# Design Principles and Challenges for Gaze + Pinch Interaction in XR

Ken Pfeuffer, *Aarhus University, Aarhus, Denmark*

Hans Gellersen, *Lancaster University, Lancaster, United Kingdom; Aarhus University, Aarhus, Denmark*

Mar Gonzalez-Franco, *Google, Seattle, Washington, United States*

*Abstract—For Extended Reality (XR) headsets, a key aim is the natural interaction in 3D space beyond what traditional methods of keyboard, mouse, and touchscreen can offer. With the release of the Apple Vision Pro, a novel interaction paradigm is now widely available where users seamlessly navigate content through the combined use of their eyes and hands. However, blending these modalities poses unique design challenges due to their dynamic nature and the absence of established principles and standards. In this article, we present five design principles and issues for the Gaze + Pinch interaction technique, informed by eye-hand research in the human-computer interaction field. The design principles encompass mechanisms like division of labor and minimalistic timing, which are crucial for usability, alongside enhancements for the manipulation of objects, indirect interactions, and drag & drop. Whether in design, technology, or research domains, this exploration offers valuable perspectives for navigating the evolving landscape of 3D interaction.*

Interaction, innovative control methods, and natural user interfaces (UIs) have long been recognized as significant challenges in achieving a truly immersive and intuitive Extended Reality (XR) experience [2], [4]. However, the input landscape so far remained unsatisfactory, plagued by usability issues such as physical fatigue, ergonomic discomfort, and complex interface designs. The challenge of getting the interface right is amplified by the fragmented input landscape – controllers, hand tracking, eye movements, and voice commands. XR operating systems can courageously strive toward universal support, but it's tricky to unify all possible inputs and combinations in one system. This leaves interaction systems often primitive, often at the exclusion of one modality in favor of another without fully considering their unity.

It is challenging to achieve a harmonious integration of multiple modalities and optimize the effectiveness across various tasks. Yet, the multimodal input trend solidifies with the arrival of the 'Gaze + Pinch' XR paradigm: *glance at a UI element with your eyes, then simply pinch with your fingers to activate it*. In the scientific community, researchers have been studying general eye-hand interaction in general since the 80s [5] and the specific "Gaze + Pinch" model since 2017 [15]. Yet, we see this as a clear innovation with a rapidly growing adoption across XR headsets such as the Microsoft Hololens 2 or Magic Leap, and especially the OS-wide integration with the Apple Vision Pro.

Comprehensive guidance on multimodal interaction design is rare, with most focusing on one modality or generalising across input devices [1], [9]. Integrating both raises questions: when to use eyes, hands, or both? How to merge the signals in time and space for optimal ease-of-use and expressiveness? In our paper, we aim to highlight our findings, establish principles, and suggest frameworks based on scientific experiments to guide designers, developers, and researchers in navigating this new interaction paradigm.

We present 5 design principles and 5 design issues, drawing insights from our eye-hand research and scientific articles in the area of human-computer interaction. Despite that in our daily lives as scientists we explore all possible future directions, our process of converging and abstraction eventually led to cover

XXXX-XXX
Digital Object Identifier 10.1109/XXX.0000.0000000

this set of principles. This abstraction effort covers the most pressing problem: the specific mechanisms of division of labor and multimodal timing that are key for usability; as well as issues of manipulation of objects with features like indirect gesture and drag & drop.

We strive for this article to captivate not only the scientific community but also to offer diverse perspectives that resonate with practitioners across various fields:

- For designers: This article explores innovative approaches to Gaze + Pinch interaction, offering valuable inspiration.
- For technologists: This article highlights the technical challenges and opportunities of Gaze + Pinch interaction.
- For researchers: This article examines the latest research on Gaze + Pinch interaction, offering valuable insights at the intersection between XR, eye-tracking, and multimodal UI.

## SUMMARY

The bulk of our work on this paper has been to prepare the essence of design principles, that include:

1) **Division of labor**: Use a clear separation of tasks: the eyes perform selection tasks, the hands do the actual manipulation work.
2) **Minimalistic timing**: One moment matters for the eyes–the moment thumb and index finger contact, the hands take over, relieving the eyes from the explicit motor control tasks.
3) **Flexible gesture**: Gaze affords lightweight and flexible gesturing, allowing transitions from one vs. two-hands to single vs. multi-target control.
4) **Infallible eyes**: eyes operate instantly, with constant accuracy. With good tracking, we can't miss or overshoot a target when we look at something.
5) **Multimodal by design**: Gaze + Pinch complements direct gestures- understand which tasks can be accomplished with one or another and provide transitions to get the best of both worlds.

