InteractionAdapt: Interaction-driven Workspace Adaptation for Situated Virtual Reality Environments

Yi Fei Cheng
Department of Computer Science
ETH Zürich, Switzerland

Christoph Gebhardt
Department of Computer Science
ETH Zürich, Switzerland

Christian Holz
Department of Computer Science
ETH Zürich, Switzerland

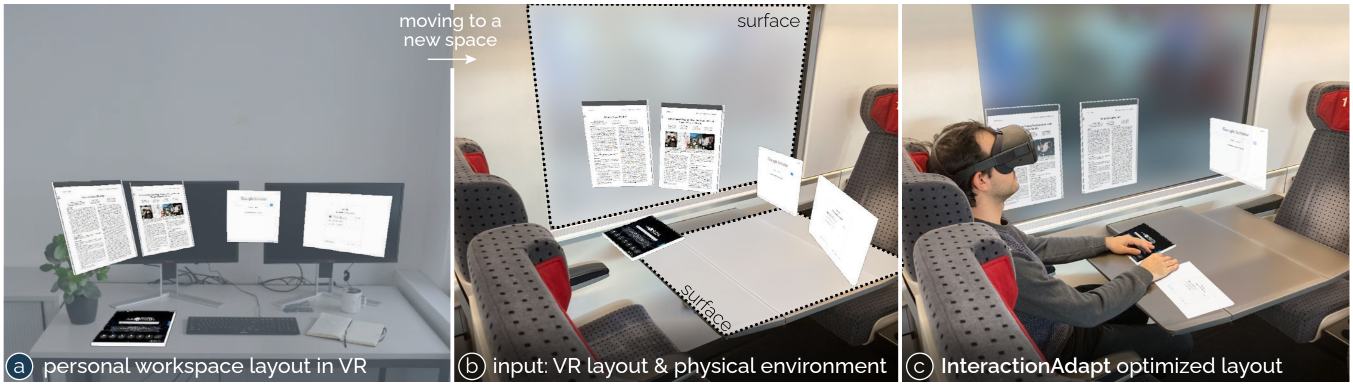


Figure 1: *InteractionAdapt* is a user interface optimization method that exploits affordances and constraints of physical spaces to adapt workspaces in Virtual Reality to situated settings for comfortable and efficient use. (a) It takes a user's personalized, physically situated, and virtual workspace as input, and (b) models the new physical environment the user is transitioning to as obstacles and affordances for UI alignment. (c) InteractionAdapt optimizes the position and orientation of virtual elements to retain the user's workspace configuration as much as possible, leveraging physical affordances for the optimized use of direct and indirect input techniques while avoiding real-world obstacles. (Real-world environment shown for illustrative purposes.)

ABSTRACT

Virtual Reality (VR) has the potential to transform how we work: it enables flexible and personalized workspaces beyond what is possible in the physical world. However, while most VR applications are designed to operate in a single empty physical space, work environments are often populated with real-world objects and increasingly diverse due to the growing amount of work in mobile scenarios. In this paper, we present *InteractionAdapt*, an optimization-based method for adapting VR workspaces for situated use in varying everyday physical environments, allowing VR users to transition between real-world settings while retaining most of their personalized VR environment for efficient interaction to ensure temporal consistency and visibility. InteractionAdapt leverages physical affordances in the real world to optimize UI elements for the respectively most suitable input technique, including on-surface touch, mid-air touch and pinch, and cursor control. Our optimization term thereby models the trade-off across these interaction techniques based on experimental findings of 3D interaction in situated physical environments. Our two evaluations of InteractionAdapt in a selection task and a travel planning task established its capability of supporting efficient interaction, during which it produced adapted layouts that participants preferred to several baselines. We further showcase the versatility of our approach through applications that cover a wide range of use cases.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; **Virtual reality**; *User interface management systems.*

KEYWORDS

Virtual Reality, Computational interaction, Adaptive user interfaces

ACM Reference Format:
Yi Fei Cheng, Christoph Gebhardt, and Christian Holz. 2023. InteractionAdapt: Interaction-driven Workspace Adaptation for Situated Virtual Reality Environments . In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23), October 29-November 1, 2023, San Francisco, CA, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3586183.3606717

UIST '23, October 29-November 1, 2023, San Francisco, CA, USA

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/3586183.3606717

1 INTRODUCTION

Virtual Reality (VR) has the potential to enhance the way we work, because it can alleviate many limitations that are present in the physical world. For instance, virtual content can be placed anywhere around information workers, enabling them to take advantage of spatial interactions otherwise unavailable in real-world environments. VR apps may also flexibly substitute the physical surroundings according to a worker's preferences. Several research prototypes (e.g., [5, 25]) and commercial applications have targeted the use of VR for such productive scenarios, such as Horizon Workrooms [57], Varjo Workspace [76], and Spatial [67].

As knowledge work becomes increasingly remote and mobile, interfaces for accomplishing productivity tasks, including VR, must evolve to operate effectively inside a wide variety of usage contexts and environments. A VR workspace designed for one physical environment may thus be sub-optimal for another—switching to a new physical environments may introduce physical constraints that hinder interaction and, worse yet, raise safety concerns.

To reconcile complex interface designs with diverse physical surroundings, prior work has adapted interface layouts to contextual factors, such as the user's state [47] or application-defined geometric constraints [22]. While prior approaches to workspace optimization are promising, they have notably neglected considering the *mode of interaction* itself, i.e., how users may interact with virtual elements inside the workspace and the physical surroundings. Crucially, this includes the extent to which *physical surroundings* may either limit or support these interactions.

Traditionally, VR users have interacted with content using a variety of techniques, such as mid-air direct touch (e.g., "poking" virtual elements [43]) and cursor-based pointing (e.g., controlling a "laser beam" [2]). However, each interaction technique that is available for use entails individual optimal interface element placements. To our knowledge, prior work has primarily focused on optimizing workspaces for visibility and direct access. This creates the opportunity to support more suitable interaction techniques that may result from the presence of physical obstacles and surfaces.

In many cases, environment features obstruct interaction, causing undesirable contact when providing input in VR. Prior work has therefore focused on interface adjustments to avoid undesirable physical collisions [9, 28] or transformed input motions in spatially-constrained settings to fit virtual environments [33, 40].

In some contexts, however, affordances in the physical environment may be *beneficial* for interaction, such as by providing passive haptic feedback during direct on-surface touch [9, 12, 55] or serving as platforms for rest (e.g., supporting elbow-anchored interactions [72]). We argue that rather than regarding physical objects as mere obstacles to interaction, adaptive interfaces could benefit from co-opting their features. To address these gaps, we introduce *InteractionAdapt*, a novel UI adaptation model that explicitly optimizes for interaction in VR.

1.1 Interaction-driven UI adaptation

InteractionAdapt determines placements for virtual elements considering *how* they may be interacted with. Our approach co-opts environment affordances for this purpose while avoiding surrounding objects that obstruct interaction. InteractionAdapt first models the physical environment as surfaces, obstacles, and mid-air placement volumes. It then evaluates the suitability of potential element placements for direct on-surface touch, mid-air touch and pinch, and remote cursor interactions, based on experimental performance insights that derived in our prior work [12]. In this process, our model factors in how physical affordances in the environment could provide haptic feedback or encourage restful postures. Our model's decisions ultimately represent optimal element placements for using an optimal interaction technique. InteractionAdapt further trades off this objective with requirements for layout spatial-temporal consistency and visibility.

To evaluate whether our method supports more ergonomic and efficient interaction, we first compare InteractionAdapt with two baseline approaches in a *target selection study*: *Surround* performed naive re-centering, and *Consistency* optimized for layout similarity while avoiding visual occlusion of elements and overlaps with physical objects. For performing selections resembling common interaction patterns in multi-window environments, participants preferred InteractionAdapt-optimized layouts and adjusted them significantly less themselves. Results also show that they perceived the layouts as a better match for both the task and environment.

Our second evaluation, a proof-of-concept study, involved a real-world knowledge work task to compare our model with *Consistency* in a *travel planning scenario*. The study examined the feasibility of our approach to support interaction performance during a task that, in addition to selections, comprised cognitive aspects and associative interaction as the target for task completion. Again, participants' feedback suggests promise in InteractionAdapt's usage of physical affordances to provide interaction benefits.

1.2 Contributions

In summary, we make the following contributions in this paper:

- An optimization-based adaptive VR layouting model that considers how affordances of the real world support different input modalities for interacting with virtual elements. Our approach fluidly transitions UI layouts between physical spaces to support situated usage by combining our consideration of interactions with factors of spatio-temporal consistency and visibility.
- Two empirical studies comparing InteractionAdapt to several baseline approaches. In a *selection task* ($N = 12$), designed to resemble multi-window interaction patterns, InteractionAdapt produced preferred and more ergonomic layouts than alternative approaches. In a second proof-of-concept study of behavior during a *travel planning task* ($N = 12$), participants' preliminary feedback illustrates the value of InteractionAdapt's consideration of interaction affordances in the physical environment.
- A series of proof-of-concept UI applications that leverage our method to allow mobile workers to use their productivity environments across diverse physical settings.

Taken together, we believe that InteractionAdapt is step towards bringing personalized VR environments *anywhere*, affording users productive and efficient use of virtual productivity tasks for prolonged periods while leveraging their physical surroundings.

2 RELATED WORK

InteractionAdapt is motivated by related work on VR interaction techniques, situated VR interfaces, and adaptive MR interfaces.

