notoriously challenging, as the design space is vast and hard to explore due to the non-continuous nature of combinatorial optimization. We hope to systematically explore this aspect, as well as task-dependent weights, in the future.

4 EVALUATION

We performed a comparative evaluation ($N = 16$) to gather insights into the efficacy of BlendMR. Specifically, we were interested in users' ability to find and read contents on BlendMR interfaces, and BlendMR's ability to reduce perceived clutter and distraction. Participants were asked to read a newspaper article on paper. At the same time, they were asked questions by an experimenter, which could be answered using virtual interfaces displayed through a video see-through MR headset. As representation of virtual contents, we compared BlendMR with a baseline approach where

applications were displayed as 2D windows anchored to the same objects, which reflects current standard MR interfaces, and allows fair comparison. All blended interfaces were direct outputs of our approach, without manual adjustment. We measure the time to complete the question-answering task, error rate, and subjective ratings for preference, ease of use, clutter, and distraction. Results show that BlendMR significantly reduces visual clutter and distractions induced by virtual user interfaces, without increasing search time or error rates compared to the baseline. Participants preferred BlendMR, largely owing to the integration of the virtual and physical, and the opportunity for functionality mapping.

4.1 Participants & apparatus

We recruited 16 paid participants (age $M = 26.9$ years, $SD = 2.7$; 9 female, 8 male) from a local university. Participants had an average experience of $M = 2.7$ ($SD = 0.77$) in using AR interfaces and $M = 2.8$ ($SD = 0.8$) for using VR interfaces, on a scale from 1 (none) to 5 (expert). All participants had corrected or corrected-to-normal vision based on self-reports. The study was approved by the local institutional review board. The study took place in a quiet experimental room, shown in Figure 8. The room was divided to simulate a living room and a work station scenario. The software was implemented using Unity 2019. The baseline interfaces were applications from the Rico dataset. The blended interfaces were the same application processed using BlendMR, but with slightly modified contents (e.g., different temperature on the weather app) across conditions to avoid answering based on memory. Participants viewed the virtual content using the Varjo XR-3 video see-through headset, as shown in Figure 9.



Fig. 8. Participant performing task in the two study scenarios.

4.2 Study design

We use a within-subject design with two independent variables, *display mode* with two levels (baseline, BlendMR) and *environment* with two levels (living room, work station), yielding four conditions total. In each condition, participants were asked to perform a primary reading task, while occasionally answering questions regarding peripheral virtual displays, described below.

4.2.1 Tasks. We use a dual-task paradigm as employed by previous research in adaptive MR interfaces [7, 9, 31] to simulate scenarios where users are engaged in their environments, and the virtual user interfaces function as peripheral information displays.

Primary task. As primary reading tasks, participants were given short texts printed on paper, and had to answer comprehension questions afterward. We did not measure any performance metrics for primary tasks, as they were used only as stand-ins to create more realistic scenarios of virtual user interfaces as peripheral displays in everyday living environments.

Secondary task and applications. As secondary tasks, the experimenter asked participants a series of questions that could be answered using the virtual interfaces that were shown using the baseline or BlendMR. For each condition, participants were asked seven questions, covering all applications. Questions were asked during random intervals between 30 and 60 seconds. Example questions include: "What is the current temperature?" and "What is a recent sporting event you can attend?".

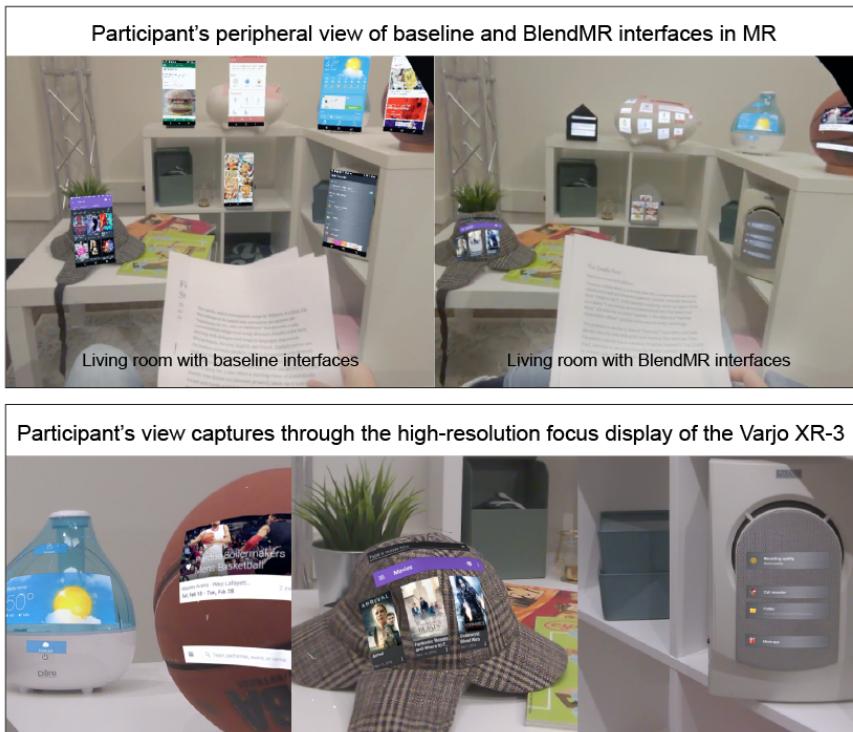


Fig. 9. Participant's MR view during the study. Left shows the 2D baseline and the BlendMR condition in the living room. Right show the interfaces displayed in the MR headset's high-resolution focus display, enabling BlendMR to display interface elements that are legible.