We also discuss 5 behavioural design issues:

1) **(Un)Learning**: The Gaze + Pinch interaction challenges conventional action by enabling pointing without physical hand motion, emphasizing the need to unlearn habitual actions for efficiency.
2) **Early and late triggers**: The eye-hand timing is key. But ideally UIs are supportive when manual commands precede or lag behind gaze fixation, as errors can result.
3) **Input Mappings**: To be efficient across near and far spaces, consider control-display ratios and speed amplification like mouse acceleration to improve control flexibility.
4) **Drag & drop sequences**: There are challenges in re-engaging with dropped objects– UIs can reduce dragging use and provide drop prevention.
5) **Continuous eye-input**: Exceptions like pinch-to-zoom showcase natural continuous eye inputs but require careful integration.

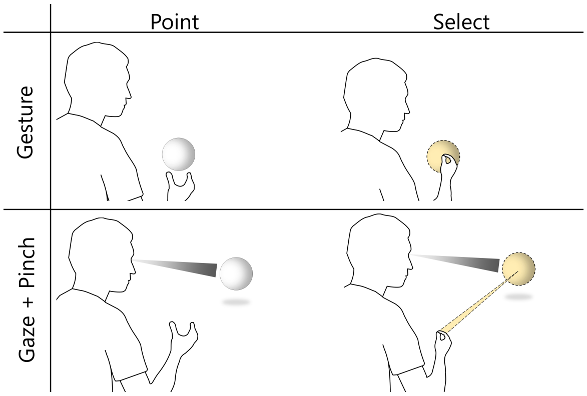

**FIGURE 1.** Basic operation of Gaze + Pinch in contrast to direct manipulation, demonstrating how similar (and, thus, easy to learn) the gestures are, as well as how the eyes extend reach to any object the user sees. Adapted from [15]

## GAZE + PINCH

Gaze + Pinch users can directly manipulate objects they look at using familiar gestures like pinch-to-select or two-handed scaling. Even when they are far from what would be considered a direct interaction space.

The term "Gaze + Pinch" originated from our 2017 paper [15], where we studied the foundations for this particular combo of eyes and hands. Gaze + Pinch stems from Pfeuffer's prior doctoral thesis on the unity of gaze and multi-touch gestures [16], superseding the earlier "Gaze-Touch" technique [12].

As per Poupyrev et al.'s taxonomy [7], Gaze + Pinch is categorized as an egocentric input method, offering a first-person view and employing the virtual pointer metaphor with the eyes as the pointer. Nevertheless, its manipulation closely mirrors direct manipulation (Figure 1), resulting in the adoption of numerous traits from the familiar virtual hand technique. Key distinctions with other main input techniques are:

- **Gaze + Pinch *vs.* Hand Gesture**: Gaze + Pinch allows users to interact with objects from a distance using the same gesture set, expanding the effective interaction area and maximizing the virtual environment's vast space.

- **Gaze + Pinch *vs.* Controller:** Gaze + Pinch liberates users from holding physical devices, enabling them to perform hand gestures on distant objects as if directly manipulating them. This makes the UI highly intuitive, tapping into the inherent spatial manipulation skills of humans.

## DESIGN PRINCIPLES

We go deeper into each of our key considerations that should guide the design and implementation of UIs based on Gaze + Pinch. In setting up these principles we strive for a balance between simple yet powerful 3D manipulation capabilities.

### Division of labor: The eyes select, the hands manipulate

Our eyes' natural role involves indicating points of interest, and we can easily look at any point at will. In contrast, the hands are adept at physical manipulation through the interplay of finger movement and hand posture. Use a clear separation of concerns: the eyes perform selection tasks, the hands do the actual confirmation or manipulation work. This avoids the pitfalls of (i) overloading the eyes with explicit motor control tasks [20] – you only actively "use" the eyes to select, (ii) physical fatigue [21] –gaze pointing minimizes the hands' physical motion needs, and (iii) supporting the (most) naturalistic roles for each modality.