2.1 VR interaction techniques

There is significant research on VR interaction techniques [7, 27, 37, 41, 43, 58]. Within VR productivity applications [35, 57, 67, 75], the most commonly supported techniques are mid-air direct interactions, like point and pinch, and distant cursor-based control. Each interaction technique has associated benefits and limitations. Direct interactions are generally faster and more accurate because they enable users to rely on their proprioception and intuitions from interacting with everyday objects; however, they have a limited interaction range [8, 43]. Remote interaction techniques, on the other hand, have an expanded interaction range, but can be imprecise depending on the task [43]. To enable users to flexibly take advantage of both direct and distant input, several VR applications, including Immersed [35] and Spatial [67], have opted to implement controls that adapt (i.e., presenting pinch versus cursor input) depending on the context (e.g., proximity to applications, hand pose). In line with this work, we argue that providing several interaction techniques enables more fluid interaction. Thus, InteractionAdapt implements an automatic input selection feature. Moreover, our adaptation formulation explicitly optimizes element placements by considering different interaction techniques. In our evaluations, we confirm that this approach yields interface adaptations that are qualitatively preferred and more supportive of ergonomic interactions.

Besides hand-driven interaction techniques, such as those mentioned above, prior research has also explored several alternative modalities involving peripheral devices. Menzner et al. [56] and Biener et al. [6], for instance, proposed integrating 2D touchscreens into VR. Zhou et al. [78] introduced a physical mouse for 3D clicking and dragging. While we acknowledge that such modalities have their strengths (e.g., beneficial ergonomics [61]), adding peripheral devices to the system may be inconvenient. InteractionAdapt is thus focused on enabling free-hand usage of VR workspaces.

2.2 Situated VR Interfaces

VR applications have traditionally been designed in a way that isolates users from their surroundings and relies on mid-air free-hand controls. However, more recent VR research has also recognized opportunities of situating experiences in physical reality [10, 28, 66, 77]. Existing approaches situate VR experiences by either visualizing features of the physical environment (e.g., RealityCheck [28], Meta Quest's "bring your own desk" feature [3]) or co-opting passive affordances in the user's surrounding environment for virtual objects (e.g., Substitutional Reality [45, 66, 68]). InteractionAdapt takes inspiration from the latter.

The surrounding physical environment can both benefit and constrain interactions in VR. For instance, physical affordances, especially surfaces, may provide beneficial passive haptic feedback for direct interactions with virtual interfaces [12, 18, 55, 70, 71, 79, 80] (i.e., collocating the virtual interface with the physical surface extends mid-air pointing into on-surface touch). Prior research has shown that direct interaction with virtual elements, particularly when collocated with physical surfaces, improves immersion [36, 73], ergonomics, [12, 46, 55, 70, 79, 80] and task performance [12, 30, 73]. Beyond providing passive haptic feedback, physical surfaces in the environment may also serve as platforms for more restful and stable interactions [55, 70–72], while supporting users in leveraging their sense of proprioception and kinesthesia [33, 69]. In our recent study [12], we demonstrated that element placements aligned with surrounding physical surfaces enable users to adopt more ergonomic postures during interaction (e.g., elbow-rested interaction [72]). InteractionAdapt integrates the experimental findings of this study in its optimization model.

Conversely, interacting around physical objects in VR may result in unintended collisions with the environment, which may not only be distracting but detrimental to a user's safety [11]. InteractionAdapt is informed by these empirical insights. It aligns a user's virtual workspace with the real world by considering the affordances of physical objects to enable ergonomic interactions while avoiding them if they are obtrusive.

2.3 Adaptive Mixed Reality (MR) Interfaces

MR (i.e., Augmented Reality (AR) or VR) interfaces interweave UIs with the real world around us and are thus context-sensitive. They depend on numerous factors relating to the user's state (e.g., [1, 23, 47, 48]) and surrounding environment (e.g., [11, 13]). Prior research has proposed a variety of methods to enable MR interfaces to be adaptive to its usage context (cf. Grubert et al. [24]).

Our work is most directly related to prior approaches focused on environment-driven MR interface adaptations [17, 19, 20, 22, 31, 39, 60]. Ens et al. [13, 14], for instance, introduced a Markov-chain-based algorithm to align widgets with the physical environment while maintaining body-centric spatial constancy and UI visibility. Lages and Bowman's work [42] adaptively attached AR windows to walls while users walked around. Luo et al.'s study [49] suggested that environment affordances influence layouting behaviors when performing collaborative sense-making tasks in AR. ScalAR [65] used a decision-tree-based approach to re-target AR layouts to new environments while obeying semantic virtual-physical relations from prior demonstrations. SemanticAdapt [11] optimized for semantic agreement between virtual and physical objects using a linear program that additionally considers factors like temporal consistency, occlusion avoidance, and task utility [11].

Our work also builds on prior adaptive MR approaches that consider ergonomics (e.g., [15, 72]). To evaluate the ergonomics of interactions, prior work have used a mixture of bio-mechanical models and heuristics [32, 52]. Evangelista Belo et al. [16] and Montano Murillo et al. [59] both leveraged these approaches to optimize for more ergonomic mid-air virtual element placements, while Luong et al. [50] investigated which input tasks are more ergonomic and endurable for controller vs. free-hand manipulation.

The aforementioned prior work either focused exclusively on mid-air interactions (e.g., [13, 16, 59]) or treated physical objects as mere obstacles and neglects considering how environment features influence interaction (e.g., [11, 65]). Furthermore, to our understanding, no prior work explicitly modeled interaction techniques when optimizing UI layouts. InteractionAdapt's main distinguishing feature from prior UI adaptation approaches, such as Cheng et al. [11] and Ens et al. [13], is its explicit consideration of different interaction modalities. Our approach exploits physical environment affordances to optimize element placements that support more ergonomic interaction with different techniques.

3 REQUIREMENTS FOR ADAPTIVE VR

Based on prior findings, we define requirements for layout adaptation in VR to support situated use in a given physical setting. First, previous research has shown that VR productivity apps benefit from supporting multiple interaction techniques. The most effective interaction technique for a specific VR interface depends on the physical environment's constraints and opportunities. VR layout adaption methods should optimize element placements by leveraging the physical environment's affordances for interaction.

Second, prior work generally highlights that direct interaction techniques are preferable for productivity tasks due to ergonomic and accuracy considerations [79], especially when supplemented with passive haptic feedback. Beyond passive haptic feedback, prior work has shown that physical surfaces can serve as platforms for restful and stable interactions [12]. Placement of virtual elements in proximity to surfaces may hence also yield benefits.

Last, prior literature on adaptive MR interfaces suggests that inter-element visibility and temporal consistency are important usability considerations to account for [11, 13, 47]. In summary, VR workspace adaptations should include the following objectives:

R1 Prioritize interaction techniques in the order of on-surface touch interaction, elbow-rested interaction, mid-air direct interaction, and distant interaction.
R2 Avoid element placements that risk inadvertent contact with physical obstacles. Particularly in constrained spaces, encourage or even require the use of a distant interaction technique through element placement.
R3 Optimize for element visibility and minimize inter-element occlusions in layouts.
R4 Enable users to rely on their spatial memory by retaining spatio-temporal consistency in the adapted layouts.

4 INTERACTIONADAPT

We introduce InteractionAdapt, a novel UI adaptation model that reconciles interface optimization with the physical affordances surrounding the user, ergonomic constraints, and the use of preferential input techniques for prolonged interaction. InteractionAdapt enables users to interact with a personal virtual workspace in a given physical environment. When switching between physical environments, our method adapts the user's virtual interface layout depending on the physical interaction affordances of its new usage context. Figure 2 shows an overview of the method.

4.1 Layout optimization

We define the problem of re-targeting a virtual workspace as follows: given an interface layout in one environment (*input environment*), determine the placement for each interface element in a new environment (*target environment*). We propose solving for the *target* interface layout with a linear program, which has shown promise in prior work on optimizing interface designs [62].

4.1.1 Inputs. Our model takes as input *element variables*, which describe the virtual elements in use, and *environment variables*, which characterize the physical environment. Input parameters are documented in Appendix A.

Virtual elements (E):. Our model considers each element's $e \in E$ position in the input environment $p_e \in \mathbb{R}^3$, size $d_e \in \mathbb{R}^3$, and usage frequency $u_e \in [0, 1]$. u_e represents how frequently each element needs to be accessed and hence prioritized in a given context (similar to respective terms in [11, 47, 63]). For example, in a knowledge work scenario, a document that a user is actively using (i.e., high u_e) should probably be placed in a position more beneficial for interaction than a time widget referenced occasionally (i.e., low u_e). u_e values can either be defined manually, using domain knowledge, or computed from historical interaction data as the fraction of time spent actively interacting with each element.

Physical environment: Our model determines placements for virtual elements considering the physical affordances of the target environment. Following the approach of Cheng et al. [11], we define an environment as a 2 m × 1 m × 2 m mid-air placement volume positioned 0.2 m in front of the user, as well as sets of physical obstacles $o \in O$ and surfaces $s \in S$. Elements can be placed within the placement volume or attached to surfaces. More precisely, we voxelize the defined placement volume and surfaces into containers $c \in C$ to which elements can be assigned. Voxel containers are computed by subdividing the volume and surfaces based on the dimensions of the smallest input element. Each container is associated with a position and size, denoted $p_c \in \mathbb{R}^3$ and $d_c \in \mathbb{R}^3$ respectively. We then recursively merge neighboring voxels to identify additional placement arrangements (since two contiguous container can either be occupied by two individual smaller elements or a single wider element). Obstacles are physical objects that users should not interact with. Currently, surfaces and obstacles are manually defined as planes and bounding boxes respectively. They are characterized by their position $p_o, p_s \in \mathbb{R}^3$ and size $d_o, d_s \in \mathbb{R}^3$.

