A Generic Architecture for Cross-Virtuality

FABIAN POINTECKER, DANIEL SCHWAJDA, DOMINIK LIST, and CHRISTOPH ANTHES, University of Applied Sciences Upper Austria, Austria

In this position paper, we provide an overview of a generic high-level architecture for cross-virtuality. The architecture is structured into six subcomponents in which we focus on the unique features and design considerations that are required for cross-virtuality. We hope that the architecture provides a basic foundation for future work in this area and that potential pitfalls can be avoided in advance.

CCS Concepts: • **Human-centered computing** → *Mixed / augmented reality*.

Additional Key Words and Phrases: cross-virtuality, mixed reality, reality-virtuality continuum, high-level architecture

ACM Reference Format:
Fabian Pointecker, Daniel Schwajda, Dominik List, and Christoph Anthes. 2021. A Generic Architecture for Cross-Virtuality. In *ISS'21 Workshop Proceedings: "Transitional Interfaces in Mixed and Cross-Reality: A new frontier?", November 14, 2021, Łódź, Poland*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.18148/kops/352-2-1l3l2wncu3tgb5

1 INTRODUCTION

With advancements in head-mounted display (HMD) technology and the recent rise of augmented (AR) and virtual reality (VR) hardware, transitioning between different areas of Milgrams Reality-Virtuality continuum (RVC) [16] has become possible [22]. This new research field, as a sub-area of Transitional Interfaces (TIs), is known as cross-virtuality (XV) [26]. Although each of the different stages on the RVC has already gained significant attention in the various research communities, moving between different stages comes with additional requirements. In this publication we give an overview of a high-level architecture of such XV systems considering the unique requirements and characteristics of XV:

- Integration and support of a transition between different stages on the RVC.
- The use of multiple input and output devices simultaneously or in combination.
- Tracking of multiple users, devices and surrounding objects in a single environment.
- Network communication to interconnect different devices, provide a common data model and a shared virtual environment.

2 ARCHITECTURE

To sketch a generic architecture for XV, we extended the Sensory Feedback Loop [19] with XV specific characteristics. The main features were classified into components and are outlined in Figure 1. Within the Sensory Feedback Loop the user can utilise one or more devices simultaneously or switch between them. In the following section the components are described in more detail.

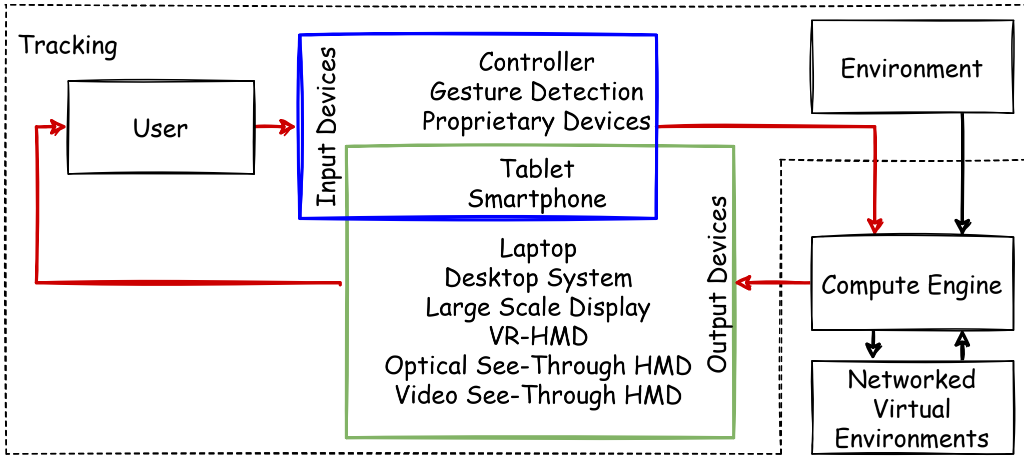

Fig. 1. Generic Architecture for XV with Sensory Feedback Loop [19] highlighted in red.

2.1 Tracking and Environment

The tracking component is used to detect the position and orientation in six degrees of freedom (6-DoF) of various entities in the physical environment. Tracking information is shared between those entities so that they can be spatially registered within XV. This enhances safety for fully immersed experiences through collision reduction, adds spatially aware interaction methods [24] and provides additional opportunities for visualisations (e.g. extending displays in physical space [25]). Possible entities include various input or output devices, the user themselves, or objects in the tracked environment.