The hands, then, make indirect gestures. This is similar to a controller in having the ability to interact at a distance, but now with intuitive pinch gestures. Indeed, there are hands-only techniques for selection and manipulation, such as the 'handray' raypointing coupled with a pinch-confirmation (e.g., used by the Meta Quest 3 and Hololens 2). Yet, assigning both selection and manipulation to the hand can be susceptible to hand jitter issues (the Heisenberg problem [18]). Our studies showed that Gaze + Pinch (and other eye-hand techniques) leads to improved performance and comfort for 3D selection over gestures alone for interaction over distance [22], [8].

### Minimalistic timing

There are many ways to mix & match the eye and hand tracking signals. A poorly-designed multimodal input fusion can amplify complexity [10], especially with the eyes that can be wandering around and accidentally select things [20]. At the same time, it is important to take advantage of the eyes' prime strength to offer instant selection.

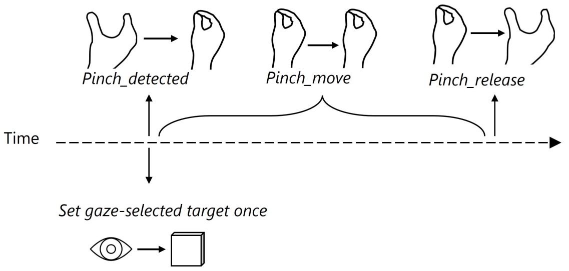

**FIGURE 2.** Eye-tracking as input is only active the moment that a pinch gesture is registered to avoid erratic behaviors when the eyes are wandering around.

The primary model for Gaze + Pinch only use a single moment in time for the eyes: the moment that the index finger and thumb have first contact, one has to fixate on the desired target (Figure ). This instantaneous approach to the interaction is key– but only for the selection. For follow-up object manipulations, such as a drag, pan, or zoom gesture, the hands take over. This affords the freedom to inspect the surroundings independently and avoids accidental actions by eye or hand inputs. For instance in drag & drop, after selection one can freely look around to locate the destination for the dragged object and follow with the hand via indirect control.

In contrast, a hands-only UI typically means you can point with your hand to the target, without continuously monitoring the target. Gaze + Pinch inverses the relationship: the eyes must be on the target but the hands can be anywhere. This is a fundamentally new behavioral pattern that users got to master. Employing gaze minimally, and using the standard gesture set, facilitates a quick adaptation of this new relationship. Beginners may find themselves more attentive to ensure their gaze is on the UI element until receiving the right feedback. More experienced users may swiftly execute a Gaze + Pinch command even without having fully perceived the target and its selection feedback.

### Flexible gesture

Hand gestures are in control of virtual objects acquired by the eyes. Particularly, the commonly used selection, manipulation and navigation commands can be covered when employing only pinch-tap and pinch-drag gestures. This flexibly extends to all atomic classes of the hand-based manipulation – one vs. two-handed interactions, and single vs. multiple target manipulation (Figure 3). Users can seamlessly shift between 1 or

