A Survey on Cross-Virtuality Analytics


1 University of Applied Sciences Upper Austria, Wels, Austria
2 Institute for Multimedia and Interactive Systems, University of Lübeck, Lübeck, Germany

Abstract

Cross-virtuality analytics (XVA) is a novel field of research within immersive analytics and visual analytics. A broad range of heterogeneous devices across the reality–virtuality continuum, along with respective visual metaphors and analysis techniques, are currently becoming available. The goal of XVA is to enable visual analytics that use transitional and collaborative interfaces to seamlessly integrate different devices and support multiple users. In this work, we take a closer look at XVA and analyse the existing body of work for an overview of its current state. We classify the related literature regarding ways of establishing cross-virtuality by interconnecting different stages in the reality–virtuality continuum, as well as techniques for transitioning and collaborating between the different stages. We provide insights into visualization and interaction techniques employed in current XVA systems. We report on ways of evaluating such systems, and analyse the domains where such systems are becoming available. Finally, we discuss open challenges in XVA, giving directions for future research.

Keywords: cross-virtuality, cross-virtuality analytics, reality–virtuality continuum, transitional interfaces

CCS Concepts: Human-centred computing → Visualization; Mixed/augmented reality; Virtual reality; Visualization design and evaluation methods

1. Introduction

Data analysis and visualization along the reality–virtuality continuum (RVC), as defined by Milgram and Kishino [MTUK95], ranges from conventional workstation-based visual analytics (VA) on 2D screens to 3D or immersive visualizations that employ augmented reality (AR), augmented virtuality (AV) and virtual reality (VR). Moving beyond traditional visual representations on 2D screens and enabling embodied ways of navigating and interacting with visualizations in 3D space is of increasing interest in many application areas such as aeronautics [UKAG19], production [BV19], education [AZPT17] and cultural heritage [BPF*18].

Within this space, cross-virtuality analytics (XVA) is a novel field of research, concerned with systems for data visualization and analysis that seamlessly integrate different visual metaphors and devices along the entire RVC to support multiple users with transitional and collaborative interfaces. Two recent workshops in this field [SKE*20, JSG*21] emphasize the importance and relevance of the topic. According to Riegler et al. [RAJ*20], XVA enables a seamless integration and transition between conventional 2D visualization, AR and VR. Its goal is to provide users with optimal visual and algorithmic support with maximum cognitive and perceptual suitability, depending on their current tasks and needs in the analysis process. Riegler et al. define XVA as a novel possibility for interactive visualization based on fluent transitions across the RVC.

Due to the proliferation of affordable, diverse display hardware, such as head-mounted VR and AR devices, a key aspect of XVA is the possibility to immerse into or view data using different visual representations and devices depending on the task at hand. Using different devices enables XVA to cover the entire spectrum of mixed reality (MR) in the sense of Milgram and Kishino [MK94], with different displays providing what we call specific stages in the RVC (such as AR, AV or VR). The usage of the term cross-virtuality (XV) instead of the widely used but ambiguously defined terms cross-reality (XR) and extended reality (see, for example, Wang et al. [WAM20], Çöltekin et al. [ÇLM*20] and Speicher et al. [SHN19]) emphasizes the following observation: Especially the early phases of the data analytics process are best supported with RVC stages closer to the virtuality side, such as VR or AV. The integration of AR and other stages closer to the reality side can be powerful additions to create an entirely new kind of visual analysis tools.
In our understanding, XVA is fully contained within the field of VA, which according to Keim et al. [KAF*08] ‘combines automated analysis techniques with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex data sets’. Additionally, XVA is also firmly embedded within the space of immersive analytics (IA), which according to Chandler et al. [CCC+15] seeks ways for ‘immersing people in their data’. IA aims at using ‘engaging, embodied analysis tools to support data understanding and decision making’, and is defined as not being tied to the use of specific techniques [MSD*18]. In contrast to these large areas covered by VA and IA, XVA specifically focuses on VA systems for IA with transitional interfaces; that is, interfaces allowing users to simultaneously interact in multiple types of spaces (AR, VR) and transition between these different contexts [GLB06] or stages. Similarly, hybrid virtual environments (HVE), defined by Wang and Lindeman as systems ‘which incorporate multiple and complementary virtual and/or physical interface elements appropriate for a set of tasks’[WL14], can be considered a superset of XVA, with XVA specifically focusing on VA applications and XV interfaces. Furthermore, working along stages in the RVC can involve one or multiple users transitioning between these stages including situations in which some users work simultaneously towards a common goal, but can be situated in different stages. Here, fields such as computer-supported cooperative work (CSCW) [Gru94], collaborative visual analytics (CVA) [HA07], collaborative immersive analytics (CIA) or collaborative virtual environments (CVE) [BBF95] can provide input to XVA. It is noteworthy that even though collaboration is of key importance in XVA, it is not a prerequisite. XVA can also enhance single-user analysis through transitional interfaces.

In this work, we provide an overview of the existing body of work in XVA at the intersection of VA, IA, CVA and transitional interfaces. The main contributions of our work are:

1. a review and classification of existing VA methods and systems for enabling analysis on devices across the RVC,
2. an analysis of the XVA transition and collaboration concepts used to link devices from across the RVC,
3. a systematic overview of the visualization and interaction techniques employed for XVA,
4. a review of existing studies on the effectiveness of these methods, and finally,
5. the identification of open challenges to guide future research endeavours in this area.

There is a wealth of recent surveys in neighbouring areas to XVA that reveal the importance and timeliness of the topic. Our work here differs in key aspects from these previous analyses. The recent survey on IA by Fonnet et al. [FP21] as well as the grand challenges in IA collected in a workshop by Ens et al. [EBC*21], provided a good starting point for our analysis, with XVA being a subtopic of IA. Our view in this survey is, however, focused on the aspects specific to XVA, such as ways for interconnecting stages on the RVC, as well as transitional interfaces. In their review on extended reality in spatial sciences, Çöltekin et al. [ÇLM*20] look at systems regarding technology, design and human factors. They focus mainly on geospatial research, whereas we address a broad range of VA domains. While they shortly discuss collaboration and the necessity of designing and implementing more collaborative systems in the future, they do not address transitional interfaces or cross-device aspects. The recent survey of collaborative work in AR by Sereno et al. [SWB*20] covers any collaborative scenario, but limits itself to AR devices. Therefore, it also does not cover cross-device aspects, while we specialize on VA systems on devices across the whole RVC. The report on the state of the art of spatial interfaces for 3D visualization by Besançon et al. [BYK*21] also provides an interesting view on interaction concepts across a wide variety of devices used in 3D visualization scenarios. In contrast to our work, it does not put a focus on cross-device aspects or collaboration.

We start our survey by motivating the research into XVA in Section 2. We describe the process we followed in our survey in Section 3. Few systems are already fully embracing the XVA definition, so the main part of our survey is dedicated to analysing individual aspects of XVA: Section 4 explores the ways how devices at different stages in the RVC can interconnect, which we refer to as levels of XVA. Section 5 elaborates on techniques for transitioning between different stages in the RVC. Section 6 explores the collaboration aspect in XVA systems. Our analysis of input data, visualization metaphors and visual interaction techniques employed in XVA systems is described in Section 7. Considerations regarding the evaluation of XVA systems can be found in Section 8. Section 9 provides an analysis of the domains and application areas in which such systems currently are most likely to be found. Finally, in Section 10, we list the most important research challenges in XVA, which we identified in our literature review.