To incorporate the physical environment into the XV application, the environment needs to be reconstructed or statically modelled [30]. This can be achieved with mobile 6-DoF tracking modules which are attached to physical objects such as tables and chairs to determine their spatial position. This allows for passive haptic feedback [23] and enables object occlusion at runtime. A more detailed reconstruction of the surroundings could be achieved by depth cameras to procedurally create a virtual replica of the physical environment [31] which is continuously updated.

2.2 Output Devices

In order to support XV on different stages on the RVC, we have identified a variety of possible visual output devices. These device types as well as their position and range can be seen in Figure 2. They either support more than one stage on the RVC or can be combined with each other to provide benefits of each technology.

- *Laptop, Desktop System and Large-Scale Display:* With this type of devices, the content is presented on monoscopic 2D displays in various sizes, functionalities and designs. Mobile devices such as laptops can be flexibly positioned in the room, depending on the application. Desktop systems can be used for computationally intensive simulations or visualisations. Large-scale displays support collaborative decision-making for multiple users [17]. If multi-touch is also available, it can be classified as a combined input and output device that allows users to collaborate simultaneously. In combination with AR, these devices can be used as augmented displays, allowing non-stereoscopic, interactive surfaces to be enhanced with computer generated content [1, 13, 25].

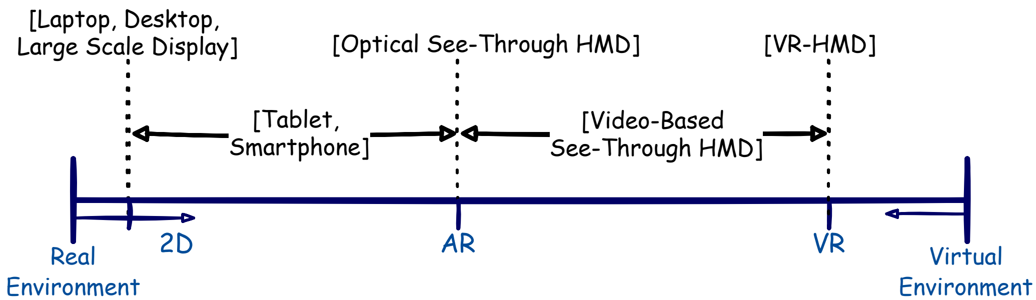

Fig. 2. Devices and device types for XV, arranged according to their position and range on the RVC.

- *Optical See-Through HMD:* These devices use a semi-transparent surface to visually combine computer-generated content with a view from the physical environment. They can only be used for AR as they are not able to isolate the user compared to conventional VR HMDs. But due to the unrestricted view of real content, cross-device interaction is possible. This enables the user to work with other displays or use other input devices at any time without having to put down the HMD. However, optical see-through HMDs still have to address several issues, such as, a restricted field of view, out-of-place augmentations, vergence-accommodation conflict and occlusion problems [4, 5, 27].
- *Video-Based See-Through HMD:* Instead of external sensors, several manufacturers of VR HMDs have incorporated wide-angle cameras into the headset to spatially track the headset and controllers. In recent years, the concept of forward-facing cameras has been expanded to allow for a video-based see-through AR mode in addition to the traditional VR mode. This allows the user to move along the RVC without changing the headset. Contrary to optical see-through devices, the physical environment is captured by cameras which leads to several advantages and disadvantages [27]. For example, video-based systems can block on a pixel-by-pixel basis and can therefore deal with occlusions more easily. They are able to perform contrast matching and can provide a higher alignment precision when augmenting the physical world in a video stream. However, the additional video processing can result in increased display latency, which could lead to cybersickness. In addition, the perceived resolution of the physical scene may be limited by the resolution of the video cameras or viewing optics. Furthermore, viewing angles need to match the viewing angles of the user's eyes in order to avoid a spatial shift. In contrast, optical systems usually have an unobstructed vision of the physical world, which means that there is no perspective conversion of the viewing angle.
- *VR HMD:* By combining VR HMDs with other devices, a remote or co-located collaboration across the RVC is possible through a shared virtual environment [3, 12] and a synchronised data layer. Piumsomboon et al. [21] combine a VR HMD with an optical see-through HMD and Grandi et al. [8] combine a VR HMD with a tablet device for handheld AR.

2.3 Combined Devices

Handheld devices such as smartphones and tablets can be used as output and input devices in an XV environment. Mobile devices can display content flexibly in space or directly within the user's reach and the touch-sensitive surface is commonly used as input, e.g. keyboard with passive haptic feedback in VR. Due to the availability of a front-facing camera on most devices, it is possible to perform handheld AR, enabling seamless switching between 2D display and AR. Out of the box, these devices usually only support 3-DoF tracking. 6-DoF can be achieved with additional tracking data from the XV tracking area. For example, with marker-based tracking or via mobile trackers which are attached to the handheld device, enabling spatial tracking. This results in additional input possibilities, such as gesture control via the position and orientation of the handheld device [11, 15].