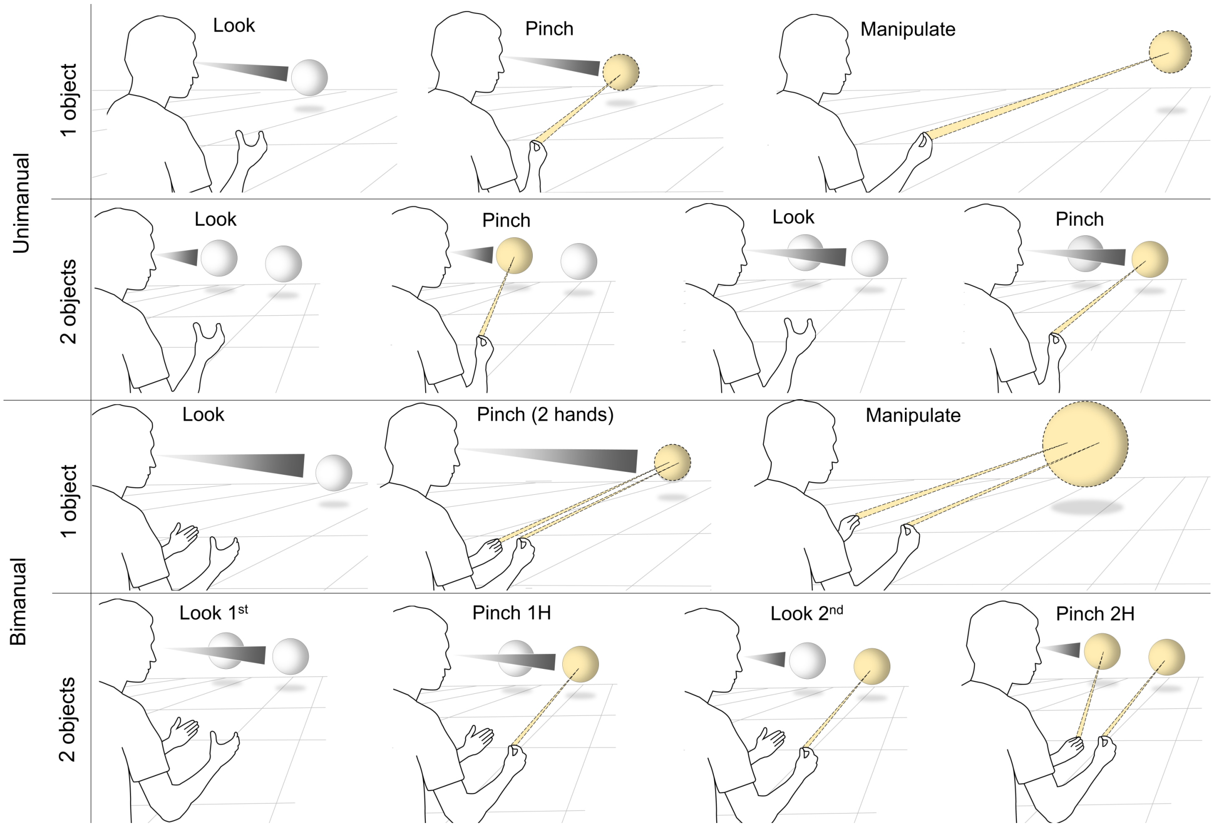

**FIGURE 3.** Fundamental interaction classes with Gaze + Pinch, natively supporting the atomic classes of one and two handed inputs, and single and multi-object interaction. Taken from [15].

two objects and hands by simply glancing and re-engaging pinch gestures as desired. This is inherited from the default hand-tracking gestures – but the integration with gaze, and the elimination of the manual pointing sub-task renders all those basic hand actions extremely lightweight and flexible to use across space.

## Infallible eyes

Our hands adhere to a speed-accuracy trade-off: faster normally means less accurate when it comes to hands [19]. In contrast, the eyes can fixate on a target in almost instant time, even if the target is in motion, with constant accuracy given by the eye-tracking sensor. From a user's perspective, the eyes are infallible: hand pointing can miss or overshoot a target, there are even Olympic competitions on target shooting, but the eyes can't miss as we are either on-target or we look elsewhere. It is crucial for an interaction system that engages eyes to support the simple way of just looking to select, without manual effort. That is of course if sensors and tracking were perfect. However, inaccurate eye-tracking can prompt users to undergo correction measures, such as squinting their eyes or adjust their head position in vain attempts to correct precision limitations, that in turn leads to increased mental exertion and longer selection times. It is perhaps that limitation that has hindered the popularization of Gaze + Pinch until now.

Some of the issues derive from intrinsics of human vision, like eye dominance, or sub-optimal eye strain when looking at targets above the horizon. We are better at looking down than up. In a way eyes are not as symmetric as people might intuitively think. And that means that even the most basic menus, such as home menus, might need to be reconsidered, perhaps they are better off at the bottom of the screen or even on a circular distribution.