We selected two sets of applications for the two environments, including seven applications for the living room (weather, sports, microphone, movies, health, newspaper, recipe), and five applications for the work station (printshare, pizza deliver, slideshare, plane tickets, market place). The applications and mapping for the living room, as well as close-up views from the high-resolution focus display of the MR headset are shown in Figure 9. Mappings between applications and objects were manually chosen based on semantic connection between the application and the objects, or because the objects were well-suited as virtual containers. Participants saw the same set of interfaces across the two display modes within each environment. To avoid that participants memorize contents, we changed the contents of the applications, while maintaining their size and appearance. For example, the weather application displayed 50 degrees in the BlendMR condition,

but displayed 47 degrees in the baseline condition. For a fair comparison, we kept the sizes of interface elements, positions of interfaces, and interfaces' distances and directions to the viewer constant across conditions.

4.2.2 Procedure. Participants were briefly introduced to the experiment setup. They then completed demographic and motion-sickness questionnaires, and signed a consent form. Upon the start of each condition, the experimenter briefly walked participants through the applications. This helped them to familiarize with application placements, but avoided content memorization. Participants completed the tasks for both display modes in one environment, and then for the other environment. The order of the environment and display mode within an environment were counterbalanced using a Latin square. We did not fully counterbalance the presentation sequence of conditions across two environments, to avoid requiring participants to switch between environment setups multiple times. Each condition ended with a post-condition questionnaire. The study ended with a post-study interview and took approximately 45 minutes per participant.

4.2.3 Measurements. Secondary task competition time, i. e., time to verbally answer a question, was recorded by the experimenter using a stopwatch. Time was measured from when the experimenter completed asking the question to when participants answered. Errors in the answers were manually recorded by the experimenter. We employed a post-condition questionnaire based on the System Usability Scale (SUS) [4] (questions on confidence, ease of use, and whether participants would like to use the tested system). We added targeted questions on information access ("Information on virtual interfaces were accessible when I needed to retrieve them."), visual clutter ("The virtual user interfaces induced visual clutter in my environment.") and visual distraction ("The virtual user interfaces induced visual distraction to my reading task."). Participants also participated in a semi-structured post-study interview, expanding on their reasoning and subjective experiences.

4.3 Results

Results indicate that BlendMR significantly decreases visual clutter and distraction, while maintaining perceived interactivity and legibility of the interfaces. For both baseline and BlendMR interfaces, participants answered all questions regarding information on the interfaces correctly. Results did not show any statistically significant difference for the time it took participants to answer the questions between the baseline and BlendMR. Participants were not timed on their primary reading task to mimic a casual interaction scenario, but were asked questions regarding how distracting the virtual interfaces were to their reading task in the post-condition questionnaires. Participants also provided further insights on the BlendMR system and how they would like to incorporate it in their lives with future everyday MR. Statistical analysis was performed using JASP 0.16.3 [27].

4.3.1 Time and error. All participants answered all questions across 4 conditions correctly. We performed a 2 (display mode) \times 2 (environment) repeated measures ANOVA to gather insights into secondary task completion time. Normality and homogeneity of variance tests were conducted before using ANOVA, and neither was violated. Results did not yield a statistically significant difference for display mode ($p = .425$) between BlendMR ($M = 6.09$ sec, $SD = 4.26$) and baseline ($M = 6.39$ sec, $SD = 3.89$). While there are inevitable imprecisions in the recorded times due to the manual process, we do not believe that this changed the results in a way that would favor either condition. Results yielded a main effect for environment ($F_{1,444} = 43.10$, $p < .001$), with participants answering questions faster for work scenario ($M = 5.03$ sec, $SD = 3.23$) than the living room ($M = 7.44$ sec, $SD = 4.47$). We believe that this is due to the different object arrangements between rooms, highlighting the variability between the two setups. The results indicate that differences between BlendMR and the baseline in secondary task competition times are likely due

to noise rather than systematic challenges in the legibility or overall presentation. This is reflected in participants' comments such as "*Important information are filtered through... and it doesn't cause problems to me comprehending the interfaces*" (P1); "*I think both of them (2D display system and BlendMR) are quite good (in legibility), everything was okay to read.*" (P12).

4.3.2 Questionnaire data. We analyzed the questionnaire data using a series of Wilcoxon signed-rank tests (Bonferroni adjusted $\alpha = 0.0083$). Results of the differences between display modes are illustrated in Figure 10. Results indicate that BlendMR reduced perceived clutter ($Z = 3.861, p < .001$) and perceived distraction ($Z = 3.977, p < .001$) compared to the baseline in a statistically significant manner. Baseline ratings for information access were higher than for BlendMR ($Z = 2.727, p = .005$). Qualitative results indicate that this is due to participants being familiar with 2D information displays. Results did not yield a statistically significant difference between BlendMR and the baseline for confidence ($p = 0.139$) or ease of use ($p = 0.623$). Participants rated BlendMR higher than the baseline for the question whether they would like to use the system frequently ($Z = -2.643, p = .007$). Results did not show any differences between the different environments.

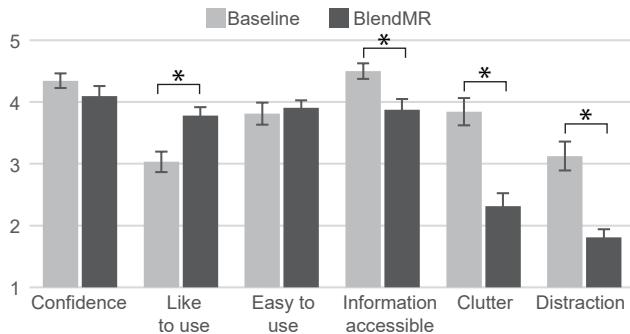


Fig. 10. Mean subjective ratings from 1 (strongly disagree) to 5 (strongly agree). Error bars indicate standard error. Lower ratings for clutter and distraction are better.