2.4 Input Devices

Input devices are typically designed for a specific stage on the RVC, and it is therefore important to consider how input modalities change as the user moves across the RVC. To avoid switching input devices at every transition, it is advisable that devices can be used on more than one stage of the RVC. Traditional multi-touch input on handheld devices or large-scaled displays can be used as planar input in 2D and AR [7, 28]. Since spatial position of input devices should be available in XV, it is possible to use multi-touch input in VR via virtual replica [32]. Spatial input can be provided by hand controllers, hand tracking or by detecting and interpreting spatial events to provide a set of supported mid-air gestures [14, 33]. For special use cases, proprietary and tangible [9, 10] devices can be implemented in XV, e.g. the Cubic Mouse [6].

2.5 User

The user accesses different input devices to interact with the application or environment. Depending on the output device, the user can either move along the RVC or use several different devices to cover multiple locations on the RVC.

For collaboration and self-awareness in XR, a user representation is important. A corresponding model can be reconstructed via inverse kinematics [2, 18] or depth sensors [29]. This is archived by accessing the tracking data of the used devices. Depending on the use case and the available tracking data, the entire user, parts of the user or certain awareness cues can be represented. Due to HMDs limited field of view, awareness cues can be especially useful in VR and AR to indicate nearby users and their viewing frustum [20].

2.6 Compute Engine

The requirements for the compute engine are highly varying, as they mainly depend on the available devices and their usage. With regard to XV, particular attention should be paid to the implementation and communication of the various devices with each other. The following tasks should be considered, data sharing, user management, device management, synchronising tracking data and rendering.

2.7 Networked Virtual Environments

To connect multiple users at different stages on the RVC, a networked virtual environment (NVE) is crucial and provides the basic interface to exchange information. Traditional AR or VR installations use NVEs for collaboration, but in XV it is additionally necessary for the communication between different output and input devices and the exchange of shared data and interactions. Common goals in NVEs are, multi-modal communication, data sharing, responsiveness and consistency.

3 ACKNOWLEDGMENTS

This publication is a part of the X-PRO project. The project X-PRO is financed by research subsidies granted by the government of Upper Austria.

REFERENCES

[1] Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs, and Harald Reiterer. 2018. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, Montreal QC, Canada, 1–12.

[2] Polona Caserman, Philipp Achenbach, and Stefan Göbel. 2019. Analysis of inverse kinematics solutions for full-body reconstruction in virtual reality. In *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 1–8.

[3] Marco Cavallo, Mishal Dholakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Dataspace: A reconfigurable hybrid reality environment for collaborative information analysis. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 145–153.

[4] Sara Condino, Marina Carbone, Roberta Piazza, Mauro Ferrari, and Vincenzo Ferrari. 2019. Perceptual limits of optical see-through visors for augmented reality guidance of manual tasks. *IEEE Transactions on Biomedical Engineering* 67, 2 (2019), 411–419.

[5] David Drascic and Paul Milgram. 1996. Perceptual issues in augmented reality. In *Stereoscopic displays and virtual reality systems III*, Vol. 2653. International Society for Optics and Photonics, 123–134.

[6] Bernd Fröhlich and John Plate. 2000. The cubic mouse: a new device for three-dimensional input. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. Association for Computing Machinery, New York, NY, USA, 526–531.

[7] Stéphane Gosset, Mickael Sereno, Lonni Besançon, and Tobias Isenberg. 2020. Tangible Volumetric Brushing in Augmented Reality. In *IEEE VIS Posters*. Posters of IEEE Visualization.

[8] Jerônimo Gustavo Grandi, Henrique Galvan Debarba, and Anderson Maciel. 2019. Characterizing asymmetric collaborative interactions in virtual and augmented realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 127–135.

[9] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. 234–241.

[10] Paul Issartel, Florimond Guéniat, and Mehdi Ammi. 2014. A portable interface for tangible exploration of volumetric data. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*. Association for Computing Machinery, New York, NY, USA, 209–210.

[11] Paul Issartel, Florimond Guéniat, and Mehdi Ammi. 2014. Slicing techniques for handheld augmented reality. In *2014 IEEE symposium on 3D user interfaces (3DUI)*. IEEE, 39–42.