4.1.2 Optimization. Our approach determines the optimal assignment of virtual elements e to containers c:

$$\mathbf{x}_{e,c} = \begin{cases} 1 & \text{if } e \text{ assigned to } c \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Element positions are set to the positions of their assigned containers. They are oriented to point towards the user in mid-air containers, and aligned to be flush with surface containers.

Our model uses a temporal consistency term T (***R4***) [11, 13] and an occlusion avoidance term V (***R3***) [11, 13] drawn from prior work. In addition, we propose a novel interaction modality term I to exploit affordances of the real-world and optimize virtual element placements for more beneficial interactions (***R1***).

The resulting model maximizes the following objective function:

$$\text{argmax}_{e,c}\ (w_t \cdot T + w_v \cdot V + w_i \cdot I) \tag{2}$$

We empirically set the weights as $w_t = 0.25$, $w_v = 0.5$, and $w_i = 0.25$. We attribute a comparatively higher weight to the the occlusion avoidance term because element overlaps hinder access to interface components, which severely compromises layout usability. We see optimizing for spatio-temporal consistency and interaction modality support as equally important, and set the remaining

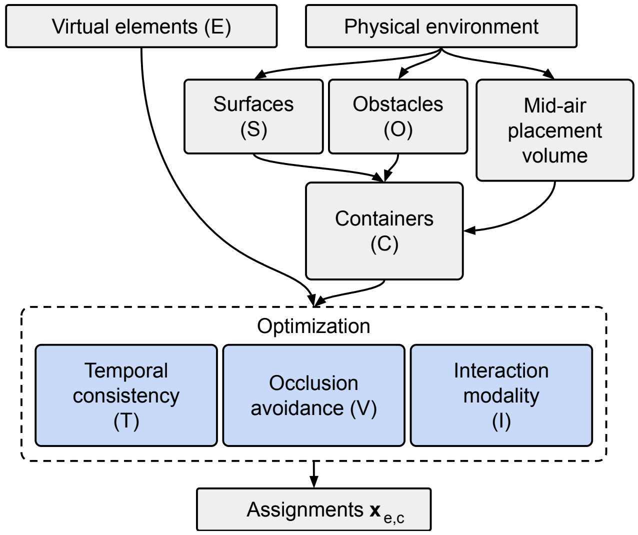


Figure 2: Our proposed model for adapting interface layouts for situated usage in virtual reality. Particularly InteractionAdapt's explicit consideration of interaction modality distinguishes it from prior work.

weights accordingly. We verified our weighting scheme through experimenting with different UIs in a variety of environments.

*4.1.3 Temporal consistency (**R3**).* To reward consistency between the *input environment* layout and the *target environment* layout, we define the sub-objective T as a summation of two components: relative position preservation T_{position} and relative order preservation T_{order}. Our T_{position} term follows the temporal consistency objective in Cheng et al. [11],

$$T_{\text{position}} = -\frac{1}{N_e}\sum_{e}\sum_{c}\|p_c - p_e\| \cdot \mathbf{x}_{\mathbf{e,c}} \tag{3}$$

N_e denotes the total number of elements. The term encodes the L2-norm between the input and potential assignment positions.

Our relative order preservation term T_{order} is inspired by Ens et al. [13]. It aims to preserve the directional spatial relationship between elements (e.g., left of, above, in front of). For the sake of clarity, we focus on describing an objective that preserves left-right relative ordering. Our approach to optimizing for up-down and front-back ordering is identical. Consider a pair of elements e_1, e_2, and their potential container assignments c_1, c_2. Relative ordering is preserved when both e_1 and c_1 are either left or right of e_2 and c_2 respectively. This behavior can be encoded as follows,

$$T_{\text{right}}^{e_1,c_1,e_2,c_2} = \begin{cases} 1 & \text{if } \left(\delta_{\text{right}}^{e_1} - \delta_{\text{right}}^{e_2}\right) \cdot \left(\delta_{\text{right}}^{c_1} - \delta_{\text{right}}^{c_2}\right) > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

where $\delta_{\text{right}}^{e_1}$, $\delta_{\text{right}}^{e_2}$, $\delta_{\text{right}}^{c_1}$, $\delta_{\text{right}}^{c_2}$ are the respective projected positions of p_{e_1}, p_{e_2}, p_{c_1}, p_{c_2} onto the user's right directional vector. $\delta_{\text{right}}^{e_1}$, $\delta_{\text{right}}^{e_2}$ represent projections in the input environment. $\delta_{\text{right}}^{c_1}$, $\delta_{\text{right}}^{c_2}$ represent projections in the output environment. Note, however, that this current formulation is quadratic. To keep our solution linear, we use the following approximation,

$$T_{\text{right}}^{e_1,c_1,e_2} = \begin{cases} 1 & \text{if } \left(\delta_{\text{right}}^{e_1} - \delta_{\text{right}}^{e_2}\right) \cdot \left(\delta_{\text{right}}^{c_1} - \delta_{\text{right}}^{e_2}\right) > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

We observed that this approach preserves relative ordering when optimized over all element pairs. T_{order} sums these values for all directions (i.e., forward, up, right), elements, and containers,

$$T_{\text{order}} = \frac{1}{3 \cdot N_e^2}\sum_{e_1,e_2,c_1}\left(T_{\text{forward}}^{e_1,e_2,c_1} + T_{\text{up}}^{e_1,e_2,c_1} + T_{\text{right}}^{e_1,e_2,c_1}\right) \cdot \mathbf{x}_{e_1,c_1} \tag{6}$$

*4.1.4 Occlusion avoidance (**R4**).* We set our occlusion avoidance term V to the negated sum of occluded assignments [11],

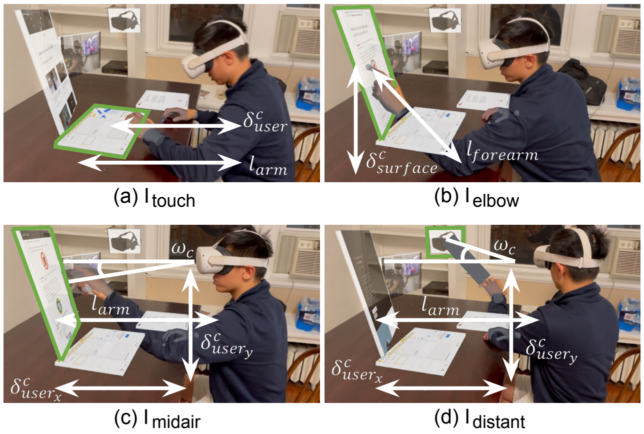
$$V = -\frac{1}{N_e}\sum_{c}\sum_{c_{\text{occluded}}} \mathbf{x}_{c,c_{\text{occluded}}} \tag{7}$$

$\mathbf{x}_{c,c_{\text{occluded}}}$ is a set of additional variables denoting whether containers c and c_{occluded}, a container c potentially occludes, are simultaneously occupied (i.e., $\mathbf{x}_{c,c_{\text{occluded}}} = 1$ if $\mathbf{x}_c = 1$ and $\mathbf{x}_{c_{\text{occluded}}} = 1$). We achieve this behavior with an indicator constraint. We compute potential occlusions with ray-casts from the user's head position.

*4.1.5 Interaction modality (**R1**).* Our interaction modality term I prioritizes usage of interactions in the order of on-surface touch, elbow-supported interaction, mid-air direct interaction, and distant interaction, each modeled respectively with the following sub-objectives (Figure 3): I_{touch}, I_{elbow}, I_{midair}, and I_{distant}.

Touch support (I_{touch}): We compute touch support as follows,

$$I_{\text{touch}}^{c} = \begin{cases} \left(1 + e^{10\cdot\left(\delta_{\text{user}}^{c} - l_{\text{arm}}\right)}\right)^{-1} & \text{if } c \text{ is a horizontal surface} \\ 0.3 \cdot \left(1 + e^{10\cdot\left(\delta_{\text{user}}^{c} - l_{\text{arm}}\right)}\right)^{-1} & \text{if } c \text{ is a vertical surface} \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$



(a) I_{touch} (b) I_{elbow} (c) I_{midair} (d) I_{distant}

Figure 3: Illustration of our interaction technique term. The green frame highlights the UI element of focus. (a) I_{touch} places elements on surfaces reachable with a bent arm position. (b) I_{elbow} places elements above surfaces with a distance smaller than the length of a forearm. (c) I_{midair} places elements reachable with a bent arm position within the user's field of view and at eye level. (d) I_{distant} places elements out of reach within the user's field of view and at eye level.

δ^c_{user} denotes the distance between container c and the user. l_{arm} is a constant representing the length of the user's upper limb. We set $l_{arm} = 0.8$ meters, which is the the average upper limb length of people [21]. As touch interaction requires a surface for haptic feedback, this term can only become non zero for containers that are surfaces. We multiply the value for vertical surface containers with a constant scalar of 0.3 to model the assumption that touch interactions with vertical surfaces are less ergonomic as users can leverage horizontal surfaces to rest. The function's inclusion of an exponential term aims to reward placements of elements on surfaces within reach for ease of access.

Elbow-rested interaction support (I_{elbow}): We compute support for elbow-rested interactions as,

$$I^c_{\text{elbow}} = \begin{cases} e^{-30\cdot(\delta^c_{\text{surface}} - l_{\text{forearm}})^2} & \text{if } c \text{ is mid-air above surface} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

$\delta^c_{\text{surface}}$ denotes the distance between a mid-air container c and the closest horizontal surface container beneath it. l_{forearm} is a constant that represents a user's elbow length (set to 0.3 adjusted from [21]). Note that $I^c_{\text{elbow}} = 0$ for containers with no horizontal surface beneath them. To afford restful elbow-anchored interactions [72], I^c_{elbow} saturates at 1 when container c is approximately a forearm's distance l_{forearm} from a surface.

