Promises and Design-Challenges of Mobile Transitional Interfaces

Daniel Schacht*
University of Lübeck

Hans-Christian Jetter†
University of Lübeck

ABSTRACT

While Transitional Interfaces (TIs) have primarily been studied in stationary settings, recent advancements in commercial technologies, such as head-mounted displays and smartphones, have made Mobile Transitional Interfaces (MobTIs) accessible to consumers. MobTIs differ from their stationary counterparts by providing users the ability to leave their tracking space and adapt to their surroundings. However, MobTIs also present challenges that require attention. The paper identifies three prominent challenges: the discoverability of MobTI content in the real world, the benefit-cost ratio of transitions, and the selection of appropriate input modalities. .

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Human-centered computing—Ubiquitous and mobile computing—Ubiquitous and mobile computing theory, concepts and paradigms—Mobile computing

1 INTRODUCTION

Transitional Interfaces (TIs) [4] represent an emerging class of user interfaces within the realm of mixed reality (MR) or cross-reality (XR) [8]. These interfaces allow users to navigate fluidly along the reality-virtuality continuum (RVC) [10] during their work. TIs hold the promise of bridging the gap between two distinct interaction paradigms: (1) interacting with objects and computers situated in the physical reality, such as using personal computers (PCs) or mobile devices, and (2) interacting within augmented and virtual realities (AR/VR), typically involving the use of head-mounted displays (HMDs) [8]. With TIs, users have the freedom to move seamlessly along the RVC, enabling them to individually select and switch between displays, input/output modalities, and representations of data or functionality that best suit their specific tasks at any given time [5, 8].

Grasset et al.'s concept and framework for TIs from 2006 defines this behavior as a transition between different *contexts* [6]. The definition of a context is defined by its position on the RV continuum (e.g., AR, VR, reality) but also by properties such as scale (e.g., macro, micro, nano in relation to the data space or virtual environment), representation (e.g., photorealistic, non-photorealistic, symbolic), or any other user parameters such as navigation mode (e.g. natural walking, teleporting).

1.1 Mobile Transitional Interfaces

Since their inception with Billinghurst et al.'s Magic Book [4] in the early 2000s, TIs have predominantly been studied in stationary settings. This is partly because the technical constraints of the hardware have restricted the use of Transitional Interfaces to lab environments and prevented their use in mobile settings. Consequently, previous investigations of TIs have primarily focused on tasks and interactions within professional environments (e.g., laboratories or offices) [3, 9, 14], or semi-public spaces (e.g., museums or libraries) [7], catering to specific goals and contexts of use. In contrast, mobile transitional interfaces (henceforth MobTIs) have the potential to offer users greater freedom, accommodating a wide range of tasks and contexts encountered in everyday activities. MobTIs differ from their stationary counterparts by providing users the ability to leave their tracking space beyond a confined area. This expansion unlocks numerous new opportunities, raises questions, and presents challenges that go beyond those of stationary TIs. The subsequent subsections discuss some of these aspects in detail.

The significance of MobTIs is underscored by recent advancements in commercial technologies. The emergence of head-mounted displays such as the Oculus Quest 3 and Apple Vision Pro, which eliminate the need for a predefined tracking space and seamlessly transition between AR and VR, has made mobile TIs accessible to consumers. By combining these devices, which cover the AR to VR range of Milgram's RVC, with smartphones and other wearables, MobTIs can cater to a wide array of real-world use-cases.

1.2 Examples of Mobile Transitional Interfaces

To further explain why this concept might bring benefits to users, we outline some exemplary scenarios where one might use a MobTI in the following.

Scenario 1. Furniture Shopping: Imagine that you want to go to a furniture store and purchase the perfect couch for your new apartment. The initial step involves browsing the website of the furniture store on your personal computer (PC) to explore the available couch options that pique your interest. You identify the ones that catch your attention and utilize the AR feature of your HMD to visualize how they would appear in your living room. Simultaneously, the HMD scans and captures your living room for future reference.

The following day, you physically visit a furniture store to personally test the comfort of the selected couches. To conveniently locate your preferred options, you employ the AR function on your HMD to guide you through the furniture store directly to the couches that have captured your interest. After thoroughly evaluating the coziness of each couch, you decide to perform a final assessment of how the chosen couch would complement your living room and whether opting for a leg extension would be advantageous. For this purpose, you leverage the VR function of your HMD to transport yourself to an immersive virtual representation of your living room, allowing you to experiment with different couch alignments and configurations.