Some examples of design considerations for Gaze + Pinch UIs include somewhat counter-intuitive aspects, such as large buttons to achieve a low error rate. In reality, large buttons invite wandering around with the eyes, potentially leading to outliers. Drawing the most salient parts to the center of the button will be welcomed by the selection mechanism, and a

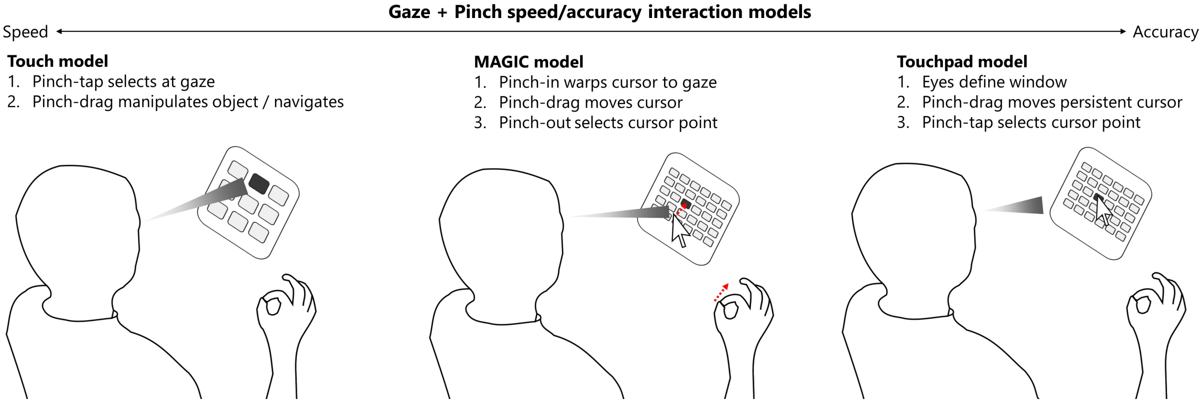
Gaze + Pinch speed/accuracy interaction models
Speed
Accuracy
Touch model
1. Pinch-tap selects at gaze
2. Pinch-drag manipulates object / navigates
MAGIC model
1. Pinch-in warps cursor to gaze
2. Pinch-drag moves cursor
3. Pinch-out selects cursor point
Touchpad model
1. Eyes define window
2. Pinch-drag moves persistent cursor
3. Pinch-tap selects cursor point

**FIGURE 4.** Interaction models for Gaze + Pinch, with different speed/accuracy trade-offs. Adopted from [15]

generous buffer space around targets makes outliers less impactful. That is, it is better to leave space between buttons than to have large buttons.

For compatibility reasons, not all UIs can be redesigned. If smaller targets are required, hand-refinement and cursor extensions can be integrated with Gaze + Pinch, although departing from the original simplicity. Figure 4 presents interaction models to accommodate both UI requirements and sensor limitations. These methods leverage the Gaze + Pinch signal in different ways to provide flexibility in interaction. The Touch model serves as the default, where a Gaze + Pinch command selects a target and a drag gesture executes manipulation. The MAGIC model (Manual and Gaze Input Cascaded [20]) enhances precision by introducing a one-time mouse cursor: upon pinch-in, a cursor appears at the gaze position, which is then moved by gestural motion; upon pinch-out, the object under the cursor is selected. The Touchpad model ensures full precision with a persistent mouse in the window. Here, eye movement is utilized solely for selecting the window, while pinch tap (click) and pinch-drag gestures control the mouse. These models can be selectively implemented by UI systems based on the precision requirements of the application context.

### Multimodal by design

The motto is to get the best of both worlds. Gaze + Pinch can also work together with the direct gestures in nearspace. This is possible via mode-switching methods that use time and space multiplexing of the inputs [13]. This can have the neat side effect that one can rapidly use direct and indirect inputs at a glance (Figure ). For time-multiplexing, imagine a user opens an app with a menu through Gaze + Pinch, which activates the app UI in front of the user. The user switches to direct touch gestures to scroll the app's content. In space multiplexing, picture holding a menu with one hand while using Gaze + Pinch commands with the other, enabling direct and indirect inputs at the same time for on-hand virtual menus. Hand menus usually position menus off the hand to prevent hand-tracking interference. Indirect gestures are spatially separated, avoiding hand overlap (Figure 6). It's one of the intriguing outcomes when UI systems support transitions between complementary modes of interaction.

## BEHAVIOURAL ISSUES

Learning a new vocabulary of interactions can feel as challenging as learning a new language from a cognitive perspective. This effect might be even stronger for folks who are experts in current paradigms. But the magic of Gaze + Pinch is that there is little learning to be done. Users employ familiar gestures—such as pinch for selection, translation, rotation, and scaling—reminiscent of direct manipulation, yet distinct in enabling interaction across a broader spatial range. This fusion of familiar and novel features, facilitated by eye gaze, defines the hybrid Gaze + Pinch technique. Perhaps, it's more about the (un)learning. And the natural interactions it unlocks beyond controllers or hand-gestures.