Direct mid-air interaction support (I_{midair}): We define support for direct mid-air interaction as,

$$\begin{aligned} I^c_{\text{midair}} = {} & (1 + e^{10\cdot(\delta^c_{\text{user},x} - l_{\text{arm}})})^{-1} \cdot \\ & (1 + e^{10\cdot(|\delta^c_{\text{user},y}| - 0.2)})^{-1} \cdot \\ & \text{smoothstep}(0.5, 1.0, \omega_c) \end{aligned} \tag{10}$$

ω_c is computed as the angular difference between the user's forward vector and the element's position. $\delta^c_{\text{user},x}$ and $\delta^c_{\text{user},y}$ define the horizontal and vertical distance of container c from the user. Our smoothstep function follows the implementation described in Vivo and Lowe [74]. $\text{smoothstep}(0.5, 1.0, \omega_c)$ returns zero for $\omega_c \geq 60°$ and saturates at one for $\omega_c \leq 30°$. The function is designed to bias placements within reach (i.e., with the 0.8 standard arm-length [21] offset on $\delta^c_{\text{user},x}$), at approximately eye level (i.e., with the 0.2 offset on $\delta^c_{\text{user},y}$, empirically adjusted from Cheng et al. [11]), and within user's field of view (i.e., 120° for modern VR headsets [51]).

Distant mid-air interaction support (I_{distant}): We define support for distant mid-air interaction as,

$$\begin{aligned} I^c_{\text{distant}} = {} & (1 + e^{-10\cdot(\delta^c_{\text{user},x} - l_{\text{arm}})})^{-1} \cdot \\ & (1 + e^{10\cdot(|\delta^c_{\text{user},y}| - 0.2)})^{-1} \cdot \\ & \text{smoothstep}(0.5, 1.0, \omega_c) \end{aligned} \tag{11}$$

The function biases placements out of reach, approximately at eye level, and within user's field of view. Constant values within this term are selected following similar considerations as in I_{midair}.

We define our final interaction technique term as the sum of values of each interaction technique,

$$\begin{aligned} I = \frac{1}{N_e} \sum_{e,c} u_e \cdot (& w_{\text{touch}} \cdot I^c_{touch} + w_{\text{elbow}} \cdot I^c_{\text{elbow}} + \\ & w_{\text{midair}} \cdot I^c_{\text{midair}} + w_{\text{distant}} \cdot I^c_{\text{distant}}) \cdot x_{e,c} \end{aligned} \tag{12}$$

We set the weights to be $w_{\text{touch}} = 0.70$, $w_{\text{elbow}} = 0.15$, $w_{\text{midair}} = 0.10$, and $w_{\text{distant}} = 0.05$. We weighted the touch support term most highly to encode the assumption that direct on-surface touch is preferred because it enables interactions at greater speed, accuracy, and comfort [55, 79]. We set the remaining weights to model a general preference for direct pointing interactions over distant interactions (i.e., $w_{\text{midair}} > w_{\text{distant}}$), especially when the interface element is placed in positions affording restful postures (i.e., accomplished by setting $w_{\text{elbow}} = 0.15$ and $w_{\text{midair}} = 0.10$). Intuitively, this term prioritizes placement of elements with high usage frequency u_e in positions where the interaction technique score is higher.

4.1.6 Constraints. The previous terms do not sufficiently constrain the optimization. Here, we describe constraints that further limit the assignment of elements to avoid trivial solutions and produce more meaningful layouts. First, we avoid element duplicates with the following constraint,

$$\sum_c x_{e,c} \leq 1, \forall e. \tag{13}$$

Second, we prevent multiple elements from simultaneously occupying the same container. This constraint is formulated as,

$$\begin{aligned} c_1 \cap c_2 \neq \emptyset \implies \sum_e x_{e,c_1} + \sum_e x_{e,c_2} \leq 1; \\ c_1, c_2 \in C, c_1 \neq c_2 \end{aligned} \tag{14}$$

Third, we require elements to fit within their assigned containers in terms of size with the following constraint,

$$d_e > d_c \implies x_{e,c} = 0 \tag{15}$$

*Interaction Obstacle Avoidance (**R2**).* Beyond common constraints of UI assignment problems (Eq. 13-15), we model a bias that avoids placing elements near physical real-world obstacles that risk unintentional collisions during interaction. Specifically, we propose two constraints which avoid placements that either intersect with obstacles or are directly behind a physical surface,

$$\begin{aligned} c \text{ contains } o \implies x_{e,c} = 0, \forall e \in E \\ \delta^c_{\text{surface}} < \zeta \implies x_{e,c} = 0, \forall e \in E. \end{aligned} \tag{16}$$

$\delta^c_{\text{surface}}$ denotes the smallest projected distance between container c from the surface it resides behind, seen from the user's perspective. ζ is a threshold we empirically set to $0.1\,m$. We determined this threshold observing that if elements are placed too closely behind surfaces, users may still be inclined to interact with them directly (i.e., through touch).

4.2 InteractionAdapt Workspace

For our evaluation and applications, we integrated our optimization model as part of an InteractionAdapt Workspace environment. Our workspace implements the functionality of popular VR productivity apps, providing tools for opening browsers, creating whiteboards, and adding widgets for input (e.g., keyboards). Unlike in prior research prototypes (e.g., [11, 47]), all virtual elements in our workspace are fully interactive.

4.2.1 Interaction techniques. Users can interact with workspace contents using three input techniques: direct touch, pinching, and remote cursor control. In line with popular productivity VR apps (e.g., Spatial [67], Immersed [35]), our system automatically selects which technique is presented based on the user's hand and head positions, the physical environment, and the position and contents of interface elements. Note that elements can be attached to surfaces. Interacting directly with surface-attached interfaces enables users to benefit from passive haptic feedback.

4.2.2 Defining surfaces and obstacles. Our workspace provides support to define surfaces and obstacles in the real world as planes and bounding boxes in VR. Users can use the pass-through mode of Quest 2 to inspect the real world and indicate 3 or 4 constraining points of physical surfaces and obstacles using the controller. The workspace saves defined objects as one virtual scene.

If users return to a previously visited environment, the correspond virtual scene can be retrieved and physically aligned using a 3-point calibration method. We acknowledge that this process could be streamlined with automated sensing approaches in future iterations (e.g., SSTNet [44]). It is critical to note, however, that simple scene geometry extraction methods will unlikely suffice for identifying interaction affordances where important considerations of safety and physical texture must also be respected.

We provide an awareness of physical surfaces and obstacles in the user's surrounding with an adaptive grid visualization that adjusts its resolution based on its proximity to the user. Both, surfaces and obstacles fade in when the user is within a range of $1m$ (i.e., slightly beyond the user's reach). Obstacle representations are further highlighted in red as the user approaches them within $0.1m$.

4.2.3 Implementation. We implemented our system for the Meta Quest 2 using Unity 2021 and solved the linear program described above using Gurobi 9.5 [26] via their Python 3.8 interface. Communication was realized via websocket.

5 SELECTION STUDY

To evaluate our approach, we first conducted an experiment investigating whether InteractionAdapt enables ergonomic and efficient input through its adapted layouts in a selection task. We compare InteractionAdapt to two baseline adaptation methods. Our evaluation is designed following the example of prior work [12, 59], which used abstracted tasks to examine the ergonomic qualities and usability of interfaces. In our study, we aim to answer the following research questions: 1) how our approach affects task performance and interaction behavior, 2) how our approach impacts user preferences, perceived usability, as well as task and environmental complementarity in layout design, and 3) the degree to which adapted layouts resemble participants' perceived optimal layout.

5.1 Study Design

We used a within-subject design with three variables: TASK (2 levels: *selection*, *layout*), ENVIRONMENT (2 levels: *seated*, *standing*), and METHOD (3 levels: *Surround*, *Consistency*, *InteractionAdapt*). We collected metrics on participants' *task execution time*, *selection errors* (i.e., number of missed clicks), *selection time*, *selection speed* (i.e., average velocity across selection trajectory), and *touch-to-cursor ratio* (i.e., ratio between touch and cursor interactions). We also recorded the number of elements participants moved during the *layout* task.

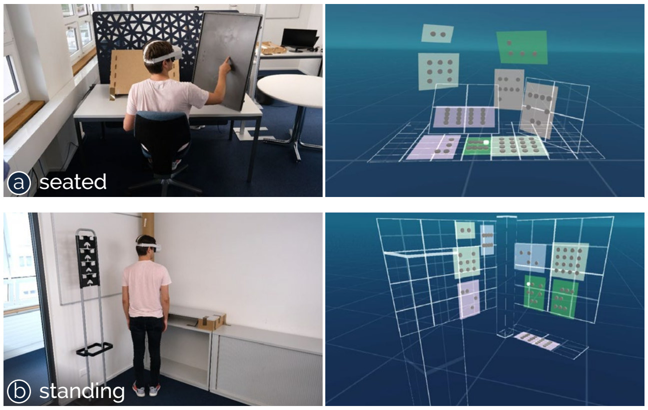


Figure 4: Physical and virtual representations of the (a) *seated* and (b) *standing* environments used in our selection study.

Lastly, we asked participants to rank our methods, rate the complementarity of adapted layouts with the environment and task, and report their task load during *selection* (i.e., averaged NASA TLX responses). ENVIRONMENT and METHOD orders were fully counterbalanced. The TASK order was fixed.