4.4 Discussion

BlendMR successfully reduced visual clutter and distraction caused by the presence of virtual interfaces, and received higher ratings. This is reflected by participants' comments during interviews: "*It was definitely less distracting and more enjoyable*" (P3); "*Blended interfaces induce less visual cluttering, and also less distractions. 2D interfaces are just like smartphone screens, so they are like smartphone screens floating in the air... so I think it is pretty annoying*" (P14); or "*I would like to use it (BlendMR) to get information I need, and it doesn't add any distraction, because objects are already there, to me, I feel like I didn't add anything redundant, but I can do more, and retrieve more information*" (P2). Both display modes were new to participants, and we aimed not to bias them towards BlendMR in our instructions. Participants' comments on the reduced distraction showed that BlendMR created a more integrated MR environment. We thus believe that peripheral digital information was unobtrusively embedded, similar to decorative physical objects, without hindering users' other daily tasks in the environment.

Integration. Participants preferred integrating digital information with the 3D world, as reflected by P12 in their comment "*Because I am 3D, why put everything in 2D in the 3D world*". 2D displays induce a sense of dis-belonging and even intrusiveness to their environments, e. g., as stated by P7: "*Mapping (information onto objects) is better. This (2D display system) is very odd, they (2D interfaces) don't look like they belong*"; or P4: "*2D displays stand out too much... I am used to a 3D world. 2D*

(interfaces) looks like a bunch of LED screens in Times Square". While all applications could simply be shown as 2D interfaces on physical planes such as tables or walls, our findings indicate an inherent value of an integrated appearance. Though participants were naturally more used to 2D information displays in their present daily lives, they appreciated that BlendMR used the 3D surfaces of existing daily objects to become "hosts" for digital information. We believe this indicates that users perceive a tighter integration of virtual content into users' physical environment positively.

Information access. Results show reduced questionnaire ratings for information access for BlendMR. This might be because participants are more familiar with 2D interfaces, as commented by six participants, e.g., "2D interfaces ... are similar to smartphone interfaces ... so they are more intuitive to read" (P14). Nevertheless, eleven participants positively commented that BlendMR's enabled efficient information access by filtering unimportant information: "It (BlendMR) displays information in a way that highlights the most important things. In a 2D interface, (there are) some things I don't need to see" (P13). We believe that this highlights that even though participants are used to conventional 2D layouts, they quickly saw the benefits of our novel representation. We believe that participants generally appreciated that BlendMR only displayed important information, but might have been skeptical that *all* important information was shown. This is reflected in the discrepancy between subjective ratings and qualitative feedback. One potential explanation is that balancing information access with the quality of presentation is challenging. In our study, one physical element was a cylindrical object, for example, which exhibited good values for the geometric sub-objective throughout due to its constant curvature. This led to some virtual information being presented at the edges of what was visible to participants, and required them to move slightly to see all contents. Even though all important information was visible, they thought that they might miss information. In the future, we plan to improve on such cases with further weight tuning or interaction to indicate which information is visible, and which is not visible

5 APPLICATIONS & ADDITIONAL RESULTS

We present three application scenarios highlighting the efficacy of our approach, shown in Figure 11 and Figure 1 (right). All applications were created using BlendMR, and displayed live in MR using the Varjo XR-3 headset. Physical objects were manually aligned, the mapping between 2D applications and the physical targets was created by a content creator.



Fig. 11. Implemented application scenarios. (a) In the meeting scenario, the red folder and gray decorative object hold virtual objects. (b) In the living room, weather applications and quick access for other apps are visible. (c) Similar pairings between 2D applications and physical objects are shown in the social interaction scenario. *All images were recorded live through the Varjo XR-3 MR headset.*

Meeting. While collaboratively working on a problem on a whiteboard (Figure 11, *left*), status messages and their calendar are integrated in objects on the shelf. While the current mapping of 2D applications to physical objects is static, we hope to expand our approach with automatic visibility control of applications users' cognitive load and task [31].

Casual interaction in a living room. A user is in their living room (Figure 11, *center*). Their environment contains peripheral elements for applications such as weather, calendar, mapped onto multiple decorative objects such as a metal box and a decorative house. The applications enable them to quickly check information such as the weather, unread messages, enable them to control a music application, see what movies were released.

Social interaction. A user is sitting in their couch at a table, with another user on the other side of the table (Figure 11, *right*). On the table, there are quick-access icons for movies on a physical hat, as well as a peripheral weather application. The interfaces are designed to blend into the environment and fade into the background. Once users require access, the interfaces are meant to change their representation to one that simplifies input and interaction.

5.1 Additional results

We further provide additional BlendMR outputs with physical objects (Figure 12), and different interfaces on the same objects (Figure 13). Our approach is able to identify suitable areas even for challenging geometry such as the desk lamp. For the red document box (2nd column), three icons that were annotated to be the most important (web pages, pictures, documents) are shown. The white lines in the MR view are artifacts of the parameterization. Our approach is able to identify suitable surfaces for all physical targets. Note that with the current weights, the placement is more conservative, meaning that we opted to present fewer but important objects. We believe the exact weights will depend on users' context and task, and hope to combine our approach with existing work on adaptive MR interfaces that automatically controls such parameters (e. g., [31]).

6 DISCUSSION

We contribute BlendMR, a novel optimization-based method that maps 2D applications onto the geometry of physical objects and takes geometric constraints, and parameters of the 2D interfaces into account. The resulting mappings are used to create unobtrusive ambient MR interfaces.