Scenario 2. Mobile Cinema: In another scenario, imagine watching a movie while traveling on a train. Watching the Movie in VR would enable a cinema-like experience. Unlike the previous example where all interactions were tied to a specific goal and part of a larger process, the transition between different contexts can also offer opportunistic advantages, even in the absence of a clearly defined objective or process. For example, you could use your mobile phone to present your ticket to the conductor and browse the Netflix app to discover the perfect movie. To enhance your viewing experience, you can engage in the VR mode of your HMD instead of watching it on the small screen of your phone. With VR, you can switch to a more immersive environment.

Moreover, VR provides the added benefit of being able to block out distractions, such as a group of noisy passengers located at the

*d.schacht@uni-luebeck.de
†hanschristian.jetter@uni-luebeck.de

end of your train compartment. If someone enters your personal space, the VR environment can seamlessly transition to AR, allowing you to continue watching the movie while remaining aware of the conductor's arrival for ticket verification or similar social interactions.

2 Promises of mobile Transitional Interfaces

A key promise of MobTIs and advantage over traditional mobile applications is based on the very idea of cross-reality interfaces themselves: They are intended to "provide users with optimal visual and algorithmic support with maximum cognitive and perceptual suitability, depending on their current tasks and needs" [14].

In a mobile context, the user's ever-changing social and physical environment affect strongly what the "current tasks and needs" are and what "optimal support" or "maximum cognitive and perceptual suitability" mean. Therefore, the much greater choice of tools and representations across the entire RV-continuum of a MobTI is an important promise for improving future human-computer interactions.

However, in a mobile setting, TIs need to incorporate adaptation and incorporation of the changing surroundings to achieve an optimal context for the user. This adaptation of an application to its surrounding has been studied for a long time, starting with Ubiquitous Computing [16] in general and, more specifically, Context-Aware Computing [15]. In the future, some of the already existing taxonomies for the behavior of applications in changing surroundings could be useful to describe and design user and system behaviors for MobTI applications. Examples of this would be the Context-Aware Software Dimensions [15] (i.e. proximate selection, contextual informations, contextual commands, automatic contextual reconfiguration, and context-triggered actions) or Generic Contextual Capabilities [11] (i.e. contextual sensing, contextual adaptation, contextual resource discovery, and contextual augmentation).

The concept of *context* from Ubiquitous Computing [2] and Context-Aware Computing [1, 15] provide thoughtful overviews of environmental characteristics that could influence the optimal adaptation of a system. However, a complication arises from the fact that both the research field of TIs and the research field of context-aware computing have defined *context* differently. One possible solution to address this problem would be to refer to the surrounding of the user as "environment" and rename all taxonomies accordingly (e.g., "contextual adaptation" would be renamed to "environmental adaptation"). Another option would be to integrate *context* in the sense of Ubiquitous Computing into the *context* of TIs. Grasset deliberately formulated his definition of *context* for TIs in a broad manner, stating that "A context not only defines a space (e.g. AR, VR, Reality), but can define a scale (e.g. macro, micro, nano, etc.), a representation (e.g. photorealistic, nonphotorealistic, symbolic), and any other user parameters (such as viewpoints and navigation mode). A context is the collection of values of parameters relevant to the application." [6]. Therefore, it is plausible to augment the TI context definition by incorporating the context definition from the domain of ubiquitous computing, without contradicting the original definition.

3 Challenges of Mobile Transitional Interfaces

Alongside the discussed benefits associated with MobTIs, numerous new design challenges emerge that we need to solve before MobTIs can provide real-world value to users. The following points represent some of the most prominent challenges. It is important to note that this list is not intended to be exhaustive. The subsequent subsections succinctly introduce three main problems and pose questions that should be addressed in future research.

3.1 Challenge 1: Benefit-Cost-Ratio of Transitions

The central hypothesis underlying this challenge posits that a transition must offer a certain level of benefit to users, prompting them to make the effort to switch actively to another context.

Each transition from one context to another requires effort, which may stem from the time it takes to transition, the physical effort involved in switching devices, the cognitive load of adapting to different interfaces and information representations, or a combination thereof. For a user, a transition is only useful, if it offers enough benefit in comparison to its necessary effort, e.g., a much more improved user experience, a more immersive interface, or an enhanced spatial understanding of the surroundings.

This cost-benefit ratio affects all TIs, but it might play a more significant role in MobTIs, as the benefits of a transition may not be immediately evident in a new application. This assumption is based on the premise that MobTI users tend to explore a wider variety of new TI applications compared to users of stationary TIs. Furthermore, stationary TIs often have explicitly defined goals that required the utilization of specific contexts.