Tasks. The study comprised two tasks, which were always executed in sequence. Participants first performed a 3D multi-window *selection* task, which involved selecting highlighted buttons on spatially distributed UI elements "as *quickly* and *precisely* as possible" (Figure 4). We designed the UI to be an abstraction of windowed environments in VR workspaces and the task interactions to resemble interaction patterns of sequential and parallel usage in multi-screen ecologies [38]. Participants' attention would first be directed to one or two elements, activated in dark green in (Figure 4, right). They then selected highlighted buttons, which either appeared sequentially inside one UI element or simultaneously in two UI elements (i.e., parallel window usage). UI elements that shared the same color were always activated at the same time. Participants performed a total of 74 selections, divided into 3 parallel window usage patterns (simultaneous interaction across 2 windows) and 1 sequential pattern (one-by-one across 3 windows).

Participants subsequently performed a *layout adjustment* task, where they adjusted method-adapted layouts (i.e., move, rotate, attach or detach from surfaces) such that it would better support them in the selection task. Participants' adjusted layouts served as a "ground truth" to compare output layouts against.

Environments. We built two environments to evaluate our adaptation approach (Figure 4), including a *seated* environment that resembled a classic workstation in an office and a *standing* environment that represented a mobile setting where a user may briefly want to perform a productivity task (e.g., in a quiet corner at an airport or a conference venue). The *seated* environment consisted of two slightly tilted displays and a desk. The *standing* environment consisted of one horizontal and two vertical surfaces, as well as two obstacles. The environments were selected as potential settings of knowledge work, and for the differences in their affordances (e.g., seated versus standing, more available space in front versus constrained). We intentionally selected environments with different

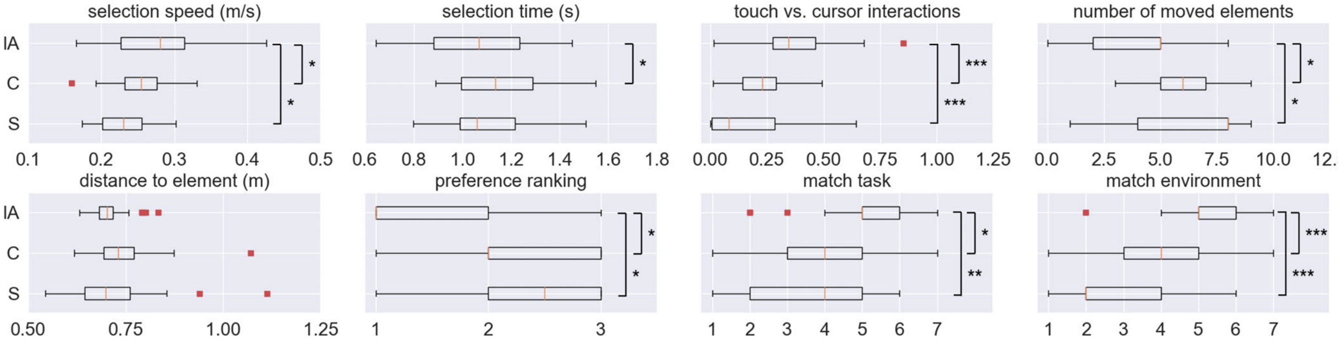


Figure 5: Boxplots for the results of our selection study. We compared the methods *InteractionAdapt* (IA), *Consistency* (C), and *Surround* (S). Significance levels: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

affordances because limited adaptations would be required otherwise. In addition, environments were selected to contain physical objects which could potentially be appropriated for interaction support. In training, participants were immersed in a third and empty environment. Here, they were seated in an office chair, faced a large open space, and performed interactions in mid-air.

Methods. We compare *Surround*, *Consistency*, and *InteractionAdapt*. *Surround* places elements at the same position relative to the user in new environments. It is comparable to how virtual environments are initialized on commercial platforms (e.g., re-centering on the Meta Quest). *Consistency* optimizes for layout similarity between environments while avoiding occlusions and placements that intersect with obstacles (i.e., objectives T and V defined in Section 4.1). *InteractionAdapt* represents our method's output.

As inputs, the usage frequency of each UI element u_e was computed as the fraction of its number of buttons with respect to all buttons. For all three environments, we manually defined the surfaces, obstacles, and mid-air placement volumes that constitute the set of containers C of our problem formulation. Participants' adjusted layouts were used as input for following adaptations.

5.2 Procedure

Prior to the study tasks, participants completed a consent form and demographic questionnaire. Participants then began by performing two practice trials in the empty environment to learn both tasks. Afterwards, participants completed the conditions of the study, performing selection and adjustment tasks in both environments for each of the three adaptation methods (completing six pairs in total). Participants completed a questionnaire after each selection task. Finally, participants ranked the three adaptation methods according to preference. They completed sessions in under an hour.

5.3 Apparatus

Study environments were implemented as InteractionAdapt workspaces. Participants interacted with the virtual environment, including adjusting element position, using the standard 3D interaction techniques described in Section 4.2.1. We manually calibrated virtual environments to the physical space (Section 4.2.2), each time verifying physical surfaces and corners indeed lined up with corresponding virtual features.

5.4 Participants

We recruited 12 participants (3 female, 9 male), ages 21–38 (M=27, SD=4.6) from a local university. Participants reported experience with VR technology, hand tracking, and video games on a 5-point Likert scale, as well as their alertness using the Stanford Sleepiness Scale [34] (0–7). Participants' median responses were VR experience = 4, hand-tracking experience = 3, gaming experience = 3, and level of alertness = 6. Participants received $20 as gratuity.

5.5 Results

To analyze dependent variables, we performed an ANOVA for normally distributed data and Friedman's test if the normality assumption was violated (tested with Shapiro-Wilk). For the latter, we treated environments as repeated measures. Post-hoc pairwise comparisons were performed using the Holm-Bonferroni test. In the following, we focus on reporting significant differences across METHODS. Results with significant differences are shown in Figure 5.

Selection Performance & Interaction Behavior. We found a main effect of METHOD on *selection speed* ($\chi^2(2) = 9.22, p = .01$). Participants moved fastest between target buttons in the *InteractionAdapt* conditions (all $p = .05$). Similarly, we found an effect of METHOD on *selection time* ($\chi^2(2) = 3.75, p = .04$). Participants had a lower selection time in *InteractionAdapt* than *Consistency* ($p = .04$). There were no differences between the two other conditions. A comparison of the ratio between touch and cursor interactions also showed significant differences ($\chi^2(2) = 19.66, p < .001$). Participants used touch more frequently in *InteractionAdapt* than *Surround* and *Consistency* (all $p < .001$). In all conditions, participants avoided unintentional collusion with physical objects, likely on account of how all physical affordances were clearly visualized. The experimenter further observed that in the *InteractionAdapt* condition, participants used surfaces as resting platforms while performing the selection task.

Layout. We found a main effect of METHOD on the number of elements participants moved in the *layout* task ($\chi^2(2) = 14.04, p < .001$). On average, participants moved fewer elements in *InteractionAdapt* than *Consistency* ($p = .009$) and *Surround* ($p = .001$). There was no significant effect of METHOD on distance of UI elements to user ($\chi^2(2) = 2.42, p = 0.30$).

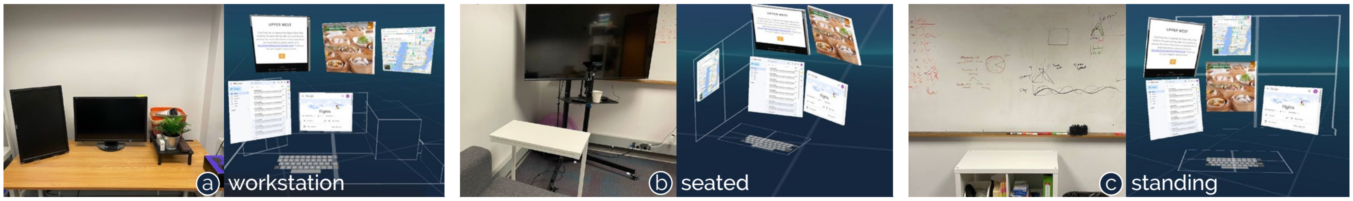


Figure 6: Physical and virtual representations of the (a) *workstation*, (b) *seated*, and (c) *standing* environments used in our proof-of-concept usage study, where participants performed a real-world knowledge work task.

Self-reports. We found a main effect of METHOD on participants' preference rankings ($\chi^2(2) = 11.31$, $p = .004$), perceived match between layout and task ($\chi^2(2) = 11.76$, $p = .003$), and match between layout and environment ($\chi^2(2) = 21.14$, $p < .001$). Participants ranked *InteractionAdapt* significantly higher than both *Consistency* and *Surround* (both $p = .01$). Participants also perceived layouts adapted by *InteractionAdapt* as more suitable for the task than both *Consistency* ($p = .02$) and *Surround* ($p = .003$). Lastly, participants perceived layouts adapted by *InteractionAdapt* as more suitable for the environment compared to the baselines (both $p < .001$). Compared to *Surround*-adapted layouts, *Consistency*-adapted layouts were also perceived to be a better match for the environment ($p = 0.007$), but not for the task ($p = 0.34$).

Performance metrics. We found no significant effect of METHOD on task load ($\chi^2(2) = 2.60$, $p = 0.27$), task execution time ($\chi^2(2) = 0.31$, $p = 0.86$), and selection errors ($\chi^2(2) = 3.52$, $p = 0.17$).

Discussion. Our results indicate that there is a preference for our approach over *Surround* and *Consistency*. Participants also reported perceiving *InteractionAdapt* adapted layouts as more suitable for both the task and environment and adjusted them least, suggesting they were closest to what participants considered optimal.