2. Motivation

The first systems, which can be considered as XVA, are the CAVERN architecture for CVE by Leigh et al. [LJD97] from 1997 and the Studierstube environment for multi-user AR by Szalavari et al. [SSFG98] from 1998. Research on XVA gained momentum recently due to the increasing availability of respective consumer-grade hardware, such as handheld devices as well as head-mounted devices (HMD) for AR and VR. XVA systems and aspects thereof are now found in highly diverse application areas at different stages of maturity. Examples of domains for which XVA techniques have been introduced, or for which XVA techniques will be of interest in the near future, include (for more details see Section 9):

1. Production and supply chain: visualization for design, assembly and testing by Zhou et al. [ZLL*19].
2. Biology and medicine: preoperative planning introduced by Pfeiffer et al. [PKP*18].
3. Material science: visual analysis of fibre reinforced composite as presented by Gall [Gal20].

All these fields often have either inherently spatial or sufficiently complex data. Thus, there is a need to combine the well-established 2D screen-based methods with approaches at other stages in the RVC, such as AR, VR or handheld devices, for optimized cognitive reception. As a representative domain of XVA, we dig deeper into the production and supply chain domain in order to discuss two different analysis scenarios in this area, where combining novel devices and techniques for data analysis across the RVC is starting to show the potential to improve the overall analysis experience.

© 2022 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd
Network analysis: The analysis of complex networks, for example from the domain of industrial manufacturing, has received enormous attention from the scientific community [LTC*17, CTXZ18]. Visualization is typically the initial step for analysing complex networks, since it supports the discovery of patterns and the interpretation of emergent properties of the system [Str01]. In their survey of visual analysis of large graphs, von Landesberger et al. count scalability issues in graph drawing, interaction techniques and collaborative visual analysis among the main challenges [vLKS*11]. It has been shown that immersive technologies can facilitate an interactive and efficient visualization of networks [SWKA19, KMLM16, KKM*20]. However, visualization and exploration of networks with immersive technology is underutilized due to various technical, complexity- and visualization-related reasons. Extending systems in the direction of collaborative and cross-device analysis facilitates an efficient and inclusive approach for understanding the complexity of networks. One example of visualization methods for networks is shown in Figure 1 (top image) by Cordeil et al. [CDK*17], which enables to find patterns, structures and complex characteristics collaboratively.

Volumetric data analysis: The most widespread methods for the analysis of volumetric data as generated, for example, by X-ray computed tomography, are currently based on 2D slices or 3D renderings of reconstructed volumes on 2D screens [FWS*19]. Recently, visual analytic systems for computed tomography data have been customized for analysing complex combinations of volumetric data together with derived, abstract information [WAG*16]. The inherent three-dimensional nature of volumetric data lends itself to being explored natively in 3D. VR-based visualizations of such data have thus been long explored [LBS13, WTHM01], but only recently gained momentum in specific, tailor-made applications [MGO*19, NKB*18, PKP*18]. The survey of spatial interfaces by Besançon et al. [BYK*21] also provides a multitude of examples how novel and hybrid interaction paradigms can improve the analysis of spatial data. Extending such systems with collaborative and cross-device aspects shows the potential to generate large benefits in terms of analysis capabilities and usability, as demonstrated by the prototype by Sereno et al. [SBI19], shown in Figure 1 (bottom image).

3. Method

To analyse the related literature in XVA, we reviewed publications at the intersection of visualization, VA, IA, interaction and collaboration. We followed a process of three phases in our literature survey: search (find relevant candidate publications), rate (evaluate the relevance based on predefined criteria) and code (extract the essence of the core-relevant papers). These steps were typically done in sequence, although with some overlap and re-iterations. The search, rating and coding were done by four senior researchers and seven junior researchers in the field of VA, human computer interaction, IA and respective application areas.

Search: For exploring the body of work in XVA, we conducted a thorough literature review. A list of seminal papers in the area was gathered from previous work of the co-authors as initial candidate list. This list informed the list of journals and conferences considered in the subsequent search. We then performed a semi-structured search in common scientific search engines and databases, such as IEEE Xplore, Science Direct, Wiley Online Library and Google Scholar, for combinations of the terms ‘XV’, ‘collaborative’, ‘collaboration’, ‘cross-device’, ‘VA’, ‘VR’, ‘AR’, ‘MR’, ‘immersive’, ‘XR’, ‘virtual environment’ and ‘transitions’. We focused our search on contributions integrating aspects of XVA, which were introduced in top level visualization, visual analysis, VR/AR, interactive surfaces and human–computer interaction conferences and journals. We also went through recent issues of these journals and conferences to check for any publications with a closely related topic that had evaded our previous search. The initial determination of inclusion in the list of candidates was based on title and abstract as well as on their core contributions. Some papers were identified as relevant later, for example, through related work of screened publications. In the end, this process resulted in a total number of 268 candidate papers, which were considered as potentially interesting. The full list of candidate papers as well as the list of considered journals and conferences is available as supplemental material [FAP*21].

Rate: For an initial classification of the relevant literature, the full set of 268 candidate papers had to be rated regarding relevance for XVA. The rating procedure was discussed and continuously refined in several meetings among all authors with the goal of identifying suitable relevance criteria. The initial criteria were collaborative aspects, XV, transitions, devices, visual metaphors, data, application areas as well as contributions. Following the definition of this initial set of criteria, every paper was analysed by three randomly assigned reviewers from the group of authors over a period of 2 months. In cases where an unanimous agreement regarding relevance could not be reached for a candidate publication, the differences in ratings were discussed in weekly meetings of all authors.
and additional opinions were considered. These continuous meetings during the rating procedure helped to refine the relevance criteria and also to unify the way in which they were applied. During this process, we realized that only a few papers fully cover all aspects of XVA. Therefore, we broadened the scope to include papers which are relevant for specific core aspects of XVA. Global criteria for inclusion were based on quality and number of citations, such that, for example, publications of (extended) abstracts with limited descriptions were excluded from further analysis. Additional relevance criteria depended on the type of paper: Papers describing VA systems were rated relevant based on whether the system was considered to feature cross-device and/or collaborative aspects. Transition techniques were considered as highly relevant due to their inherent XR nature. Papers describing interaction techniques or devices as well as user studies and other evaluations were judged based on whether the respective technique or study is useful in the context of XVA. Out of the set of 268 candidate papers, 118 were finally identified as core relevant (see the supplemental material for the full list).