Current MR interfaces are typically presented as 2D windows, either floating in space [7, 31] or aligned to the world [44]. We believe that for a future where MR content is ubiquitous and continuously available, we need to reconsider how to present virtual applications to users. Few works consider scenarios where physical objects become "hosts" for elements of 2D applications, likely because current MR applications focus on displaying task-relevant interfaces, rather than ambient information. We believe that BlendMR opens a path to explore this novel MR representation.

Optimization objectives. Our optimization sub-objectives work together to best map interface elements onto physical objects for peripheral ambient displays in future everyday MR environments. While our current implementation only uses non-interactive components from the RICO dataset, we could easily extend this to interactive components from other MR applications. Virtual buttons, for example, could be activated by finger proximity, avoiding that users have to touch the physical objects. Alternatively, virtual functionality could be mapped onto physical objects that exhibit suitable affordances. Future works should investigate the interactivity of these new MR interfaces, as well as their tangible qualities. In considerations of tangible interactions of BlendMR interfaces, the functionality of physical objects becomes an important factor that could be included as a separate

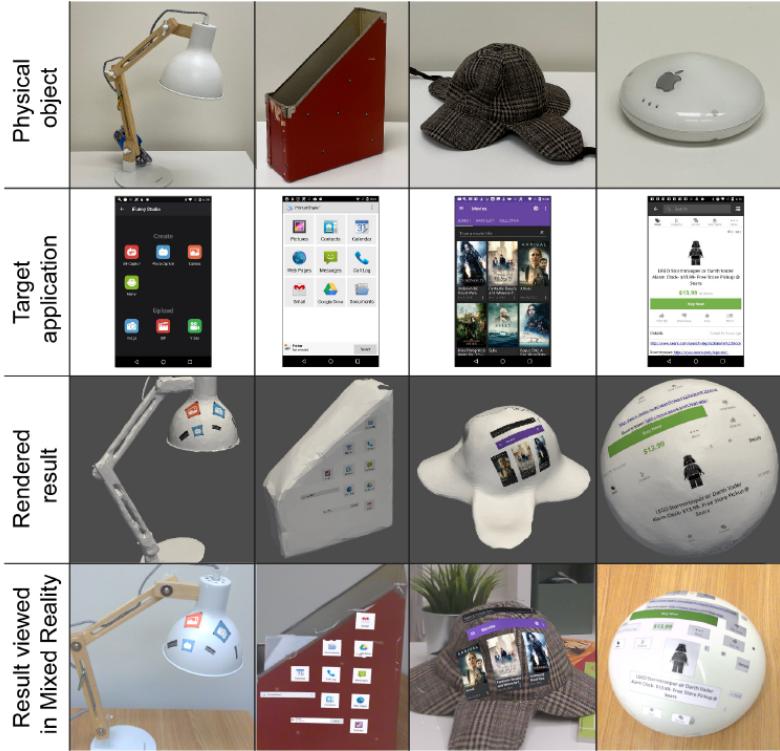


Fig. 12. Additional results with physical objects and applications. The last row is the live view from a MR headset (Varjo XR-3).

sub-objective. Similarly, the content of the UI elements should be considered. For example, a button element has high interactivity, and may be placed on an object in ways that avoid false activations when users interact with the object in its original functionality, e.g., holding a mug's handle to drink coffee. Finally, we hope to expand BlendMR with further sub-objectives regarding readability, including the color contrast between the UI elements and objects' surfaces, and investigate how object curvature impacts readability [58].

Personalization. We chose an optimization-based approach that blends 2D applications and physical objects, as we do not believe that it is feasible for content creators to anticipate what objects are in users' environments. Our current implementation assumes that physical objects are generally suitable for hosting 2D applications. We do not yet take aspects such as personal preference into account. While a decorative object might be a suitable host for 2D applications when users are working, for example, they might not want to overlay the physical object with virtual information when they are in a social context. Incorporating this type of personalization and context awareness will be interesting for future work. Further, we hope to investigate future end-to-end systems where users can indicate the importance of individual elements of their application themselves, either through manual annotations, or a mixed-initiative approach, as well as give users more control over when to trigger BlendMR. Such an end-to-end system would provide high levels of personalization for end users, but would require more manual labor. We hope to explore this tradeoff in the future. In general, however, based on the results of our approach and the comparative



Fig. 13. Additional results of different applications mapped on the same objects, as seen through MR headset (Varjo XR-3).

evaluation, we are confident that BlendMR provides a suitable and beneficial alternative to the traditional 2D window design of MR applications.

Evaluation. Our current evaluation aimed at evaluating whether BlendMR can correctly provide a representation that decreases clutter and distraction, while providing efficient access to information. Our evaluation study could be extended with other display modalities for further comparison with BlendMR interfaces. For example, 2D displays that do not occlude physical objects could be used as a comparison, as occlusion was mentioned a few times by participants to be undesired in their environments. In the future, we hope to explore cases where the importance scores are imperfect, and result in adjustments. This, however, we believe is partly independent of the visual integration. Similarly, we hope to expand our evaluation with more dynamic settings in the future, e. g., varying the number of physical objects and applications. This will require integrating multiple systems, e. g., for automatic placement and user tracking. Our current work has successfully shown that the concept and implementation are valuable for users. Integration into a more complex system is a challenging and interesting next step that will lead to a new set of challenges for BlendMR and adaptive MR systems as a whole. Further, to evaluate the generalizability of our approach, we also hope to systematically evaluate the system on a wide range of objects and interfaces (e. g., objects with variations in size; interfaces with different complexity).