[12] Kiyoshi Kiyokawa, Haruo Takemura, and Naokazu Yokoya. 1999. A collaboration support technique by integrating a shared virtual reality and a shared augmented reality. In *IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028)*, Vol. 6. IEEE, 48–53.

[13] Tahir Mahmood, Erik Butler, Nicholas Davis, Jian Huang, and Aidong Lu. 2018. Building multiple coordinated spaces for effective immersive analytics through distributed cognition. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*. IEEE, 1–11.

[14] Ali M Malkawi and Ravi S Srinivasan. 2005. A new paradigm for Human-Building Interaction: the use of CFD and Augmented Reality. *Automation in construction* 14, 1 (2005), 71–84.

[15] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. 2012. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 13–22.

[16] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.

[17] Dorit Nevo, Saggi Nevo, Nanda Kumar, Jonas Braasch, and Kusum Mathews. 2015. Enhancing the visualization of big data to support collaborative decision-making. In *2015 48th Hawaii International Conference on System Sciences*. IEEE, 121–130.

[18] Mathias Parger, Joerg H Mueller, Dieter Schmalstieg, and Markus Steinberger. 2018. Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. 1–10.

[19] Ken Pimentel and Kevin Teixeira. 1993. Virtual reality through the new looking glass. (1993).

[20] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. 2019. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI* 6 (2019), 5.

[21] Thammathip Piumsomboon, Youngho Lee, Gun Lee, and Mark Billinghurst. 2017. CoVAR: a collaborative virtual and augmented reality system for remote collaboration. In *SIGGRAPH Asia 2017 Emerging Technologies*. 1–2.

[22] Fabian Pointecker, Hans-Christian Jetter, and Christoph. Anthes. 2020. Exploration of Visual Transitions Between Virtual and Augmented Reality. In *Workshop on Immersive Analytics: Envisioning Future Productivity for Immersive Analytics // @CHI 2020 Honolulu*.

[23] Fabian Pointecker, Hans-Christian Jetter, and Christoph. Anthes. 2020. Visualising Geospatial Network Data using Virtual Reality Technology. In *Workshop on Immersive Analytics: Envisioning Future Productivity for Immersive Analytics // @CHI 2020 Honolulu*.

[24] Roman Rädle, Hans-Christian Jetter, Mario Schreiner, Zhihao Lu, Harald Reiterer, and Yvonne Rogers. 2015. Spatially-aware or spatially-agnostic? Elicitation and evaluation of user-defined cross-device interactions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 3913–3922.

[25] Patrick Reipschläger and Raimund Dachselt. 2019. DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays. In *International Conference on Interactive Surfaces and Spaces (ISS)*. ACM, New York, NY, USA, 29–41.

[26] Andreas Riegler, Christoph Anthes, Hans-Christian Jetter, Christoph Heinzl, Clemens Holzmann, Herbert Jodlbauer, Manuel Brunner, Stefan Auer, Judith Friedl, Bernhard Fröhler, Christina Leitner, Fabian Pointecker, Daniel Schwajda, and Shailesh Tripathi. 2020. Cross-Virtuality Visualization,

Interaction and Collaboration.. In *XR@ ISS*.

[27] Jannick P. Rolland, Richard L. Holloway, and Henry Fuchs. 1995. Comparison of Optical and Video See-through, Head-Mounted Displays. In *Photonics for Industrial Applications*, Hari Das (Ed.). International Society for Optics and Photonics, SPIE, Boston, MA, 293–307.

[28] Mickael Sereno, Lonni Besançon, and Tobias Isenberg. 2019. Supporting volumetric data visualization and analysis by combining augmented reality visuals with multi-touch input. In *EuroVis Posters*. 21–23.

[29] Ari Shapiro, Andrew Feng, Ruizhe Wang, Hao Li, Mark Bolas, Gerard Medioni, and Evan Suma. 2014. Rapid avatar capture and simulation using commodity depth sensors. *Computer Animation and Virtual Worlds* 25, 3-4 (2014), 201–211.

[30] 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*. Association for Computing Machinery, New York, NY, USA, 3307–3316.

[31] Misha Sra, Sergio Garrido-Jurado, Chris Schmandt, and Pattie Maes. 2016. Procedurally generated virtual reality from 3D reconstructed physical space. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*. Association for Computing Machinery, New York, NY, USA, 191–200.

[32] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. Tabletinvr: Exploring the design space for using a multi-touch tablet in virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13.

[33] Rensu P Theart, Ben Loos, and Thomas R Niesler. 2017. Virtual reality assisted microscopy data visualization and colocalization analysis. *BMC bioinformatics* 18, 2 (2017), 1–16.