**Code:** In multiple brainstorming sessions interweaving with the rating phase, we identified the most important aspects of XVA in the surveyed literature. In an open coding phase, we then investigated the core-relevant papers, focusing on the identified sub-topics of XVA as reflected in the following sections—levels of XV, transition techniques, collaboration, visualization and interaction techniques, evaluation and application domains. The authors therefore split up into smaller groups of 2–4 coders for these sub-topics, each of which integrated both senior and junior researchers, according to their respective expertise. Subsequently, the sub-groups coded the core-relevant papers in sub-categorizations that were either derived bottom-up from the analysed papers, or based on existing taxonomies on the respective topic. In this open coding phase, 28 papers that were not referenced by any of the sub-topics, were removed from the core-relevant paper list. As indicated above, we discovered in the rating phase that not many systems exist yet which fully embrace our definition of XVA. Therefore, we decided to also add papers which instead cover important aspects to be considered in future XVA systems. Under these considerations, in this phase of the process, we added 34 papers from our candidates list, which initially were not rated as core relevant. Furthermore, we integrated an additional 29 papers resulting from research into the respective sub-topics. The supplemental material provides lists of the references at each step of the process. Figure 2 provides a quick overview of the distribution of publication years of those papers that are referenced in our survey. It is clearly visible that work in the area of XVA has drastically increased in recent years. The detailed findings of the coding for each explored sub-topic are discussed in the following sections.

### 4. Different Levels of XV

MR applications often consider only a single stage within the RVC, either VR or AR, and in a few cases AV. Many approaches have been made to describe the RVC, ranging from fully real to fully virtual and thus allow for a more precise classification. Examples are the Milgram–Weiser Continuum by Newman et al. [NBP07], which introduces an axis ranging from monolithic to ubiquitous computing, or the One Reality framework by Roo and Hachet [RH17a] with six levels of incremental augmentation providing more sub-stages in the RVC. Approaches enabling different stages in the RVC to work together are of specific interest for XVA. They can either interconnect different stages in a single application, or they can employ networked devices for linking different stages. These approaches all share the property that multiple stages in the RVC are integrated, allowing for XV interaction. We refer to the degree to which XV is embraced in these approaches as levels of XV. Approaches with a low level of XV merely combine different stages in a single application. They do not consider XV-specific aspects such as the respective spatial positions of these stages or the seamless transition between different stages. The more such XV-specific aspects are embraced, the higher the level. From our survey over the existing literature, we have identified four categories of interconnections between stages. The identified categories, listed from lowest to highest level, are:

1. **Spatially agnostic XV:** approaches that combine different stages in the continuum in a single application without considering the spatial reference of these stages to each other.
2. **Augmented displays:** (as introduced by Reipschlager and Dachselt [RD19]) approaches that augment one stage in the RVC with the help of another stage and put them in a spatial relation.
3. **Networked XV:** approaches that interconnect users in different stages in the continuum with each other.
4. **Transient XV:** approaches that allow true shifting throughout the RVC.

We provide examples from our core-relevant papers and also note the different input and output devices in combination with their interaction techniques used in the examples. Table 1 gives an overview of the classification of the discussed publications according to their level of XV.

#### 4.1. Spatially agnostic XV

We adapt the concept of spatially agnostic displays as introduced by Rödle et al. [RJS15] in the domain of XV. They mainly deal with mobile devices not including MR systems. In the context of XVA a change of a parameter or dataset in one stage (for example, a desktop system) will cause a manipulation of the data in a different stage (for example, a VR visualization). Both systems are not spatially related to each other.
### Levels of XV

<table>
<thead>
<tr>
<th>Levels of XV</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatially agnostic XV (n = 7)</td>
<td>[WBR<em>20, SBI19, CCB</em>19, GPV<em>15, CML</em>12, WGA*16, KSEM17]</td>
</tr>
<tr>
<td>Augmented displays (n = 13)</td>
<td>[RD19, K097, WBS20, LSBD21, DACJ13, RFS<em>18, BHm</em>18, MBD<em>18, RFD20, BLD21, GAWK16, NJ19, PLE</em>19]</td>
</tr>
<tr>
<td>Networked XV (n = 9)</td>
<td>[PLLB17, GDm19, RH17b, CRHG17, BJ20, PGW<em>14, GLB05, CDH</em>19b, CDH*19a]</td>
</tr>
<tr>
<td>Transient XV (n = 6)</td>
<td>[KTY99a, ESE06, BKP01, RH17a, RBCH18, BIF04]</td>
</tr>
</tbody>
</table>

### 2D displays extended by AR

In our literature survey, a wide variety of motivations to combine 2D displays with AR devices was observed. One of them is to exploit the benefits of conventional desktop-based systems such as high-precision input and extensive functionality, while using stereoscopic displays in AR for a better perception of 3D data. Wang et al. [WBR*20] introduce such an approach by complementing a 2D desktop display with an optical see-through HMD as an additional stereoscopic screen for data visualization tasks. Users can arrange visualizations between those two devices and work with data on both displays simultaneously. Unlike in many other scenarios in literature, the desktop mouse is also used as input device for the AR workspace in this case. Utilizing stereoscopic see-through HMDs as additional screen to a 2D display can also be found in multi-user environments to improve sense-making and to customize views according to user privileges. Kim et al. [KSEM17] demonstrate this in a scenario for presentations in which the audience is equipped with optical see-through HMDs while the presenter shows static 2D visualizations on a projector or wall screen. In addition to conventional augmentations of the projector visualization, users can open a details-on-demand view in parallel on their HMD by gesture and voice input. The level of detail that is shown in this view is determined by the privileges of the user mapped to the HMD. Alternative input for optical see-through HMDs, which mostly use hand tracking, can be provided by multi-touch 2D displays such as tablets. These can be used to manipulate and annotate 3D visualizations that are shared across multiple optical see-through HMDs as demonstrated by Sereno et al. [SB19]. By manipulating an orthogonally rendered copy of the visualization on the tablet that is used as personal workspace, those changes are applied on the 3D visualization that is shared across all users wearing an HMD. Interaction between 2D displays and AR is, however, not unidirectional. Cordeil et al. [CCB*19] introduce the IA toolkit for data visualization designers that is capable of maintaining coordinated views distributed at different stages in the RVC. Multiple display types such as AR and VR HMDs as well as 2D displays are supported. However, only one XV use case connecting an AR application with a 2D desktop display is provided. After selecting a data point in a 3D visualization that is displayed on a stereoscopic AR headset with the hand tracking provided by the HMD, a details-on-demand view for the selected data point is opened on the 2D display.

### 2D displays extended by VR

In VR, different 2D displays such as tabletops and mobile devices are frequently combined with CAVE-like environments [CSD*92] to combine the benefits of 2D visualizations and immersive environments. For instance, Gebhardt et al. [GPV*15] suggest a factory planning tool that provides an editable overview of the factory layout on a tabletop display. Changes in the layout can instantly be observed during a walkthrough of the 3D representation of the factory layout within the CAVE. Touch-enabled tabletop displays can also be used to navigate through complex volumetric data that are displayed in a CAVE. Coffey et al. [CML*12] establish this by combining a large-scale stereoscopic detail view of the dataset in the CAVE with a second projection above a tabletop surface showing a miniature overview of the whole dataset. A cutting plane whose current position in the dataset is reproduced visually in this miniature overview is projected on the tabletop surface, imitating a shadow from the projection above the tabletop. The cutting plane on the tabletop can be moved and manipulated by multi-touch interactions to navigate through the detailed data in the CAVE while the overview projection above the tabletop gives context about the current global position of the cutting plane in the dataset. Similarly, tablet devices can be used instead of wands in CAVE environments to select and manipulate 3D objects with touch-based interaction methods. This can lead to improved object selection speed [WGA*16]. The low presence of other VR display technologies such as HMDs in combination with 2D displays might be caused by the relatively high degree of isolation from the real environment when wearing an HMD and is most likely to be found in collaborative XV scenarios as further outlined in Section 4.3. For instance, in their DataSpace prototype, Cavallo et al. [CDH*19a] combine multiple wall screens with a stereoscopic VR HMD worn by a user at a distant location. The content of the wall screens is re-produced in VR, and touch input is emulated using the wands of the HMD, offering an interaction modality with the remote users who work with the wall screens.