Mapping virtual contents to physical objects. Our current approach only considers the shape of the physical objects, not their texture. This can lead to virtual content overlaying important information

on the physical objects. This was not an issue for the everyday objects in our tests, since many of them had uniform textures. Future versions of our approach, however, should consider the physical texture, specifically how much information it holds, as an additional parameter. Our formulation would enable a straightforward inclusion as part of the overall objective.

As our approach tightly integrates 2D user interfaces into 3D physical objects, contents can be warped to fit objects' geometries, and rearranged for better utilization of objects' limited surface areas. This might negatively impact readability and spatial memory. We plan to investigate different combinations of weights to mitigate these challenges, e. g., by increasing the weights on the geometry sub-objective, and incorporating new sub-objectives that prioritize layout consistency [10]. In our implementation and evaluation, we have encountered a limited number of failure cases with objects of high curvatures and very limited surface areas. As shown in Section 3.4, as all sub-objectives work together to produce an optimal mapping, their weight could be further fine-tuned for improved or task-dependent results. Lastly, we hope to embed our approach for dynamic representation of virtual contents into existing adaptive MR placement systems [7]. While we focus on the specific representation of virtual content, such as combination would enable us to study higher-level actions such as transitions between environments or when no suitable host object is present. We believe that our approach is beneficial for users as it utilizes physical objects to *integrate* digital content. BlendMR thus transforms MR interfaces into *always-on ambient displays* that can *unobtrusively* deliver peripheral information in future everyday MR.

View-dependent adaptation. In our current implementation, we perform a single mapping to integrate the 2D application into the geometry of the physical target. This means the mapping does not adapt to users' position in real-time. We believe that frequently changing this mapping would distract users. Our approach can create ambient information displays, similar to decorative objects. Our approach allows users to get familiar with a specific setting, and to leverage their spatial memory. This goes in line with automatic approaches for placement of physical displays such as the approach by Fender et al. [16]. Future approaches might consider a more dynamic approach, where users' position is continuously taken into account. We believe, however, that in our setting, infrequent changes are more beneficial for users since they are less distracting. Our approach is optimized for this setting and uses the geometry of the physical objects and the hierarchy of the 2D applications as main influencing factors, in contrast to prior approaches that largely focused on users' viewing direction [17].

Legibility and clarity of MR content. Lastly, the output of our approach is limited by the resolution of the MR headset. While we use a high-resolution headset (Varjo XR-3), reading text and working with high-resolution graphics is still a major challenge for all MR headsets. The participants in our evaluation did not report challenges of legibility or clarity. Our algorithm is designed agnostic of display output, and can be adjusted to its capabilities by adjusting the minimum size constraint. We hope to test our approach with headsets that feature different resolutions in the future to explore the benefits and limitations of our system.

We believe that BlendMR presents a viable alternative and complement to current approaches to displaying MR interfaces. Our computational method to map 2D interfaces to physical objects provides a path to enabling users to take advantage of this representation by removing the need for manual design.

7 CONCLUSION

We present BlendMR, an optimization-based approach to automatically integrate existing 2D applications into the geometry of physical objects for peripheral interfaces in MR. Our approach takes geometric constraints and application parameters into account to produce an optimal mapping

that balances displaying as many important elements of applications as possible without leading to information overload or legibility issues. We believe that our approach provides a viable alternative to traditional 2D windows of MR applications, with the goal to enable users to always access virtual objects and information without being overloaded. We present a technical evaluation and a series of application scenarios to demonstrate the feasibility and applicability of our approach. We believe that BlendMR is a step towards more beneficial and less obtrusive presentation of digital contents for future always-on MR scenarios.

ACKNOWLEDGMENTS

We thank Divya Kartik, Qianru Zhang, Faria Huq, Rita Dang, Ron Chew, Zhuoneng Wang for help with the first prototypes of the system, and Yiming Zhang for insightful discussions.