While we did not observe differences between conditions with respect to task load and performance metrics (e.g., accuracy, time of completion), the results revealed several interesting insights relating to the participants' interaction behaviors. Participants moved significantly faster when working with layouts adapted using InteractionAdapt and used direct interactions more frequently. This can not be attributed to our method simply moving the interface closer to the participant as there was no significant difference in the average distance of UI elements to participants across conditions. We argue that this phenomenon can be accredited to *InteractionAdapt*'s capability of leveraging physical affordances that enabled participants to benefit from passive haptic feedback and adopt more comfortable, environment-supported positions. In the *InteractionAdapt* condition, participants were further observed leveraging surfaces as resting platforms for their wrist, forearm, and elbow. The participants' self-reported measures and observed interaction behavior differences collectively suggest that *InteractionAdapt* may enable more ergonomic and efficient interaction.

6 PROOF-OF-CONCEPT USAGE STUDY

To further explore the benefits and limitations of InteractionAdapt, we conducted a second proof-of-concept study to examine the applicability of our method in a real-world knowledge-work task. For this, 12 participants experienced interfaces adapted with InteractionAdapt in the context of travel planning, which we designed based on prior work (e.g., [11, 53]). Contrary to our selection study, in this study virtual elements contained semantically meaningful content, which affects users' spatial memory [13, 47]. In real-world knowledge tasks, metrics such as task completion time and interaction behavior are primarily influenced by participants' problem-solving approach as opposed to the interface layout. Hence, in this study, we focused on participants' qualitative impressions. We compare the effect of the interaction term (*InteractionAdapt*) to optimizations without it (*Consistency*). We excluded *Surround* since it was least preferred in our selection study (Section 5).

Procedure. We used a mixed factorial design with METHOD (2 levels: *Consistency*, *InteractionAdapt*) as a within-subject independent variable and ENVIRONMENT (2 levels: *seated*, *standing*) as a between-subject independent variable. We designed the study with the following narrative in mind: while working, participants occasionally move from their personal workspace to new physical contexts.

Participants first designed a layout consisting of five virtual displays to perform a travel planning task at a *workstation* environment (Figure 6a). Half of the participants ($N = 6$) then moved to a *standing* (Figure 6b) environment and the other half moved to a *seated* (Figure 6c) environment. We note that these environments differ from those used in our selection study. In their respective second environments, all participants experience both *Consistency* and *InteractionAdapt* adapted layouts (defined in Section 5.1).

For the travel planning task, participants were asked to first check their email for instructions, and then, in any order, search for a flight, explore a map for a museum, and compare two restaurant menus. We designed this task to include representative interactions, such as reading, navigation, cross-referencing information across windows, and manipulation [64]. Between conditions, participants repeated the same task but for a different travel destination.

For our *InteractionAdapt* model, the usage frequency of each input element u_e was defined to depict the interaction patterns we expected to emerge from the travel task (e.g., keyboard requiring most frequent access, followed by the map, then restaurant websites). The adaptation method order was counterbalanced to account for learning effects. We concluded the study with a semi-structured interview asking participants for their impressions of the adapted interfaces.

Participants. We recruited 12 new participants (5 female, 6 male, 1 did not specify), ages 21–37 ($M = 25$, $SD = 4$) from a local university. Participants' median responses to our demographics questions (Section 5.4) were: VR experience = 3, hand-tracking experience

= 3, gaming frequency = 4, and alertness = 5. Participants received $10 as a gratuity for their time.

Results. On average, participants completed one instance of the travel planning task in 183.6 s ($SD = 64.8$). Several participants completed the task more quickly than anticipated due to their familiarity with the destination.

9 participants preferred *InteractionAdapt* over *Consistency*. The strongest point of consensus within participant observations was that *InteractionAdapt* generated beneficial layouts by aligning frequently used elements with surfaces ($N = 7$; P2, "I liked how the map which I used most was attached to a surface [...]"). All participants who observed this further noted that they appreciated the passive haptic feedback that came from this arrangement decision ($N = 7$; P11, "I liked how I could use the wall. It felt like a touchscreen."). Beyond touch interaction, several participants remarked that *InteractionAdapt* generally produced layout adaptations that were more comfortable to use ($N = 3$; P8, "The window was placed in a position that was comfortable for my arm"). *InteractionAdapt*'s placements of high usage frequency elements directly in front of users was also appreciated ($N = 3$). Similar to the selection study, participants avoided unintentional collisions in both conditions. We also observed that participants indeed used surfaces more as resting platforms in the *InteractionAdapt* condition. The three participants who reported preferring *Consistency* over *InteractionAdapt* justified this by citing their preference for either a fixed layouts between environments or specific element placement requirements. This suggests that adaptation behaviors may need to be personalized.

Figure 7: Possible use cases for InteractionAdapt: (a) searching for research literature, (b) planning a vacation, and (c) browsing social media. *Left:* Manually created layouts. *Right:* Layouts adapted by InteractionAdapt.

7 APPLICATIONS

We demonstrate InteractionAdapt in three example applications that involve different elements and environments (Figure 7).

Mobile Knowledge Work. First, we demonstrate how InteractionAdapt could support work on the go. In the initial desktop environment, we show a VR interface for literature research, consisting of an excel sheet, two search engines, two document readers, and a keyboard (Figure 7a, *left*). We then show an InteractionAdapt-adapted layout at a printing station (Figure 7a, *right*). In the adapted layout, the keyboard has notably been placed on a flat surface so the researcher can comfortably type while waiting for their printing documents. Both browser windows have also been attached to vertical surfaces to enable interactions that benefit from passive haptic feedback. All elements retain approximately the same relative spatial positions to support reliance on spatial memory. The demonstrated scenario is representative of where we expect adaptation to benefit most, namely in settings where users briefly want to perform a task. These situations often do not allow users to clear their space for a set-up typically required by current VR applications. An adaptive workspace addresses this requirement.

Travel Planning. Figure 7b (*left*) shows a VR interface for travel planning. The layout consists of a keyboard, a map, two booking websites, a translation widget, and a word processor. We envision a user getting up for coffee. Here, our adaptive interface enables them to continue their task while waiting at their kitchen counter, which is more spatially constrained (Figure 7b, *right*). Elements are either attached to a surface for direct interaction (e.g., placement of the blue booking website), or placed at a distance to avoid unintentional collisions (e.g., placement of the translation application).

Browsing Social Media. In this application scenario, we imagine a user waiting for an order at a coffee shop. The layout consists of a weather website, a video viewer, a news website, and a social media platform (Figure 7c, *left*). After the user moves to a seating area, the interface is automatically adapted (see Figure 7c, *right*). Elements were either aligned to the table in front of the user for touch or placed directly in front at a height accessible by elbow-rested interactions to support more ergonomic input. This scenario again illustrates how VR may be used more ubiquitously.

8 DISCUSSION AND FUTURE WORK

We developed InteractionAdapt to enable immersive interfaces to fluidly transition between physical spaces. In the following, we highlight the design implications of our study results and discuss remaining open questions related to this direction.

Design implications. In our selection study, we observed that our explicit consideration of how physical environment affordances may support different interaction modalities resulted in an increased usage of touch interactions. In our usage study, participants similarly voiced a liking for using performing touch interactions on environment surfaces. Combined, these insights suggest that environment affordances can indeed support more ergonomic interactions, especially through providing passive haptic feedback.

Also, from the feedback of participants who preferred *Consistency* over *InteractionAdapt*, we surmise that aligning optimization objectives with individual user preferences is critical to enabling desirable adaptations. Adaptation approaches need to be personalized, such that they could account for individualized requirements like preferences for fixed layouts.

Incorporating additional contextual factors. Our work explores the trade-off between physical interaction affordance, temporal consistency, and occlusion avoidance in adaptive VR interfaces. It is complementary to prior work that has investigated other factors of the design space of adaptive VR interfaces, such as the user's physical surroundings [11, 12], cognitive load [47], social acceptability [54], and ergonomics [12, 16, 50, 59]. We believe there are research opportunities in both expanding and consolidating the design space. For instance, future research could investigate the social factors of nomadic VR usage, e.g., optimizing adaptations to fit space-constrained settings [40] and not to invade others' personal space [54]. Explicitly considering the impact of inter-element relationships is also an interesting direction. For example, it may be desirable to place related elements closer to each other (e.g., word processor and reference document).It would furthermore be valuable to understand trade-offs between factors, such as when considerations of semantics and ergonomics are in conflict.

Moreover, we believe that several features of our own implementation warrant more in-depth exploration. First, while we currently define usage frequency in terms of proportion of active interactions (e.g., number of selections), a more nuanced characterization of the property could benefit generated layouts. Additionally, while we only implicitly factor in user adjustments with our consistency term, explicit specifications of user preferences could alternatively be encoded as hard constraints. Third, future adaptive approaches could also explore co-opting environment features and objects beyond surfaces, like table edges [29] and pens, as UI elements like tools, buttons, or valuators [30]. Fourth, our proposed approach primarily targets situated usage of workspaces in VR. It would be worthwhile investigating to what extent our formulation can produce feasible adaptations while users are not stationary (e.g., while walking [42]). Last but not least, in InteractionAdapt, we performed interface layout adaptations considering the potential usage of direct touch and cursor input techniques. Prior work has proposed a plethora of other spatial interaction techniques, such as balloon selection [4]. Exploring alternative sets of interaction techniques in interface layout optimization is thus interesting.

Optimization method. InteractionAdapt formulates interface layout adaptation as a linear program [62]. Prior work has explored a multitude of alternative approaches, including artificial potential fields [42], simulated annealing [17], and decision trees [65]. We believe that a comparative evaluations of different computational methods would make for valuable future work.

Additional studies. Our approach would further benefit from both controlled and longitudinal evaluations. Future studies can focus on quantitatively examining interaction behaviors, including the influence of different adaptation approaches on body posture and surface usage. It may be especially interesting to observe and analyze these behaviors with full-body motion capture and contact or pressure sensing on environment features. Additionally, prior work suggested that preserving some degree of spatio-temporal consistency is important because it enables users to leverage their spatial-temporal memory [47]. The trade-off between optimizing for familiarity and optimal interaction remains under-explored, and studies analyzing visual search time during tasks may extend upon our current understanding. Last but not least, we believe more in-depth investigations of how participants adjust interface elements may beneficially inform future adaptation approaches.