### 4.2. Augmented displays

Reispschlager and Dachselt [RD19] introduced the concept of augmented displays, which they describe “as the extension of an non-stereoscopic, interactive surface [sic][...], with two- or three- dimensional content using personal AR devices”. This concept in most cases combines 2D surfaces with AR, but can also be applied for VR systems.

Few applications exist which fully support the idea of augmenting displays. An early approach extruding planar projected CAD data out of a display with the help of a tailored optical see-through HMD is described by Kijima and Ojika [K097]. They already provide keyboard and mouse input as well as a tracked wand for interaction. The focus of their publication is mainly on the technical realization of the system setup. Extrusion from desktop displays has also been shown by Wu et al. [WBS20]. In their MergeReality system shown in Figure 3(left), they use an optical see-through HMD with a traditional screen and a tablet. The interaction is performed with hand...
tracking and gesture recognition. An interesting aspect in their approach shows the manipulation of the augmentation of the environment, by augmenting real-world light sources for example. Mobile 2D displays have been explored as source for content extrusion as well. For instance, Normand and McGuffin discussed several applications of 2D augmentations that are arranged co-planar to a smartphone screen to extend the rather limited screen space [NM18]. In their work, a video-based see-through HMD with hand tracking is used to register optical markers attached to a smartphone and to augment the area surrounding the mobile screen with 2D visualizations that follow the smartphone as it moves. Input can be provided by touch input either on the smartphone display or mid-air on augmentations via hand tracking. Prouzeau et al. [PLE*19] focus on the interconnection between different datasets. They augment screens and interconnect them with real-world tangible objects. An interesting aspect is the recognition of real-world objects to avoid occlusion of displayed links by such a real-world object. Reipschläger and Dachselt [RD19] combine a tablet for interaction and a 2D data display with an optical see-through HMD to extend data into space. The application scenario of their tool DesignAR, shown in Figure 3(right), is the creation of 3D content by sketching on the tablet in 2D and extruding the outline into 3D space. They use touch and pen input on the 2D surface to intuitively manipulate the datasets. With the use of an optical see-through HMD, the data can be placed at an arbitrary location in space after extrusion. A variety of solutions to augment tablet devices with the help of HMDs in the context of VA is provided by Langner et al. [LSBD21]. They discuss in detail how to layout data next to or on top of the devices. Interestingly, they also allow to interconnect multiple devices and display links between multiple devices.

The Mockup Builder by De Araújo et al. [DACJ13] combines an interactive back-projected tabletop system in the real environment with an additional stereoscopic representation realized by active stereo with shutter glasses. Touch gestures on the tabletop as well as mid-air gestures are recognized. In their prototype, they show how sketching and modelling can be performed on a 2D surface. Similar to DesignAR, the 2D shapes are subsequently extruded to 3D models in the 3D space where they can be further manipulated. The system restricts the 3D dimensions of the data to the area of the tabletop surface.

For more generalized purposes also including data visualization, Riemann et al. [RFS*18] developed Overtop, a combination of a tabletop system and an optical see-through HMD. They separate the space as below, on and above a table. The interaction is processed with the help of a depth sensor and supports tangible objects. A dedicated display augmentation for IA, based on a tabletop in combination with a video see-through HMD is presented by Butscher et al. [BHM*18]. As shown in Figure 4(left), they present multidimensional data to the users in 3D space, which are interconnected with parallel coordinates. The table acts as an interaction device to scroll through different plots for 3D display. The representation of the single plots is available on the table surface, as well as in the augmentation. Both representations are spatially linked. Besides the augmentation of interactive tabletops and tablets, which are limited to small user groups, approaches for augmenting large scale wall displays have been developed. Mahmood et al. [MBD*18] combine a large display with optical see-through HMDs to show multiple coordinated view visualization widgets to the users, as shown in Figure 4(right). They also allow the display to be extended with additional virtual displays on the same plane or into space. The input is generated by gaze interaction and voice commands. Reipschläger et al. [RFD20] analysed the challenges of large scale wall displays in regard to perception, effective multi-user support and managing data density and complexity. They implemented a prototype including an optical see-through HMD to overcome these challenges by providing spatial alignment of displays, visualizations and objects in space. These alignments include hinged visualizations and curvature into the AR space. A combination of augmenting tablets, a wall display and the environment is achieved by Büschel et al. [BLD21]. They use optical see-through HMDs to display motion data for analysis of a spatial interactive game. They provide in situ visual analysis of the spatial interaction and devices, with the help of heatmaps, scatter plots and trajectories.

In 3DScoveR [GAWK16], a CAVE-like environment is used as a planar information display. Three walls of the CAVE show different 2D representations of the same dataset. The additional stereoscopic display allows the spatial linkage of these datasets. Interaction is performed with a wand but is extended to make use of a spatially tracked tablet interface [WGA*16]. The tablet allows for text search inside the dataset. To overcome the limitations of display space, Nishimoto et al. [NJ19] use an optical see-through HMD to extend a CAVE-like display with the help of augmentations. When the user’s visual field of view moves out of the bounds of the display, an augmentation of the content is provided.

The analysed approaches of augmented displays can be categorized based on their display type as shown in Table 2.
The desktop and screen systems implement input with the help of wands and gesture tracking. Tablet systems work with touch gestures or pen devices, similar to the tabletop approaches. Some systems [RF10, RD20, LSBD21] point out the advantages of having the planar display as shared space but additionally providing the augmentation as potential private space. Most approaches use optical see-through HMDs for implementing the augmentation. The main use cases for augmented displays are sketching and 3D model creation [RD19, DACJ13] and VA [PLE19, LSBD21, MBD18, RD20, BLD21, GA21]. From the 13 publications analysed in this section, ten were published in the last 3 years, demonstrating the importance and the current trend in this subcategory of XV.

### 4.3. Networked XV

The XV aspect of Networked XV scenarios is realized by connecting multiple users on different stages in the RVC (such as VR and AR users). Respective applications are all multi-user applications. Common scenarios include: Telepresence systems, where real and virtual environments are interconnected; remote support applications, where real and AR environments are interconnected or classical networked virtual environments, staying within one stage of the RVC, which are coupling two or more VR applications. In this section, we focus on less common approaches approaching different stages in the RVC and point out interconnections that have not been explored intensively before. Most of the approaches discussed here are also analysed in Section 6 regarding their collaborative aspects.