REFERENCES

- [1] Ronald T. Azuma and Chris Furmanski. 2003. Evaluating Label Placement for Augmented Reality View Management. In *2003 IEEE / ACM International Symposium on Mixed and Augmented Reality (ISMAR 2003), 7-10 October 2003, Tokyo, Japan*. IEEE Computer Society, 66–75. <https://doi.org/10.1109/ISMAR.2003.1240689>
- [2] Blaine Bell, Steven Feiner, and Tobias Höllerer. 2001. View management for virtual and augmented reality. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology, UIST 2001, Disney's BoardWalk Inn Resort, Walt Disney World, Orlando, Florida, USA, November 11-14, 2001*. ACM, New York, NY, 101–110. <https://doi.org/10.1145/502348.502363>
- [3] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. 2008. Tangible augmented reality. *ACM Siggraph Asia* 7, 2 (2008), 1–10. <https://doi.org/10.1145/1508044.1508051>
- [4] John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
- [5] Runze Cai, Nuwan Nanayakkaraawasam Peru Kandage Janaka, Shengdong Zhao, and Minghui Sun. 2023. ParaGlassMenu: Towards Social-Friendly Subtle Interactions in Conversations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–21. <https://doi.org/10.1145/3544548.3581065>
- [6] Jieshan Chen, Amanda Swearngin, Jason Wu, Titus Barik, Jeffrey Nichols, and Xiaoyi Zhang. 2022. Towards Complete Icon Labeling in Mobile Applications. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI '22). Association for Computing Machinery, New York, NY, USA, Article 387, 14 pages. <https://doi.org/10.1145/3491102.3502073>
- [7] Yifei Cheng, Yukang Yan, Xin Yi, Yuanchun Shi, and David Lindlbauer. 2021. SemanticAdapt: Optimization-based Adaptation of Mixed Reality Layouts Leveraging Virtual-Physical Semantic Connections. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) (UIST '21). Association for Computing Machinery, New York, NY, USA, 282–297. <https://doi.org/10.1145/3472749.3474750>
- [8] Yi Fei Cheng, Tiffany Luong, Andreas Fender, Paul Streli, and Christian Holz. 2022. ComforTable User Interfaces: Surfaces Reduce Input Error, Time, and Exertion for Tabletop and Mid-air User Interfaces. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (ISMAR '22)*. <https://doi.org/10.1109/ISMAR55827.2022.00029>
- [9] Shakiba Davari, Feiyu Lu, and Doug A. Bowman. 2022. Validating the Benefits of Glanceable and Context-Aware Augmented Reality for Everyday Information Access Tasks. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 436–444. <https://doi.org/10.1109/VR51125.2022.00063>
- [10] Niraj Ramesh Dayama, Kashyap Todi, Taru Saarelainen, and Antti Oulasvirta. 2020. GRIDS: Interactive Layout Design with Integer Programming. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376553>
- [11] Biplob Deka, Zifeng Huang, Chad Franzen, Joshua Hirschman, Daniel Afargan, Yang Li, Jeffrey Nichols, and Ranjitha Kumar. 2017. Rico: A Mobile App Dataset for Building Data-Driven Design Applications. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) (UIST '17). Association for Computing Machinery, New York, NY, USA, 845–854. <https://doi.org/10.1145/3126594.3126651>
- [12] Stephen DiVerdi, Tobias Höllerer, and Richard Schreyer. 2004. Level of Detail Interfaces. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '04)*. IEEE Computer Society, USA, 300–301. <https://doi.org/10.1109/ISMAR.2004.38>
- [13] Ruofei Du, Alex Olwal, Mathieu Le Goc, Shengzhi Wu, Danhang Tang, Yinda Zhang, Jun Zhang, David Joseph Tan, Federico Tombari, and David Kim. 2022. Opportunistic Interfaces for Augmented Reality: Transforming Everyday Objects into Tangible 6DoF Interfaces Using Ad Hoc UI (CHI EA '22). Association for Computing Machinery, New York, NY, USA, Article 183, 4 pages. <https://doi.org/10.1145/3491101.3519911>

- [14] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matt Whitlock, and Yalong Yang. 2021. Grand Challenges in Immersive Analytics. In *CHI '21: CHI Conference on Human Factors in Computing Systems, Virtual Event / Yokohama, Japan, May 8-13, 2021*. ACM, New York, NY, 459:1–459:17. <https://doi.org/10.1145/3411764.3446866>
- [15] Barrett Ens, Eyal Ofek, Neil D. B. Bruce, and Pourang Irani. 2015. Spatial Constancy of Surface-Embedded Layouts across Multiple Environments. In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction, SUI 2015, Los Angeles, CA, USA, August 8-9, 2015*. ACM, New York, NY, 65–68. <https://doi.org/10.1145/2788940.2788954>
- [16] Andreas Rene Fender, David Lindlbauer, Philipp Herholz, Marc Alexa, and Jörg Müller. 2017. HeatSpace: Automatic Placement of Displays by Empirical Analysis of User Behavior. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (Québec City, QC, Canada) (UIST '17)*. Association for Computing Machinery, New York, NY, USA, 611–621. <https://doi.org/10.1145/3126594.3126621>
- [17] Andreas Rene Fender and Jörg Müller. 2019. SpaceState: Ad-Hoc Definition and Recognition of Hierarchical Room States for Smart Environments. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces (Daejeon, Republic of Korea) (ISS '19)*. Association for Computing Machinery, New York, NY, USA, 303–314. <https://doi.org/10.1145/3343055.3359715>
- [18] Krzysztof Z Gajos, Daniel S Weld, and Jacob O Wobbrock. 2010. Automatically generating personalized user interfaces with Supple. , 910–950 pages. <https://doi.org/10.1016/j.artint.2010.05.005>
- [19] Ran Gal, Lior Shapira, Eyal Ofek, and Pushmeet Kohli. 2014. FLARE: Fast layout for augmented reality applications. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. ieeexplore.ieee.org, 207–212. <https://doi.org/10.1109/ISMAR.2014.6948429>
- [20] Christoph Gebhardt, Brian Hecox, Bas van Opheusden, Daniel Wigdor, James Hillis, Otmar Hilliges, and Hrvoje Benko. 2019. Learning Cooperative Personalized Policies from Gaze Data. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (New Orleans, LA, USA) (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 197–208. <https://doi.org/10.1145/3332165.3347933>
- [21] Google/Ipsos, U.S. 2016. How People Discover, Use, and Stay Engaged With Apps. <https://www.thinkwithgoogle.com/marketing-strategies/app-and-mobile/app-discover-process-statistics/> Retrieved Sept. 14 2022.
- [22] Jens Grubert, Tobias Langlotz, Stefanie Zollmann, and Holger Regenbrecht. 2016. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE transactions on visualization and computer graphics* 23, 6 (2016), 1706–1724. <https://doi.org/10.1109/TVCG.2016.2543720>
- [23] Gurobi Optimization, LLC. 2022. Gurobi Optimizer Reference Manual. <https://www.gurobi.com>
- [24] Paul S Heckbert. 1986. Survey of Texture Mapping. *IEEE computer graphics and applications* 6, 11 (Nov. 1986), 56–67. <https://doi.org/10.1109/MCG.1986.276672>
- [25] Steven Henderson and Steven Feiner. 2010. Opportunistic Tangible User Interfaces for Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 16, 1 (2010), 4–16. <https://doi.org/10.1109/TVCG.2009.91>
- [26] Alec Jacobson and Daniele Panozzo. 2018. libigl: A simple C++ geometry processing library. <https://libigl.github.io/>
- [27] JASP Team. 2022. JASP (Version 0.16.3)[Computer software]. <https://jasp-stats.org/>
- [28] Mohamed Kari, Tobias Grosse-Puppendahl, Luis Falconeri Coelho, Andreas Rene Fender, David Bethge, Reinhard Schütte, and Christian Holz. 2021. TransformMR: Pose-Aware Object Substitution for Composing Alternate Mixed Realities. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 69–79. <https://doi.org/10.1109/ISMAR52148.2021.00021>
- [29] Wallace Santos Lages and Doug A. Bowman. 2019. Walking with adaptive augmented reality workspaces: design and usage patterns. In *Proceedings of the 24th International Conference on Intelligent User Interfaces, IUI 2019, Marina del Ray, CA, USA, March 17-20, 2019*. ACM, New York, NY, 356–366. <https://doi.org/10.1145/3301275.3302278>
- [30] Bruno Lévy, Sylvain Petitjean, Nicolas Ray, and Jérôme Maillot. 2002. Least squares conformal maps for automatic texture atlas generation. *ACM transactions on graphics* 21, 3 (July 2002), 362–371. <https://doi.org/10.1145/566654.566590>
- [31] David Lindlbauer, Anna Maria Feit, and Otmar Hilliges. 2019. Context-Aware Online Adaptation of Mixed Reality Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology (New Orleans, LA, USA) (UIST '19)*. Association for Computing Machinery, New York, NY, USA, 147–160. <https://doi.org/10.1145/3332165.3347945>
- [32] David Lindlbauer and Andrew D Wilson. 2018. Remixed reality: manipulating space and time in augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13. <https://doi.org/10.1145/3173574.3173703>
- [33] Rong Liu and Hao Zhang. 2004. Segmentation of 3D meshes through spectral clustering. In *12th Pacific Conference on Computer Graphics and Applications, 2004. PG 2004. Proceedings*. 298–305. <https://doi.org/10.1109/PCCGA.2004.1348360>