### (Un)Learning

When we want to acquire a target, grab something, we intuitively move toward it. This is partially intuitive because of the physics of the real world, it might even be coded on a very deep layer on our brain. Gaze + Pinch commands don't need this movement anymore, which can be considered almost counter-

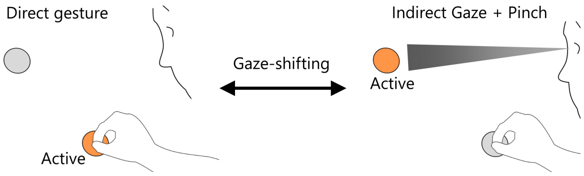

**FIGURE 5.** Best of both worlds: UIs can aim to support both gesture and Gaze + Pinch control through UI transitions. Adopted from [13]

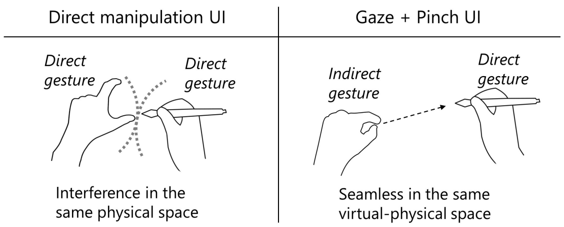

**FIGURE 6.** Mixing direct and indirect inputs allows for novel bimanual dynamics without physical interference.

intuitive. Hence it is important to balance the advantage of using a new way that initially requires rethinking of the action process, over just using the hands without the eyes. In a sense, Gaze + Pinch does not mean learning a new way of interaction – it's about unlearning the common way: don't move your hands, just confirm right where you are with your hand what your eyes have selected. Thus, the learning effort is rather negligible. But changing the nature of action might have consequences we haven't fully looked into yet. Transitioning from XR experiences with Gaze + Pinch to those without introduces a perceptible discord. It is open how this big change might affect the behavior of a new generation who may grow up using this UI. What long-term impacts could this have? These questions remain open and necessitate careful consideration.

## Early and late triggers

A problem of any multimodal interaction, and as such a problem that might arise in Gaze + Pinch, is that the eyes may leave the target before a manual command is registered, or the command is issued just before the gaze lands on the target [6]. It's possible to have a predictive and generous timing, e.g., using the last fixation ( 200–300ms of a stable gaze), rather than the current gaze coordinate (see fixation detection methods [3]). An error in this space of multimodal integration is defined in neuroscience as a body semantic violation [11]. If the early or late trigger frequency is known, e.g., through knowledge about user context and application, the timing can be adapted.

## Input Mappings

Direct manipulation means a 1:1 control-display mapping from hand to object. With Gaze + Pinch, after selection the user's hand indirectly controls an object. Using a 1:1 mapping between physical hand motion and virtual object makes the interaction feel slow. What one can do is amplify the speed in the transfer function with the object distance. And this is possible because it is naturally a relative interaction paradigm, more akin to the mouse than any other form of natural input. This means, we can even use visual angle to determine dragging speed as a distance-independent metric: if your hand moves by 5 degrees in your FoV, the remote object corresponds with 5 degree motion. This works well for objects at a distance but may be confusing when targets are near– here the UI can revert to a 1:1 transfer function.

## Drag & drop sequences

In real life, dropping an object means, hopefully your hand is right there for you to pick it up again. Same happens with direct interactions in XR. But with Gaze + Pinch, you can look away after dropping and finding it again means you have to look back. This can be a hassle, especially if turning your body is involved. So, when designing UIs, it's crucial to think about what type of tasks are supported – ideally, most tasks require only a single action to finish drag & drop, and for sequences of manipulations consider potential enhancements. Hand tracking systems can make sure not to disengage the target from the control of a pinch gesture if just briefly undetected, to keep dragging robust and avoid object loss.