Piumsomboon et al. [PLLB17] interconnect users of an optical see-through AR HMD with users of a VR HMD. Their focus lies on different asymmetric user representations, interaction techniques (eye gaze, head gaze, hand gestures) and a variety of collaboration aspects. To improve remote collaboration, Grandi et al. [GDM19] provide interaction between co-located AR and VR users. The AR users interact with augmentations on tablet devices, leading to completely different interaction metaphors, compared to the VR users equipped with a wand. An interesting aspect is the support of concurrent object manipulation, which is not often implemented due to consistency issues. The tablet interface allows for touch gestures and spatial manipulation by altering the transformation in space. The VR side works with an adapted HOMER technique [BH97] and a Spindle technique [CW15] metaphor. Roo and Hatchet [RH17b] use spatial AR (an AR sub-category, where augmentations are projected onto real objects, as described by Bimber and Rashkar [BR05]) in combination with VR to establish interaction between multiple co-located users, as shown in Figure 5. In their setup, both stages in the RVC work with the same interaction modalities and use a tracked wand, which is uncommon in such settings due to different hardware setups. It is also possible for the users to move from the VR domain into the AR domain by simply removing the headset.

Pick et al. [PGW14] interconnect two stereoscopic projection systems. Their approach allows for viewpoint sharing by integrating a tracked tablet to take snapshots or use a direct mapping of the head tracking. Thus, they stay in the same stage in the RVC but only share certain viewpoints. The virtual environments that both sides explore are not identical. The display of the shared viewpoint can be considered as an augmentation of a snapshot of one virtual environment inside of the remote-networked virtual environment. To communicate between real and virtual environment Clergeaud et al. [CRH17] establish a connection between a physical group meeting in a real environment, supported with spatial AR projections and a remote VR user. They suggest using nearby CAVE-like installations in order to allow participants to move from the physical space into the VR environment. Butcher et al. [BJR20] created VRA, a web-based software framework for VA, which also supports VR and AR devices besides browser-based data display, as shown in Figure 6. The data presented in VRIA has also basic multi-user support. Similar approaches exist for device abstraction in VR space, as for example, the NomadVR framework [GK19].

Grasset et al. [GLB05] provide a conceptual framework approach to interconnect multiple stages in the RVC, which they call Mixed Space. They perform multiple experiments comparing a VR single user, VR collaboration and a Mixed Space condition, using a VR HMD, a video-based see-through HMD and a tablet. Navigation and interaction tasks in different conditions are explored. They also provide an outlook into transitional XV. A high-end VA setup was

---

**Table 2: Types of augmented displays and their publications.**

<table>
<thead>
<tr>
<th>Augmented display</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desktop and screen (n = 3)</td>
<td>[KO97, WBS20, PLE19]</td>
</tr>
<tr>
<td>Tablet (n = 3)</td>
<td>[RD19, WBS20, LSBD21]</td>
</tr>
<tr>
<td>Tabletop (n = 3)</td>
<td>[DACJ13, RFS18, BHM18]</td>
</tr>
<tr>
<td>Wall and multi-display (n = 5)</td>
<td>[MBD18, RFD20, BLD21, NJ19, GA21]</td>
</tr>
</tbody>
</table>

---

**Figure 5:** Representation of an AR user in a spatial augmented reality scenario (on the left) interconnected with a VR user (on the right) in the scenario described by Roo and Hatchet. Republished with permission of ACM, from Roo and Hatchet [RH17b]; permission conveyed through Copyright Clearance Center, Inc.

**Figure 6:** Data analysis in VR (left) and with a browser interface (right). © P.W.S. Butcher, N.W. Johnson and P. D. Ritsos, reprinted, with permission, from Butcher et al. [BJR19].
VR mode. The other person is displayed as a computer-generated avatar, and the position in the virtual environment is changed according to the input. To improve tracking accuracy especially for AR, they implemented a look-up-table to correct the distorted magnetic field and compensated the computational delay with a Kalman filter.

Benko et al. [BIF04] introduce VITA (Visual Interaction Tool for Archaeology), a collaborative XV visualization of an archaeological excavation. In addition to a video see-through HMD that supports AR and VR, they use a multi-touch projected table, a large display and a tracked handheld display. A tracked glove is used for gesture input and can be used to drag objects from 2D into the 3D environment.

To support more anchor points on the continuum, Eissele et al. [ESE06] added AV in addition to AR and VR in their smart production application. They used an optical see-through HMD to support these different stages and a single button in combination with the user’s viewing direction to map all required interactions. A stationary-optical tracking system is used to track the position and orientation of the user’s head and physical objects. The application starts in VR to explain the machine and the installation environment. The user can select an installation position, and the HMD switches to AR mode. The system then augments the real machine with the virtual parts in the correct position. For the final check, the user can start the AV mode, which shows an assembled real machine to recheck with the user’s own machine. The usability test showed that the average assembly time is shorter in the group with XV than in the group with traditional technical drawings.

Another way to move along the continuum is to incorporate the real world, as Billinghurst et al. [BKP01] demonstrate with the MagicBook. They implemented a collaborative environment where users can switch between reality, AR and VR, as shown in Figure 9. A real, physical book is used as the main interface object and can be read like a regular book. A video see-through handheld display with an inertial tracker is used to support AR and VR. In AR mode, computer vision is used to place the 3D virtual content on the real book. When the user wants to explore a particular scene in more detail, the user can flip a switch on the handheld in order to fly into the VR environment. The user is then able to explore the scene in an egocentric view. Tracking is switched from the computer vision module to the inertial tracking system. Similar to the navigation technique of Eissele et al. [ESE06], a button on the handheld is used to navigate in the direction the user is looking.

© 2022 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd
4.5. XV challenges and recommendations

XV systems face many challenges due to the diversity of hardware and the applied interaction techniques. Especially when network aspects are introduced, additional technical problems of responsiveness and scalability of these systems become visible. We focus on the key challenges pointed out in our analysed publications and highlight some recommendations:

1. Displays and rendering: Extending large displays with augmentations can be used to overcome limitations of the displays, for example, by providing an overview even when close to the viewer [RFD20]. Details of volumetric data are less visible on 2D screens and are therefore typically presented in AR/VR using 3D visualizations [CML*12]. The degree of immersion into the data should be flexible and be adaptable to the context [CDH*19a]. Augmentation has to be distinguishable from real-world objects in terms of occlusion and colour [PLE*19]. Rendering performance for AR can be rather low on mobile devices with limited GPUs [CCB*19].

2. Interfaces and interaction: Cavallo et al. emphasize the need for seamless integration of the heterogeneous devices found in XV applications to avoid disruptions when interacting across devices that might also be located at different stages in the RVC [CDH*19a]. Furthermore, those devices should allow multi-modality to incorporate personal preference, spatial position and environment-specific properties. Besançon et al. refer to this in their survey on spatial interfaces for 3D visualizations as hybrid interaction and point out that despite the potential benefits of combining different input paradigms, most publications focus only on the combination of touch and haptic interaction [BYK*21]. One reason might be that although multi-modal sensory input can contribute to better user experience, it also increases system complexity and puts additional challenges on computational resources. This is particularly the case in XV scenarios, which are already exposed to a high load in rendering [MMG*21]. Trade-offs between different displays and input devices have to be made, and input modalities should be chosen depending on application context. For instance, typical XV interaction with mid-air gestures via hand tracking lacks haptic feedback, therefore, touch interfaces might still be helpful when manipulating augmentations [RFD20]. Compared to wands, touch interaction can provide higher selection speed and accuracy [WGA*16], while touch interactions can provide lower task completion times and error rates when compared to mid-air interaction based on hand tracking [NM18]. Other input devices specifically designed for 2D interaction (such as a mouse) are experienced as rather uncomfortable in 3D space [WBR*20]. Switching between a life-size 3D model in VR and a world-in-miniature visualization in AR while providing additional 2D information provides useful insights and relationships for the user task [BIF04]. When physical objects are decoupled from the virtual representation, additional considerations should prevent accidentally bumping into them. Proximity estimation can be used to display wireframes around the physical object when the user gets too close to an object [RH17a].