- [34] Feiyu Lu and Doug A Bowman. 2021. Evaluating the potential of glanceable ar interfaces for authentic everyday uses. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 768–777. <https://doi.org/10.1109/VR50410.2021.00104>
- [35] Feiyu Lu, Shakiba Davari, and Doug Bowman. 2021. Exploration of Techniques for Rapid Activation of Glanceable Information in Head-Worn Augmented Reality. In *Proceedings of the 2021 ACM Symposium on Spatial User Interaction (Virtual Event, USA) (SUI '21)*. Association for Computing Machinery, New York, NY, USA, Article 14, 11 pages. <https://doi.org/10.1145/3485279.3485286>
- [36] Feiyu Lu, Shakiba Davari, Lee Lisle, Yuan Li, and Doug A. Bowman. 2020. Glanceable AR: Evaluating Information Access Methods for Head-Worn Augmented Reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 930–939. <https://doi.org/10.1109/VR46266.2020.00113>
- [37] Feiyu Lu and Yan Xu. 2022. Exploring spatial ui transition mechanisms with head-worn augmented reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16. <https://doi.org/10.1145/3491102.3517723>
- [38] Weizhou Luo, Anke Lehmann, Hjalmar Widengren, and Raimund Dachsel. 2022. Where Should We Put It? Layout and Placement Strategies of Documents in Augmented Reality for Collaborative Sensemaking. In *CHI '22: CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April 2022 - 5 May 2022*. ACM, New York, NY, 627:1–627:16. <https://doi.org/10.1145/3491102.3501946>
- [39] Kalle Lyytinen and Youngjin Yoo. 2002. Ubiquitous computing. *Commun. ACM* 45, 12 (2002), 63–96.
- [40] Yuewen Ma, Jianmin Zheng, and Jian Xie. 2015. Foldover-Free Mesh Warping for Constrained Texture Mapping. *IEEE transactions on visualization and computer graphics* 21, 3 (March 2015), 375–388. <https://doi.org/10.1109/TVCG.2014.2366101>
- [41] Jérôme Maillet, Hussein Yahia, and Anne Verroust. 1993. Interactive texture mapping. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques (Anaheim, CA) (SIGGRAPH '93)*. Association for Computing Machinery, New York, NY, USA, 27–34. <https://doi.org/10.1145/166117.166120>
- [42] Michael R. Marner, Ross T. Smith, James A. Walsh, and Bruce H. Thomas. 2014. Spatial User Interfaces for Large-Scale Projector-Based Augmented Reality. *IEEE Computer Graphics and Applications* 34, 6 (2014), 74–82. <https://doi.org/10.1109/MCG.2014.117>
- [43] Fabrice Matulic, Wolfgang Büschel, Michael Ying Yang, Stephan Ihrke, Anmol Ramraika, Carsten Rother, and Raimund Dachsel. 2016. Smart Ubiquitous Projection: Discovering Surfaces for the Projection of Adaptive Content. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016, Extended Abstracts*. ACM, New York, NY, 2592–2600. <https://doi.org/10.1145/2851581.2892545>
- [44] Benjamin Nuernberger, Eyal Ofek, Hrvoje Benko, and Andrew D. Wilson. 2016. SnapToReality: Aligning Augmented Reality to the Real World. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*. ACM, New York, NY, 1233–1244. <https://doi.org/10.1145/2858036.2858250>
- [45] Antti Oulasvirta. 2017. User Interface Design with Combinatorial Optimization. , 40–47 pages. <https://doi.org/10.1109/mc.2017.6>
- [46] Antti Oulasvirta, Niraj Ramesh Dayama, Morteza Shiripour, Maximilian John, and Andreas Karrenbauer. 2020. Combinatorial Optimization of Graphical User Interface Designs. *Proc. IEEE* 108, 3 (2020), 434–464. <https://doi.org/10.1109/JPROC.2020.2969687>
- [47] Antti Oulasvirta, Per Ola Kristensson, Xiaojun Bi, and Andrew Howes. 2018. *Computational Interaction*. Oxford University Press. <https://doi.org/10.1093/oso/9780198799603.001.0001>
- [48] Seonwook Park, Christoph Gebhardt, Roman Rädle, Anna Maria Feit, Hana Vrzakova, Niraj Ramesh Dayama, Hui-Shyong Yeo, Clemens Nylandsted Klokmose, Aaron Quigley, Antti Oulasvirta, and Otmar Hilliges. 2018. AdaM: Adapting Multi-User Interfaces for Collaborative Environments in Real-Time. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI 2018, Montreal, QC, Canada, April 21-26, 2018*. ACM, New York, NY, 184. <https://doi.org/10.1145/3173574.3173758>
- [49] Ken Pfeuffer, Yasmeen Abdrabou, Augusto Esteves, Radiah Rivu, Yomna Abdelrahman, Stefanie Meitner, Amr Saadi, and Florian Alt. 2021. ARTention: A design space for gaze-adaptive user interfaces in augmented reality. *Computers & graphics* 95 (April 2021), 1–12. <https://doi.org/10.1016/j.cag.2021.01.001>
- [50] Xun Qian, Fengming He, Xiyun Hu, Tianyi Wang, Ananya Ipsita, and Karthik Ramani. 2022. Scalar: Authoring semantically adaptive augmented reality experiences in virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18. <https://doi.org/10.1145/3491102.3517665>
- [51] Ramesh Raskar and Kok-Lim Low. 2001. Interacting with Spatially Augmented Reality (*AFRIGRAPH '01*). Association for Computing Machinery, New York, NY, USA, 101–108. <https://doi.org/10.1145/513867.513889>
- [52] Pedro V. Sander, John Snyder, Steven J. Gortler, and Hugues Hoppe. 2001. Texture mapping progressive meshes. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques (SIGGRAPH '01)*. Association for Computing Machinery, New York, NY, USA, 409–416. <https://doi.org/10.1145/383259.383307>
- [53] Hyewon Seo and Frederic Cordier. 2010. Constrained Texture Mapping using Image Warping. *Computer graphics forum: journal of the European Association for Computer Graphics* 29, 1 (March 2010), 160–174. <https://doi.org/10.1111/j.1467->