## Continuous eye-selection

The minimal use principle for eye-based input is a principle rule, but there are cases where it can extend naturally to continuous eye inputs. For example, for zooming into a map. (1) The conservative default would be to set the zooming pivot to the gaze position at the initial pinch-in event, and then allow the hand position to adjust the pivot. This makes zooming like direct manipulation, where the physical input position remains at the virtual position on the underlying map. (2) An alternative model is to use the eyes continuously as input in parallel to the hand gesture. When the eyes focus on a different area during zooming, the zooming target adjusts accordingly. The on-line re-positioning can lead to more accurate zooming for goal-oriented navigation tasks, and in turn, reduce the need for panning gestures that correct zooming operations afterward [14]. Doing this will put more responsibility on the eyes, which for goal-oriented zooming can feel natural. Additionally, the choice between the two models can be informed by the task requirements, i.e., how much the eyes are expected to remain at the area of interest when performing a zoom gesture.

## CONCLUSION

The eye-hand interaction design for 3D experiences is a novel space that is gradually gaining momentum. Thanks to the groundwork laid by scientists and researchers, we're well-equipped to explore this field further and anticipate exciting UX developments. Our focus on distilling the essentials of Gaze + Pinch interaction provides a deeper understanding of how to achieve the right balance to achieve a simple-to-use but expressive UI, drawing on basic design principles as well as practical considerations from experience.

Eye-tracking technology can transform how we use our hands, opening up new possibilities for XR interaction. Since the inception of the Gaze + Pinch concept, researchers have been relentlessly advancing the space of multimodal UIs, including our work on advanced selection [8], [22] and one-handed inputs [17]. We are excited to see how these ideas play out, and what else lies ahead to advance our interactive experience through a symbiosis of our eyes and hands.

## REFERENCES

1. Apple (2023). Eyes. Developer Documentation. https://developer.apple.com/design/human-interface-guidelines/eyes. Accessed 3-11-24.
2. Azuma, R. T. (2016). The most important challenge facing augmented reality. Presence, 25(3), 234-238.
3. Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., and Nyström, M. (2017). One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. Behavior research methods, 49, 616-637.
4. Billinghurst, M., Clark, A., and Lee, G. (2015). A survey of augmented reality. Foundations and Trends in Human–Computer Interaction, 8(2-3), 73-272.
5. Richard A. Bolt. 1981. Gaze-orchestrated dynamic windows. In Proceedings of the 8th annual conference on Computer graphics and interactive techniques (SIGGRAPH '81). Association for Computing Machinery, New York, NY, USA, 109–119.
6. Manu Kumar, Jeff Klingner, Rohan Puranik, Terry Winograd, and Andreas Paepcke. 2008. Improving the accuracy of gaze input for interaction. In Proceedings of the 2008 symposium on Eye tracking research & applications (ETRA '08). ACM, New York, NY, USA, 65–68.
7. LaViola Jr, J. J., Kruijff, E., McMahan, R. P., Bowman, D., and Poupyrev, I. P. (2017). 3D user interfaces: theory and practice. Addison-Wesley Professional.
8. Mathias N. Lystbæk, Peter Rosenberg, Ken Pfeuffer, Jens Emil Grønbæk, and Hans Gellersen. 2022. Gaze-Hand Alignment: Combining Eye Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality. Proc. ACM Hum.-Comput. Interact. 6, ETRA, Article 145 (May 2022), 18 pages.
9. Microsoft (2023). Eye-gaze-based interaction on HoloLens 2. https://learn.microsoft.com/en-us/windows/mixed-reality/design/eye-gaze-interaction. Accessed 3-11-24.
10. Sharon Oviatt. 1999. Ten myths of multimodal interaction. Commun. ACM 42, 11 (Nov. 1999), 74–81.
11. Padrao, G., Gonzalez-Franco, M., Sanchez-Vives, M. V., Slater, M., and Rodriguez-Fornells, A. (2016). Violating body movement semantics: Neural signatures of self-generated and external-generated errors. Neuroimage, 124, 147-156.
12. Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-touch: combining gaze with multi-touch for interaction on the same surface. In Proceedings of the 27th annual ACM symposium on User interface software and technology (UIST '14). ACM, New York, NY, USA, 509–518.
13. Ken Pfeuffer, Jason Alexander, Ming Ki Chong, Yanxia Zhang, and Hans Gellersen. 2015. Gaze-Shifting: Direct-Indirect Input with Pen and Touch Modulated by Gaze. In Proceedings of the 28th Annual ACM Symposium on User Interface Software &amp; Technology (UIST '15). ACM, New York, NY, USA, 373–383.
14. Ken Pfeuffer, Jason Alexander, and Hans Gellersen. 2016. Partially-indirect Bimanual Input with Gaze, Pen, and Touch for Pan, Zoom, and Ink Interaction. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16). ACM, New York, NY, USA, 2845–2856.
15. Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + pinch interaction in virtual reality. In Proceedings of the 5th Symposium on Spatial User Interaction (SUI '17). ACM, New York, NY, USA, 99–108.
16. Pfeuffer, K. (2017). Extending touch with eye gaze input. [Doctoral Thesis, Lancaster University]. Lancaster University.