3. Hardware: Early publications showed that restricted field of view, darkness of the optical see-through display, the low brightness of augmented objects, weight of HMDs, extensive wiring...

The One Reality concept by Roo and Hachet [RH17a] enables the user to move along the RVC, as shown in Figure 10. Just like MagicBook [BKP01], it starts in the real environment allowing to interact with physical objects. Spatial AR is used to display additional content onto physical objects, provided by three projectors. As a supplement, they projected onto a tracked paper, to simulate a see-through handheld by coupling the virtual camera position to the paper position. For VR, they used a tracked HMD with additional hand tracking to provide visual feedback and interact directly with purely virtual content. In the final stage, the body is decoupled from the environment to change position, orientation and scale in space. To navigate in the virtual environment, teleportation-based interaction with a wand controller is used. A depth camera allows collaborators to be reconstructed as avatars in VR mode. In a similar setup, Roo et al. [RBCH18] focus on user accuracy with spatial AR using projectors and VR using an HMD. As input device, they use a wand with an additional optical tracking system to achieve higher accuracy. To switch between spatial AR and VR, the user has to take off or put on the HMD.

Figure 11 provides an overview of the described transient XV publications from this sub-section, including the stages where they are able to move along the RVC. In order to make this movement possible, a transition technique is required, which performs a transfer that is more than an immediate change of environment. Section 5 deals with such possible transition techniques.
and slightly out-of-place augmentations can cause discomfort during use [BIF04, ESE06]. But still immersive environments—collaborative environments in particular—have to balance a wide variety of technological trade-offs between hardware complexity, image quality, resolution, field of view, depth rendering, visual acuity, perception issues and cost [CDH+19a].

4. Transient XV: Our literature search revealed that all six papers identified as transient XV use individual visualization elements, but are outside the scope of traditional IA. In particular, the highest level of XV has not yet been applied to the visualization of traditional trees, graphs and networks. Further research needs to clarify whether and how IA can be applied to transient XV. Similarly, recent work on transient XV by Roo and Hachet [RH17a] and Roo et al. [RBCH18] require the HMD to be set down to move along the RVC. The HMDs used by Kiyokawa et al. [KTY99a] and Benko et al. [BIF04] already supported AR and VR, which is desirable for this level of XV, as otherwise the devices would have to be changed during the transition. Recently, commercial HMDs have become available that use a video-based see-through mode, such as the HTC Vive Pro Eye [HTC21] and the Varjo XR [Var21]. These allow not only VR but also AR as well as interaction with the real environment without switching HMDs.

5. Transition Techniques

Transient XV allows a true shift along the RVC. To perform this shift, a transition technique is required. This transition technique guides the user during the shift from one stage to another. A first classification of possible transition techniques was introduced by Pointecker et al. [PJA20] in 2020. They mentioned that a transition should be primarily seamless, so that users remain focused on their tasks. However, there are also examples from the past where no transition is used, instead the environment is changed all at once. This was the case in one of the first publications in this area by Kijima and Ojika [KO97] where they used the position and orientation of the head to switch instantly between the desktop and AR environment. Similarly, Kiyokawa et al. [KTY99a], Benko et al. [BIF04] and Eissele et al. [ESE06] presented systems to switch between VR and AR, but they did not use any transition techniques, instead they simply switched to the desired environment. All four approaches show an early need for combined AR and VR visualizations, but provide a hard cut between the different stages. Possible reasons why they did not use a seamless visual transition are not stated by the authors. It could be due to hardware limitations or the lack of explored transition techniques in literature. Since combined MR devices have not been widely used in recent years, the amount of research for transitions between reality, AR and VR is limited. Nevertheless, transitions from other domains, such as scene changes in VR can be adopted for a transition along the RVC. Out of the 118 papers rated as relevant, transition techniques were used in ten cases to move between scenes or between stages. Table 3 gives an overview of the identified transition techniques, which are more than a simple immediate change of environment.

5.1. Portal

A portal transition is an easily understandable transition as it works the same way as a real door that gives access to a new environment. To change the scene or the environment, the user needs to physically walk through the portal. The transition process is complete when the user exits the portal on the other side. All the reviewed portal implementations have a preview that is rendered within the portal and offer the possibility to have a look at the target environment. Slater et al. [SSMM98] and Steinicke et al. [SBH+09] created a replica of their laboratory and changed to the virtual target environment via a portal. Slater et al. [SSMM98] used a door that served as a portal, while Steinicke et al. [SBH+09] used a large circular portal, as shown in Figure 12(left). Clergeaud et al. [CRHG17] used a small ring-shaped portal and a door to support the collaboration between spatial AR and VR. The ring-shaped portal is used to interact and to move elements between the environments. An additional door is used to change the entire environment. Another portal implementation can be found in the work by Husung and Langbehn [HL19], where an oval portal opens in front of the user. After the users have passed through, the portal closes behind them. In contrast to the previously mentioned portals, the implementation by Nam et al. [NMT+19] uses the portal metaphor to represent different virtual environments in VR. Three portals are clipped together, creating three wedges with different environments. The user cannot walk through the portal but can increase or decrease the size of the portals, which affects the set of perceived environments. George et al. [GTH20] implemented a stationary sky portal and a portable virtual phone to provide glimpses of other stages on the RVC, similar to windows or screens to other worlds. The portals are used as an in-between state to display additional information from the virtual or real environment. However, unlike the implementations of Steinicke et al. or Clergeaud et al., the portal is not used to change the entire environment by walking through it. Instead, the user is able to teleport directly to the other environment without using any transition technique.

Table 3: Types of transition techniques and their publications.

<table>
<thead>
<tr>
<th>Transition technique</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portal (n = 6)</td>
<td>[SSMM98, SBH+09, CRHG17, HL19, NMT+19, GTH20]</td>
</tr>
<tr>
<td>Fade (n = 3)</td>
<td>[HL19, MBHM17, SWOG17]</td>
</tr>
<tr>
<td>Offscreen transition (n = 2)</td>
<td>[SWOG17, VF17]</td>
</tr>
<tr>
<td>Other techniques (n = 4)</td>
<td>[BKP01, MBHM17, SWOG17, HL19]</td>
</tr>
</tbody>
</table>

Figure 12: Portal transition (left, © 2009 IEEE, reprinted, with permission, from Steinicke et al. [SBH+09]) and Fading transition (right, republished with permission of ACM, from Husung and Langbehn [HL19]; permission conveyed through Copyright Clearance Center, Inc.) between VR scenes.
5.2. Fade
Fading is a more hidden transition where the transition technique gradually changes the surrounding environment. Over time, the visibility of virtual objects changes slowly to reveal the target environment and block the source environment. Husung and Langbehn [HL19], as well as Men et al. [MBHM17], implemented a fading transition, which gradually changes the VR environment first to black and then to the target environment. Due to the black fading, both environments are not visible at the same time, and the user’s field of view is completely black for a short time. In addition, Husung and Langbehn [HL19] implemented the technique without fading to black, as shown in Figure 12(right). In this case, the transparency of the source environment increases, while the transparency of the target environment decreases. As a result, both environments are visible at the same time in the middle of the transition process. The user study indicated that the visual flow is not interrupted as much without fading to black as with it, resulting in higher continuity. In contrast, Sisto et al. [SWOG17] only fade the appearance of the floor and the sound effects of the environment.