8659.2009.01584.x

- [54] 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* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 3307–3316. <https://doi.org/10.1145/2702123.2702389>
- [55] Misha Sra, Sergio Garrido-Jurado, and Pattie Maes. 2018. Oasis: Procedurally Generated Social Virtual Spaces from 3D Scanned Real Spaces. *IEEE transactions on visualization and computer graphics* 24, 12 (Dec. 2018), 3174–3187. <https://doi.org/10.1109/TVCG.2017.2762691>
- [56] Keisuke Suzuki, Sohei Wakisaka, and Naotaka Fujii. 2012. Substitutional Reality System: A Novel Experimental Platform for Experiencing Alternative Reality. *Scientific reports* 2, 1 (June 2012), 1–9. <https://doi.org/10.1038/srep00459>
- [57] Tomu Tahara, Takashi Seno, Gaku Narita, and Tomoya Ishikawa. 2020. Retargetable AR: Context-aware augmented reality in indoor scenes based on 3D scene graph. In *2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 249–255. <https://doi.org/10.1109/ISMAR-Adjunct51615.2020.00072>
- [58] Chunxue Wei, Difeng Yu, and Tilman Dingler. 2020. Reading on 3D Surfaces in Virtual Environments. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)* (2020), 721–728. <https://doi.org/10.1109/VR46266.2020.00095>
- [59] Mark Weiser and John Seely Brown. 1996. Designing calm technology. *PowerGrid Journal* 1, 1 (1996), 75–85.
- [60] Craig Wisneski, Hiroshi Ishii, Andrew Dahley, Matt Gorbet, Scott Brabe, Brygg Ullmer, and Paul Yarin. 1998. Ambient displays: Turning architectural space into an interface between people and digital information. In *Cooperative Buildings: Integrating Information, Organization, and Architecture: First International Workshop, CoBuild'98 Darmstadt, Germany, February 25–26, 1998 Proceedings* 1. Springer, 22–32. https://doi.org/10.1007/3-540-69706-3_4
- [61] Jason Wu, Xiaoyi Zhang, Jeff Nichols, and Jeffrey P Bigham. 2021. Screen Parsing: Towards Reverse Engineering of UI Models from Screenshots. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) (*UIST '21*). Association for Computing Machinery, New York, NY, USA, 470–483. <https://doi.org/10.1145/3472749.3474763>
- [62] Fanbo Xiang, Zexiang Xu, Milos Hasan, Yannick Hold-Geoffroy, Kalyan Sunkavalli, and Hao Su. 2021. Neutex: Neural texture mapping for volumetric neural rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 7119–7128. <https://doi.org/10.1109/CVPR46437.2021.00704>
- [63] Ya-Ting Yue, Yong-Liang Yang, Gang Ren, and Wenping Wang. 2017. SceneCtrl: Mixed Reality Enhancement via Efficient Scene Editing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) (*UIST '17*). Association for Computing Machinery, New York, NY, USA, 427–436. <https://doi.org/10.1145/3126594.3126601>

Received 2023-07-01; accepted 2023-09-22