To facilitate the development of TIs where users can freely choose their contexts, a deeper understanding of the relationship between the costs and benefits of transitions is required. The following points raise additional specific questions related to this topic:

1. What are the most effective ways to communicate the benefits of a transition to users, enabling them to make informed decisions about utilizing alternative contexts?
2. Can users potentially misjudge the benefits of certain contexts due to a legacy bias, leading to fewer transitions [12]?
3. Is it possible to measure the average effort required for a particular transition?
4. Which variables influence the effort associated with a specific transition?
   (a) Personal characteristics (e.g., size, age, gender, technology affinity)
   (b) Environmental characteristics (e.g., available space, indoor or outdoor, private or public, weather conditions)
   (c) The nature of the transition (e.g., device switch, interface change, animated visual transitions within a single device [13])

3.2 Challenge 2: Discoverability of Mobile TI Content

One challenge posed by MobTIs is the discoverability of TI content in the real world. While there may be use cases where users are aware of when and where to switch between contexts to accomplish their goals (e.g., examples 1 and 2), there could also be location-based applications in the user's vicinity that they are unaware of. In such cases, it may be appropriate to provide the user with a notification on their mobile phone, assuming they are carrying their head-mounted display (HMD) in a backpack. This notification would alert them to the presence of potential interactions nearby for which taking out and wearing the HMD could prove worthwhile.

This challenge raises questions regarding when and how to inform the user in order to strike the right balance between providing useful information and avoiding excessive disruptions. The following points highlight some further detailed questions related to this topic:

1. Should the user receive an alert on their phone when passing by a location-based MobTI interaction?
2. Should the user only receive alerts if the applications are associated with apps they have installed on their phone or websites they have used on their PC?
3. Should these alerts only be displayed if the user is actively using the corresponding app while within a certain range of the location?

4. Should users have the option to customize which services are allowed to send them alerts?
5. (a) If users can only receive alerts from apps they have already installed, the discoverability of new MobTI interactions may be significantly diminished.
   (b) However, receiving notifications from every application could lead to information overload.
6. Furthermore, the question arises as to how the discovery of MobTI applications would appear while the user is in a context wearing an HMD.

3.3 Challenge 3: Input Modalities for Mobile Transitional Interfaces

Mobile applications may require different input modalities compared to lab-based systems, depending on the specific use cases. Some commonly used inputs for head-mounted displays (HMDs), such as controllers or hand gestures, may not be suitable for certain situations due to various reasons.

One reason is that certain forms of interaction might not be socially acceptable or could lead to feelings of awkwardness. For instance, users may feel uncomfortable using mid-air gestures in crowded places due to social norms.

Another reason is that users may want to interact with their environment while simultaneously using the application. If a user is using their input abilities to interact with an application, they would not be able to use those same abilities to interact with their surroundings. For example, a user cannot use voice control while having a conversation or use gesture control while eating ice cream.

This challenge impacts all TIs to some extent, but it may have a more pronounced effect on MobTIs. In stationary TIs, the focus is typically on the TI application itself rather than additional tasks or content in the surrounding environment. However, due to the highly interactive nature of the environment in MobTIs, users can encounter various situations where they engage in additional interactions with their surroundings alongside the application.

The following points raise further specific questions related to this topic:

1. Should users themselves have the authority to decide which form of interaction is most appropriate for the environment?
2. Should the application analyze the environment and make decisions on behalf of the user?
3. Can multi-modal interactions help address these challenges?
4. Which environmental characteristics should influence the choice of input modality?

4 Conclusion

This paper provided some initial thoughts on the new class of cross-reality systems called MobTIs and how they are different from "traditional" TI systems that are primarily based in laboratory or home settings.

One of the key promises of MobTIs is their ability to adapt to the surrounding environment. To achieve this, researchers can build upon the existing scientific groundwork in the areas of Augmented Reality (AR), Virtual Reality (VR), and TIs. Additionally, they can draw upon the knowledge and content from the fields of Ubiquitous Computing and Context-Aware Computing. However, MobTIs also encounter challenges that must be addressed to maximize their value. These challenges encompass the (1) discoverability of MobTI content in the real world, the assessment of (2) the benefit-cost ratio associated with transitioning between contexts, and (3) the selection of appropriate input modalities. Each challenge poses specific questions that warrant exploration in future research, such as effectively informing users about nearby MobTI interactions, efficiently communicating the advantages of transitions, and determining suitable input modalities based on environmental characteristics.