5.3. Offscreen transition
In the offscreen transition, objects from the source environment are immediately replaced with objects in the respective region of the target environment, but only in areas outside of the user’s field of view. In the best case, the user does not notice much, since only objects currently invisible to the user are transitioned. Such a technique is implemented by Sisto et al. [SWOG17]. The HMD-based VR application by Valkov and Flagge [VF17] is using the offscreen transition to switch from a replica of the real environment to the actual target environment. This technique only works reliably if the user explores the environment. If the user keeps looking in the same direction, not all objects can switch to the target environment. To prevent this, Valkov and Flagge [VF17] built in a timer that changes objects after a certain amount of time, regardless of which direction the user is facing.

5.4. Other transition techniques
In addition to the three more common transition techniques explored above, there are several that were used in only one publication. Besides Fade, Men et al. [MBHM17] describe three other transition techniques: SimpleCut, Vortex and FastMovement. SimpleCut uses a cutting plane that gradually moves through the entire environment, cutting away the old environment and revealing the new one. The Vortex transition is a visual effect that creates the illusion of a rapidly spinning vortex that picks up the user and teleports to the new environment. With FastMovement, the camera and thus the field of view, moves very quickly to the target environment. However, this transition may be difficult to implement when switching between different stages in the RVC. The included study reveals that rather invisible transitions such as Fade and SimpleCut maximize plausibility and sense of presence. In addition to the offscreen transition, Sisto et al. [SWOG17] also investigate Morphing in their work, which involves gradually changing the shape and texture of objects. In another technique, they change the size and position of objects so that objects in the old environment become smaller and those in the new environment become larger. This also allows objects to be moved away from the user. The Fragmentation transition blasts objects, breaking the objects into tiny pieces. The user study concludes that only 22% of the transitions were noticed and the changes had no effect on the user’s tasks. Husung and Langbehn [HL19] examine Orb and Transformation in addition to Fade and Portal. The Orb technique is similar to a portal, it renders the new environment inside an orb. However, it is not possible to walk through it, instead the orb needs to be dragged over the head to change the environment. The Transformation transition spreads a rift around the user until the old environment is completely gone and the user is in the new environment. A user study showed that the Orb and Portal transition techniques performed significantly better than the other techniques in terms of presence, continuity and user preference. A transition technique that switches between different stages is used in the MagicBook by Billinghurst et al. [BKPO1]. They used a flight metaphor to transition from AR to VR. The user perceives the transition as flying from an exocentric AR view to an egocentric VR environment. The effect of this transition on the user is not described in more detail.

5.5. Transition challenges and recommendations
Our analysis revealed the following key challenges and recommendations for future work with regards to transitions:

1. Adaptation: Out of ten papers that deal with a transition that is more than an immediate change, there are only two papers that use a transition between different stages in the RVC [BKPO1, GTH20]. It is therefore important to adapt these transition techniques, which are already established for transitions between VR scenes, and use them for transitions along the entire RVC. Especially for the highest level of XV, a seamless transition is crucial. Therefore, further research is necessary to understand how such a transition needs to be designed to best support the user.

2. Transition metaphor: Depending on the employed transition technique, different metaphors are conveyed to the user, which can influence the perception of the transition. Future work should therefore investigate the desired effect of the transitions on the user in more detail in order to be able to use them in a targeted manner. A more invisible transition such as Fade or SimpleCut should be used when maximum continuity of presence is desired, and a visible transition such as FastMovement or Vortex when the continuity of the experience should be broken [MBHM17]. Steinicke et al. [SBH*09] describe that in their implementation of portals, the wormhole metaphor was perceived by participants exactly as intended. Future work should therefore use transitions in a way that provides the desired effect on the user.

6. Collaboration in XVA
To understand the state-of-the-art of collaboration in XVA, we have analysed 64 papers concerned with collaboration within the RVC. While we initially focused strictly on XVA, we subsequently broadened our scope by including papers dealing with collaboration domains beyond data visualization and analytics. We also included papers on collaboration that does not span different stages in the RVC but happens entirely within either AR or VR. We considered this
Table 4: The included papers classified by the Time–Space matrix.

<table>
<thead>
<tr>
<th>Time–Space quadrant</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same time, same space (n = 37)</td>
<td>[BCBM18, BVC16, BHMR16, BSZ<em>18, BRNB19, BHM</em>18, CDH<em>19b, CDK</em>17, EGP<em>20, FBW17, FYS</em>19, GOR<em>19, GDM19, FPS</em>12, IJS<em>15, KGZ</em>11, KTY99b, KTY99a, KBKBC19, KWFK20, LCPE19, MAG<em>16, MZW</em>19, NIA<em>18, NWE</em>19, PDE<em>19, PIR</em>17, PBC17, RJS<em>15, RLR</em>20, RH17a, SCI04, SI19, SSFG06, TTP<em>06, ZPR</em>16a, GSFR17]</td>
</tr>
<tr>
<td>Same time, different space (n = 20)</td>
<td>[AV05, BSYB20, CLST15, CRHG17, GLB05, GGG98, JSH<em>13, MRR16, MRR17, MZW</em>19, NIA<em>18, ORF</em>16, PLLB17, PLB18, PC17, PBC18, RZF<em>07, RZP</em>19, RAI*20, SNLW01, UKAG19]</td>
</tr>
<tr>
<td>Different time, same space (n = 4)</td>
<td>[AKA<em>20, NJA</em>18, PGW*14, UKAG19]</td>
</tr>
<tr>
<td>Different time, different space (n = 6)</td>
<td>[DDC<em>14, JKK</em>17, KFC<em>19, LXYM20, NIA</em>18, SRJR15]</td>
</tr>
</tbody>
</table>

broadening of the analysis scope as justified since important challenges such as awareness, coupling and territoriality (see below) are equally relevant for XVA as for MR collaboration in general.

For further classification, we have used Johansen’s Time–Space matrix [Joh88] with its two dimensions time and space. In the time dimension, collaboration can occur synchronously (at the same time) and asynchronously (at different times). In the space dimension, collaboration can take place co-located (in the same space) or remote (in different spaces). For example, business meetings around a conference table happen same time, same space while collaboration via e-mail typically happens different time, different space. Each paper was assigned to one or, when necessary, multiple quadrants of the matrix: different time, same space (four papers); different time, different space (six papers); same time, same space (37 papers); same time, different space (20 papers), as shown in Table 4.