17. Ken Pfeuffer, Jan Obernolte, Felix Dietz, Ville Mäkelä, Ludwig Sidenmark, Pavel Manakhov, Minna Pakanen, and Florian Alt. 2023. PalmGazer: Unimanual Eye-hand Menus in Augmented Reality. In Proceedings of the 2023 ACM Symposium on Spatial User Interaction (SUI '23). Association for Computing Machinery, New York, NY, USA, Article 10, 1–12.

18. Dennis Wolf, Jan Gugenheimer, Marco Combosch, and Enrico Rukzio. 2020. Understanding the Heisenberg Effect of Spatial Interaction: A Selection Induced Error for Spatially Tracked Input Devices. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). ACM, New York, NY, USA, 1–10.

19. Zhai, S., Kong, J., and Ren, X. (2004). Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. International journal of human-computer studies, 61(6), 823-856.

20. Zhai, S., Morimoto, C., and Ihde, S. (1999). Manual and Gaze Input Cascaded (MAGIC) Pointing. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 246–253). ACM.

21. 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 (CHI '14). ACM, New York, NY, USA, 1063–1072.

22. Uta Wagner, Mathias N. Lystbæk, Pavel Manakhov, Jens Emil Grønbæk, Ken Pfeuffer, and Hans Gellersen (2023). A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. ACM.

**Prof. Dr. Ken Pfeuffer** is an Assistant Professor in the Ubiquitous Computing and Interaction group at the computer science department at Aarhus University, Denmark. He leads the XI research group on topics of human-computer interaction, in particular AR/VR/XR, eye & hand interaction, UI design, and adaptive UI. Before, he was at Florian Alt's Usable Security and Privacy group in Munich, at Lancaster University in the UK at Hans Gellersen's interactive systems group. He also interned at both Microsoft and Google Research in US and received Best Paper Honorable Mention awards at ACM UIST and SUI.

**Prof. Dr. Hans Gellersen** is a Professor of Interactive Systems at Lancaster University in the UK, and at Aarhus University in Denmark. He received his PhD in computer science in 1996 from the University of Karlsruhe in Germany. Hans started his career in ubiquitous computing, with early work on computing in everyday objects, systems that blend physical and digital interaction, and techniques that facilitate cross-device interaction. His recent research interests include eye-tracking, gaze for interaction and multimodal interaction techniques that leverage eye movement in concert with other modalities. Hans was a founder of the Ubicomp conference series in 1999 and serves on the Editorial Board of ACM TOCHI. He has received Best Paper awards including from the ACM CHI and UIST conferences and the TOCHI journal. In 2021, he was awarded an ERC Advanced Grant by the European Research Council for his work on gaze and eye movement in interaction, and in 2022 he received a Humboldt Prize by the Alexander-von-Humboldt Foundation in recognition of his lifetime's research achievements.

**Dr. Mar Gonzalez-Franco** is a computer scientist and neuroscientist. She is currently a research manager at Google AR & VR where she works on creating a new generation of Immersive Tech. She leads the Blended Interactions Research and Devices team (BIRD), in their work they explore multi-device and multi-modal futures enabled by ML and AI. From foundational work to full products. Before that she was a principal researcher at Microsoft Research where she advanced fields like Haptics, Avatars, embodied interaction and Virtual Reality. She has open sourced some of the most used avatar libraries like the Microsoft Rocketbox, or the VALID avatars. Among her most significant recognitions are 2022 Times invention of the year, and the IEEE VGTC Virtual Reality new significant researcher award. Contact her at margonzalezfranco@gmail.com.