References

[1] G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles. Towards a Better Understanding of Context and Context-Awareness. In G. Goos, J. Hartmanis, J. Van Leeuwen, and H.-W. Gellersen, eds., *Handheld and Ubiquitous Computing*, vol. 1707, pp. 304–307. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999. Series Title: Lecture Notes in Computer Science. doi: 10.1007/3-540-48157-5_29

[2] G. D. Abowd and E. D. Mynatt. Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction*, 7(1):29–58, Mar. 2000. doi: 10.1145/344949.344988

[3] H. Bai, P. Sasikumar, J. Yang, and M. Billinghurst. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pp. 1–13. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376550

[4] M. Billinghurst, H. Kato, and I. Poupyrev. The magic book: A transitional AR interface. *Computers & Graphics*, 25(5):745–753, 2001. doi: 10.1016/S0097-8493(01)00117-0

[5] B. Fröhler, C. Anthes, Christoph, F. Pointecker, J. Friedl, D. Schwajda, A. Riegler, Andreas, S. Tripathi, C. Holzmann, M. Brunner, H. Jodlbauer, H.-C. Jetter, and C. Heinzl. A Survey on Cross-Virtuality Analytics. *Computer Graphics Forum*, pp. 465–494, 2022. doi: 10.1111/cgf.14447

[6] R. Grasset, J. Looser, and M. Billinghurst. Transitional interface: concept, issues and framework. In *Proceedings of the 5th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '06, pp. 231–232. IEEE Computer Society, USA, 2006. doi: 10.1109/ismar.2006.297819

[7] R. Hammady, M. Ma, and C. Strathearn. Ambient Information Visualisation and Visitors' Technology Acceptance of Mixed Reality in Museums. *Journal on Computing and Cultural Heritage*, 13(2):1–22, June 2020. doi: 10.1145/3359590

[8] H.-C. Jetter, J.-H. Schröder, J. Gugenheimer, M. Billinghurst, C. Anthes, M. Khamis, and T. Feuchtner. Transitional Interfaces in Mixed and Cross-Reality: A new frontier? In *Interactive Surfaces and Spaces*, ISS '21, pp. 46–49. Association for Computing Machinery, New York, NY, USA, Nov. 2021. doi: 10.1145/3447932.3487940

[9] R. Langner, M. Satkowski, W. Büschel, and R. Dachselt. MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17. ACM, Yokohama Japan, May 2021. doi: 10.1145/3411764.3445593

[10] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994. Publisher: The Institute of Electronics, Information and Communication Engineers.

[11] J. Pascoe. Adding generic contextual capabilities to wearable computers. In *Digest of Papers. Second International Symposium on Wearable Computers (Cat. No.98EX215)*, pp. 92–99. IEEE Comput. Soc, Pittsburgh, PA, USA, 1998. doi: 10.1109/ISWC.1998.729534

[12] T. Plank, H.-C. Jetter, R. Rädle, C. N. Klokmose, T. Luger, and H. Reiterer. Is Two Enough?!: Studying Benefits, Barriers, and Biases of Multi-Tablet Use for Collaborative Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pp. 4548–4560. ACM Press, Denver, Colorado, USA, 2017. doi: 10.1145/3025453.3025537

[13] F. Pointecker, J. Friedl, D. Schwajda, H.-C. Jetter, and C. Anthes. Bridging the Gap Across Realities: Visual Transitions Between Virtual and Augmented Reality. In *2022 IEEE Interational Symposium on*

Mixed and Augmented Reality (ISMAR). Singapore, Oct. 2022. (to appear).

[14] A. Riegler, C. Anthes, H.-C. Jetter, C. Heinzl, C. Holzmann, J. Herbert, M. Brunner, S. Auer, J. Friedl, and B. Fröhler. Cross-Virtuality Visualization, Interaction and Collaboration. In *International Workshop on Cross-Reality (XR) Interaction co-located with 14th ACM International Conference on Interactive Surfaces and Spaces*, 2020.

[15] B. Schilit, N. Adams, and R. Want. Context-Aware Computing Applications. In *1994 First Workshop on Mobile Computing Systems and Applications*, pp. 85–90. IEEE, Santa Cruz, California, USA, Dec. 1994. doi: 10.1109/WMCSA.1994.16

[16] M. Weiser. The Computer for the 21 st Century. *Scientific American*, 265(3):94–105, 1991. Publisher: Scientific American, a division of Nature America, Inc.