Personalizing interface adaptation. Weighing the importance of different factors in UI adaptation methods according to user preference is an open research question in HCI. Current methods model general preferences with respect to the trade-off between different design decisions. Fine tuning such models to the preference of individual users is challenging as it requires a way to inform the system on how a user prioritizes different factors. Due to the non-linear nature of the objective function of adaptive UI methods, it is also unclear how to weight factors to attain UIs that follow the preference of an individual user. In future work, we are interested to solve this problem by closing the loop between a chosen prioritization of factors and the individual preference towards a resulting UI.

9 CONCLUSION

We have presented InteractionAdapt, an approach to adapting immersive workspaces between different physical environments for situated usage by leveraging optimization-based layout reconfiguration to fit the user's current physical surroundings. The novel optimization term in our method exploits environment features to support multiple interaction techniques and adapt the UI accordingly, including on-surface touch, mid-air touch and pinch, and cursor control, whose prioritization stems from our detailed experimental evaluation of these techniques in situated VR settings [12]. In our evaluation of InteractionAdapt during a selection study, our method generated UI adaptations that were preferred and required fewer manual adjustments to serve for efficient interaction. In a second proof-of-concept study, we evaluated InteractionAdapt in a real-world productivity task where participants planned a travel schedule. The results of our studies indicate the potential of our approach in supporting more efficient and ergonomic interaction.

We believe that our approach is a step towards automatically integrating VR interfaces into users' physical environment. We argue that this will be key to enabling mobile VR knowledge work for everyone—using one's personalized UI environment, anywhere and in any setting that affords productivity.

REFERENCES

[1] Rawan Alghofaili, Michael S Solah, Haikun Huang, Yasuhito Sawahata, Marc Pomplun, and Lap-Fai Yu. 2019. Optimizing Visual Element Placement via Visual Attention Analysis. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 464–473. https://doi.org/10.1109/VR.2019.8797816

[2] Ferran Argelaguet and Carlos Andujar. 2009. Efficient 3D Pointing Selection in Cluttered Virtual Environments. *IEEE Computer Graphics and Applications* 29, 6 (2009), 34–43. https://doi.org/10.1109/MCG.2009.117

[3] Harry Baker. 2021. *How To Set Up Your Desk In VR On Oculus Quest 2.* https://uploadvr.com/how-to-set-up-desk-vr-quest-2/

[4] Hrvoje Benko and Steven Feiner. 2007. Balloon Selection: A Multi-Finger Technique for Accurate Low-Fatigue 3D Selection. In *2007 IEEE Symposium on 3D User Interfaces.* https://doi.org/10.1109/3DUI.2007.340778

[5] Verena Biener, Snehanjali Kalamkar, Negar Nouri, Eyal Ofek, Michel Pahud, John J Dudley, Jinghui Hu, Per Ola Kristensson, Maheshya Weerasinghe, Klen Čopič Pucihar, et al. 2022. Quantifying the Effects of Working in VR for One Week. *IEEE Transactions on Visualization and Computer Graphics* (2022).

[6] 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).

[7] Doug A. Bowman, Donald B. Johnson, and Larry F. Hodges. 1999. Testbed Evaluation of Virtual Environment Interaction Techniques. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (London, United Kingdom) *(VRST '99)*. Association for Computing Machinery, New York, NY, USA, 26–33. https://doi.org/10.1145/323663.323667

[8] Dino Caesaron, Chiuhsiang Joe Lin, Ilma Mufidah, and Anastasia Febiyani. 2022. Evaluating Direct Pointing and Indirect Cursor Interactions with Fitts' Law in Stereoscopic Environments. *Journal of Engineering & Technological Sciences* 54, 1 (2022).

[9] Lung-Pan Cheng, Eyal Ofek, Christian Holz, Hrvoje Benko, and Andrew D. Wilson. 2017. Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3718–3728. https://doi.org/10.1145/3025453.3025753

[10] Lung-Pan Cheng, Eyal Ofek, Christian Holz, and Andrew D. Wilson. 2019. VRoamer: Generating On-The-Fly VR Experiences While Walking inside Large, Unknown Real-World Building Environments. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 359–366. https://doi.org/10.1109/VR.2019.8798074

[11] 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

[12] Yi Fei Cheng, Tiffany Luong, Andreas Rene 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)*. 150–159. https://doi.org/10.1109/ISMAR55827.2022.00029

[13] Barrett Ens, Eyal Ofek, Neil 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* (Los Angeles, California, USA) *(SUI '15)*. Association for Computing Machinery, New York, NY, USA, 65–68. https://doi.org/10.1145/2788940.2788954

[14] Barrett Ens, Eyal Ofek, Neil Bruce, and Pourang Irani. 2016. *Shared Façades: Surface-Embedded Layout Management for Ad Hoc Collaboration Using Head-Worn Displays*. Springer International Publishing, Cham, 153–176. https://doi.org/10.1007/978-3-319-45853-3_8

[15] Barrett M. Ens, Rory Finnegan, and Pourang P. Irani. 2014. The Personal Cockpit: A Spatial Interface for Effective Task Switching on Head-Worn Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 3171–3180. https://doi.org/10.1145/2556288.2557058

[16] João Marcelo Evangelista Belo, Anna Maria Feit, Tiare Feuchtner, and Kaj Grønbæk. 2021. XRgonomics: Facilitating the Creation of Ergonomic 3D Interfaces. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 290, 11 pages. https://doi.org/10.1145/3411764.3445349

[17] João Marcelo Evangelista Belo, Mathias N. Lystbæk, Anna Maria Feit, Ken Pfeuffer, Peter Kán, Antti Oulasvirta, and Kaj Grønbæk. 2022. AUIT – the Adaptive User Interfaces Toolkit for Designing XR Applications. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22)*. Association for Computing Machinery, New York, NY, USA, Article 48, 16 pages. https://doi.org/10.1145/3526113.3545651

[18] Cathy Mengying Fang, Ryo Suzuki, and Daniel Leithinger. 2023. VR Haptics at Home: Repurposing Everyday Objects and Environment for Casual and On-Demand VR Haptic Experiences. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 312, 7 pages. https://doi.org/10.1145/3544549.3585871

[19] Andreas Fender, Philipp Herholz, Marc Alexa, and Jörg Müller. 2018. Optispace: automated placement of interactive 3D projection mapping content. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–11.

[20] Andreas 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

[21] Andris Freivalds. 2011. *Biomechanics of the upper limbs: mechanics, modeling and musculoskeletal injuries*. CRC press.

[22] 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)*. 207–212. https://doi.org/10.1109/ISMAR.2014.6948429

[23] 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

[24] Jens Grubert, Tobias Langlotz, Stefanie Zollmann, and Holger Regenbrecht. 2017. Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 23, 6 (2017), 1706–1724. https://doi.org/10.1109/TVCG.2016.2543720

[25] Jens Grubert, Eyal Ofek, Michel Pahud, and Per Ola Kristensson. 2018. The Office of the Future: Virtual, Portable, and Global. *IEEE Computer Graphics and Applications* 38, 6 (2018), 125–133. https://doi.org/10.1109/MCG.2018.2875609

[26] Gurobi. 2021. *Gurobi - the fastest solver*. https://gurobi.com/

[27] Chris Hand. 1997. A survey of 3D interaction techniques. In *Computer graphics forum*, Vol. 16. Wiley Online Library, 269–281.

[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* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300577

[29] Fengming He, Xiyun Hu, Jingyu Shi, Xun Qian, Tianyi Wang, and Karthik Ramani. 2023. Ubi Edge: Authoring Edge-Based Opportunistic Tangible User Interfaces in Augmented Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 461, 14 pages. https://doi.org/10.1145/3544548.3580704

[30] Steven Henderson and Steven Feiner. 2009. Opportunistic tangible user interfaces for augmented reality. *IEEE Transactions on Visualization and Computer Graphics* 16, 1 (2009), 4–16.

[31] Anuruddha Hettiarachchi and Daniel Wigdor. 2016. Annexing Reality: Enabling Opportunistic Use of Everyday Objects as Tangible Proxies in Augmented Reality. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1957–1967. https://doi.org/10.1145/2858036.2858134

[32] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. ACM, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130

[33] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. Passive Real-World Interface Props for Neurosurgical Visualization. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, Massachusetts, USA) *(CHI '94)*. ACM, USA, 452–458. https://doi.org/10.1145/191666.191821

[34] E Hoddes, V Zarcone, and W Dement. 1972. Stanford sleepiness scale. *Enzyklopädie der Schlafmedizin* 1184, 1 (1972).

[35] Immersed. 2021. *Immersed*. https://immersed.com/

[36] Brent Edward Insko. 2001. *Passive haptics significantly enhances virtual environments*. The University of North Carolina at Chapel Hill.

[37] J. Jankowski and M. Hachet. 2015. Advances in Interaction with 3D Environments. *Computer Graphics Forum* 34, 1 (2015), 152–190. https://doi.org/10.1111/cgf.12466 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.12466

[38] Tero Jokela, Jarno Ojala, and Thomas Olsson. 2015. A Diary Study on Combining Multiple Information Devices in Everyday Activities and Tasks. 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, 3903–3912. https://doi.org/10.1145/2702123.2702211

[39] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 637–644. https://doi.org/10.1145/2642918.2647383

[40] Mohamed Kari and Christian Holz. 2023. HandyCast: Phone-Based Bimanual Input for Virtual Reality in Mobile and Space-Constrained Settings via Pose-and-Touch Transfer. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 528, 15 pages. https://doi.org/10.1145/3544548.3580677

[41] Panayiotis Koutsabasis and Panagiotis Vogiatzidakis. 2019. Empirical Research in Mid-Air Interaction: A Systematic Review. *International Journal of Human–Computer Interaction* 35, 18 (2019), 1747–1768. https://doi.org/10.1080/10447318.2019.1572352 arXiv:https://doi.org/10.1080/10447318.2019.1572352

[42] Wallace S. 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* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 356–366. https://doi.org/10.1145/3301275.3302278

[43] Joseph LaViola. 2017. *3D User Interfaces* (2nd edition ed.). Addison-Wesley Professional.

[44] Zhihao Liang, Zhihao Li, Songcen Xu, Mingkui Tan, and Kui Jia. 2021. Instance Segmentation in 3D Scenes using Semantic Superpoint Tree Networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2783–2792.

[45] Chuan-en Lin, Ta Ying Cheng, and Xiaojuan Ma. 2020. *ARchitect: Building Interactive Virtual Experiences from Physical Affordances by Bringing Human-in-the-Loop*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376614

[46] Robert W. Lindeman, John L. Sibert, and James K. Hahn. 1999. Towards Usable VR: An Empirical Study of User Interfaces for Immersive Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. ACM, USA, 64–71. https://doi.org/10.1145/302979.302995

[47] 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

[48] 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* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 550, 16 pages. https://doi.org/10.1145/3491102.3517723

[49] Weizhou Luo, Anke Lehmann, Hjalmar Widengren, and Raimund Dachselt. 2022. Where Should We Put It? Layout and Placement Strategies of Documents in Augmented Reality for Collaborative Sensemaking. 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 627, 16 pages. https://doi.org/10.1145/3491102.3501946

[50] Tiffany Luong, Yi Fei Cheng, Max Moebus, Andreas Fender, and Christian Holz. 2023. Controllers or Bare Hands? A Controlled Evaluation of Input Techniques on Interaction Performance and Exertion in Virtual Reality. *IEEE Transactions on Visualization & Computer Graphics* (2023).

[51] Sebastian Marwecki, Andrew D. Wilson, Eyal Ofek, Mar Gonzalez Franco, and Christian Holz. 2019. Mise-Unseen: Using Eye Tracking to Hide Virtual Reality Scene Changes in Plain Sight. 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, 777–789. https://doi.org/10.1145/3332165.3347919

[52] Lynn McAtamney and E. Nigel Corlett. 1993. RULA: a survey method for the investigation of work-related upper limb disorders. *Applied Ergonomics* 24, 2 (1993), 91–99. https://doi.org/10.1016/0003-6870(93)90080-S

[53] Mark Mcgill, Aidan Kehoe, Euan Freeman, and Stephen Brewster. 2020. Expanding the bounds of seated virtual workspaces. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 3 (2020), 1–40.

[54] Daniel Medeiros, Mark McGill, Alexander Ng, Robert McDermid, Nadia Pantidi, Julie Williamson, and Stephen Brewster. 2022. From Shielding to Avoidance: Passenger Augmented Reality and the Layout of Virtual Displays for Productivity in Shared Transit. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3640–3650. https://doi.org/10.1109/TVCG.2022.3203002

[55] Manuel Meier, Paul Streli, Andreas Fender, and Christian Holz. 2021. TapID: Rapid touch interaction in virtual reality using wearable sensing. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 519–528.

[56] Tim Menzner, Travis Gesslein, Alexander Otte, and Jens Grubert. 2020. Above surface interaction for multiscale navigation in mobile virtual reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 372–381.

[57] Meta. 2022. *Horizon Workrooms*. https://www.oculus.com/workrooms/

[58] Mark R Mine. 1995. Virtual environment interaction techniques. *UNC Chapel Hill CS Dept* (1995).

[59] Roberto A. Montano Murillo, Sriram Subramanian, and Diego Martinez Plasencia. 2017. Erg-O: Ergonomic Optimization of Immersive Virtual Environments. 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, 759–771. https://doi.org/10.1145/3126594.3126605

[60] Aziz Niyazov, Barrett Ens, Kadek Ananta Satriadi, Nicolas Mellado, Loic Barthe, Tim Dwyer, and Marcos Serrano. 2023. User-Driven Constraints for Layout Optimisation in Augmented Reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 35, 16 pages. https://doi.org/10.1145/3544548.3580873

[61] Eyal Ofek, Jens Grubert, Michel Pahud, Mark Phillips, and Per Ola Kristensson. 2020. Towards a Practical Virtual Office for Mobile Knowledge Workers. (2020). https://doi.org/10.48550/ARXIV.2009.02947

[62] 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

[63] Seonwook Park, Christoph Gebhardt, Roman Rädle, Anna Maria Feit, Hana Vrzakova, Niraj Ramesh Dayama, Hui-Shyong Yeo, Clemens N. 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* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173758

[64] Leonardo Pavanatto, Chris North, Doug A Bowman, Carmen Badea, and Richard Stoakley. 2021. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 759–767.

[65] 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 *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 65, 18 pages. https://doi.org/10.1145/3491102.3517665

[66] 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

[67] Spatial. 2022. *Spatial*. https://spatial.io/

[68] Patrick L. Strandholt, Oana A. Dogaru, Niels C. Nilsson, Rolf Nordahl, and Stefania Serafin. 2020. *Knock on Wood: Combining Redirected Touching and Physical Props for Tool-Based Interaction in Virtual Reality*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376303

[69] Paul Streli, Rayan Armani, Yi Fei Cheng, and Christian Holz. 2023. HOOV: Hand Out-Of-View Tracking for Proprioceptive Interaction Using Inertial Sensing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 310, 16 pages. https://doi.org/10.1145/3544548.3581468

[70] Paul Streli, Jiaxi Jiang, Andreas Rene Fender, Manuel Meier, Hugo Romat, and Christian Holz. 2022. TapType: Ten-finger text entry on everyday surfaces via Bayesian inference. In *CHI Conference on Human Factors in Computing Systems*. 1–16.

[71] Paul Streli, Jiaxi Jiang, Juliete Rossie, and Christian Holz. 2023. Structured Light Speckle: Joint egocentric depth estimation and low-latency contact detection via remote vibrometry. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery. https://doi.org/10.1145/3586183.3606749

[72] Rafael Veras, Gaganpreet Singh, Farzin Farhadi-Niaki, Ritesh Udhani, Parth Pradeep Patekar, Wei Zhou, Pourang Irani, and Wei Li. 2021. Elbow-Anchored Interaction: Designing Restful Mid-Air Input. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15.

[73] Raquel Viciana-Abad, Arcadio Reyes Lecuona, and Matthieu Poyade. 2010. The Influence of Passive Haptic Feedback and Difference Interaction Metaphors on Presence and Task Performance. *Presence* 19, 3 (2010), 197–212. https://doi.org/10.1162/pres.19.3.197

[74] Patricio Gonzalez Vivo and Jen Lowe. 2015. The book of shaders. *Dosegljivo: https://thebookofshaders. com* (2015).

[75] vSpatial. 2021. *The workplace of the future is here.* https://vspatial.com/

[76] Varjo Workspace. 2022. *Varjo*. https://varjo.com/use-center/get-to-know-your-headset/varjo-workspace/

[77] Jackie Yang, Christian Holz, Eyal Ofek, and Andrew D. Wilson. 2019. DreamWalker: Substituting Real-World Walking Experiences with a Virtual Reality. In *Proc. 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1093–1107. https://doi.org/10.1145/3332165.3347875

[78] Qian Zhou, George Fitzmaurice, and Fraser Anderson. 2022. In-Depth Mouse: Integrating Desktop Mouse into Virtual Reality. In *CHI Conference on Human Factors in Computing Systems*. 1–17.

[79] Daniel Zielasko, Marcel Krüger, Benjamin Weyers, and Torsten W. Kuhlen. 2019. Passive Haptic Menus for Desk-Based and HMD-Projected Virtual Reality. In *2019 IEEE 5th Workshop on Everyday Virtual Reality (WEVR)*. 1–6. https://doi.org/10.1109/WEVR.2019.8809589

[80] Daniel Zielasko, Benjamin Weyers, and Torsten W. Kuhlen. 2019. A Non-Stationary Office Desk Substitution for Desk-Based and HMD-Projected Virtual Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 1884–1889. https://doi.org/10.1109/VR.2019.8797837

A NOTATION

Element variables

Variable	Description
$E = (e_1, ..., e_n)$	Set of virtual interface elements to place in the layout
$N_e \in \mathbb{Z}^+$	Number of virtual elements in E
$p_e \in \mathbb{R}^3$	Position of element e in the input environment
$d_e \in \mathbb{R}^3$	Size of element e
$u_e \in [0, 1]$	Usage frequency of element e

Environment variables

Variable	Description
$O = (o_1, ..., o_n)$	Set of considered obstacles in the environment
$N_o \in \mathbb{Z}^+$	Number of obstacles in O
$p_o \in \mathbb{R}^3$	Position of physical obstacle o
$d_o \in \mathbb{R}^3$	Size of physical obstacle o
$S = (s_1, ..., s_n)$	Set of considered physical surfaces in the environment
$N_s \in \mathbb{Z}^+$	Number of physical surfaces in S
$p_s \in \mathbb{R}^3$	Position of physical surface s in the environment
$d_s \in \mathbb{R}^3$	Size of physical surface s
$C = (c_1, ..., c_n)$	Set of containers that are available for placement
$p_c \in \mathbb{R}^3$	Position of container c
$d_c \in \mathbb{R}^3$	Size of container c

Table 1: Description of model inputs.