6.1. Being co-located in physical versus virtual space

The great majority of 56 papers is concerned with synchronous collaboration. This reflects the growing need for synchronous collaborative teamwork for solving complex tasks in multi-role teams (such as sense-making or analysis of data), no matter if co-located, remote or partially distributed across different locations, as in the work by Neumayr et al. [NJA*18]. However, during further classification attempts, we noticed that in the context of XV or XR systems, the meaning of same space or different space becomes complex and ambiguous.

Users can be perceptually separated by being immersed in different virtual environments or by differing stages in the RVC even when they are spatially co-located. For example, two co-located users might interact with virtual 3D objects, one in VR using an HMD and the other in AR using a tablet, as for example in the work by Grandi et al. [GDM19] (see Figure 13). In principle, the VR-user could be geographically co-located in the same physical space or room but without having any visual or auditory contact to the AR-user. In fact, the AR-user might not be present in the VR-user’s perceived environment at all. On the other hand, users can be located at different geographical locations and in different rooms but still perceive each other through audiovisual representations as if they were co-located (for example, in AR as in the work by Orts-Escolano et al. [ORF*16] or in VR as in the work by Buck et al. [BRNB19]). Therefore, future work in XVA must establish a more precise terminology and new models that differentiate between physical and perceptual co-location and support collaboration across this additional dimension.

In order to understand how this new and different notion of co-location can affect collaboration, Grasset et al. [GLB06, GLB05] were the first to study how team members would use their ability to move either individually or collaboratively between different stages (such as between AR and VR) during collaborative problem solving. Also, early research on scenarios where users switch between multiple viewpoints as well as manipulate one’s view of other users in a collaborative environment are highly relevant for collaborative XVA [BBF*95]. Nonetheless, scenarios of users sharing the same physical space but being separated by different stages during collaboration will still require further research, especially with regard to information transfer, awareness and awareness cues that are key to successful XVA.

6.2. Awareness cues for XVA

To enable efficient team work, collaborative systems rely on awareness among team members. That is, they require a common understanding of their situation, task, workspace, presence and similar aspects of collaboration. Systems that employ awareness cues to achieve this contribute to a reduction of effort, increased efficiency, and fewer errors during collaborative tasks [ND14].

© 2022 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd
Therefore, in addition to the well-documented challenges of establishing awareness during collaboration in general [GG98, ND14] and, in particular, for symmetric MR collaboration (such as AR-only [BIF04] or VR-only [CDK*17]), XVA collaboration also requires XV awareness cues [PDE*19]. In parts, this was already explored in remote collaboration between different networked stages, as mentioned in Section 4, with systems such as CoVAR [PLLB17], MiniMe [PLH*18] or the work of Bai et al. [BSYB20]. Plumsomboon et al. [PDE*19] showed that awareness cues should be more than just virtual visualizations of each collaborator’s head and hands, and proposed combining field-of-view frustums and head-gaze rays in XR environments. Bai et al. also explore sharing eye gaze and hand gestures [BSYB20]. However, awareness cues for same space, same time, but different stage will additionally need to account for new challenges such as collision avoidance between users immersed in VR and/or other users or objects that are physically present in the same physical space but not necessarily in the same virtual environment. Another open challenge is the representation of remote users in a degree of fidelity which is reasonable for XVA collaboration. Self-representation also is an issue when using VR HMDs, but when collaborating throughout the RVC, this problem becomes more prominent. The level of detail of the user representation should depend on the current stage in the RVC. In AR, a simple arrow can indicate position and orientation of collaborators. In VR, a 3D reconstruction of the user, which mimics his gestures, should be used [RH17a].

6.3. Coupling styles and territoriality in XVA collaboration

The aforementioned new notion of co-location in XVA could create new research opportunities by transferring other concepts from traditional co-located collaboration into XV. For example, similarly to Neumayr et al.’s work on hybrid collaboration [NJA*18], future work will need to introduce new coupling styles for a detailed analysis of collaborative practices, such as pointing gestures or other spatial activities between users distributed across different stages. The concept of territoriality was originally introduced for describing personal, shared and storage spaces during co-located collaboration around tablespots by Scott et al. [SCI04]. Later, this concept was adopted for auditory spaces in hybrid collaboration by Neumayr et al. [NJA*18]. So, territoriality could prove to be highly relevant for sharing and discussing digital objects across stages.

6.4. XVA collaboration challenges and recommendations

Our analysis revealed three challenges that future research on collaborative XVA needs to address:

1. Future work in XVA must establish a more precise terminology and new models for understanding the intricacies of same space collaboration that differentiate between physical and perceptual co-location and account for the possible distribution of team members across different stages.

2. A critical challenge for collaborative XVA is to visualize improved and novel XV awareness cues that share multi-modal information (such as gestures, eye gaze or spatial audio) to enable efficient communication and spatial interaction across stages. Other challenges to be solved include collision avoidance or pointing gestures for users in shared physical but different perceptual spaces.

3. Future work could draw strength from existing work on co-located collaboration by transferring concepts such as coupling styles or territoriality into XVA settings.

7. Visualization and VA

In this section, we analyse 37 publications, which introduce or discuss VA systems, methods and techniques. Papers merely focusing on specific inputs, a specific interaction or transition technique, or papers performing mainly evaluations of aspects of an existing system, were considered as out of scope for this section. The 37 qualified papers and the stages in the RVC, which they operate in, are shown in Figure 14. In this section, we analyse the input data such systems are built upon, the respective visual metaphors they introduce and apply, together with the interaction methods employed with respect to the visual metaphors.

7.1. Input data

In terms of input data in visualization, a variety of taxonomies have been introduced. An early and widely used taxonomy regarding general input data for visualization was presented by Shneiderman et al. [Shn96] who distinguish data into 1D, 2D, 3D, temporal, multi-dimensional, tree and network data categories. As the analysed input data in XVA tends to be heterogeneous and multi-variate, we considered other taxonomies as well. Kehrer et al. [KH13] in their survey on VA of multi-faceted scientific data introduced their view on different facets of heterogeneous scientific data and pointed out, that although data and model scenarios are becoming multi-faceted, the heterogeneity given in spatial, temporal and multivariate data also represents novel opportunities. The classification of visualization techniques by Ward et al. [WGK15], which we use in Section 7.2, also reflects a data classification at its basic level, therefore we...
provide a common classification of XVA contributions on these two aspects in Table 5.

More than half (20) of the 37 qualified publications were analysing spatial or geospatial data, closely followed by 19 dealing with multivariate data. This was to be expected, as the visual analysis of spatial data, especially of 3D data, lends itself to an analysis in native 3D. Furthermore, visual analysis typically deals with multivariate data or tabular data. As for tabular data, applying immersive analysis with the help of multivariate visualization methods is an extension to classical 2D screen analysis, often requiring novel visual metaphors or interaction techniques. Nine systems also deal with visual analyses of network and tree data. All nine works in this area introduce visualization techniques for arbitrary graph data, while three also specialize on networks from communities; no system was found dealing with tree based data structures. Finally, in our survey, we did not find a single paper describing a tool that analysed texts and logs data using XVA. This might be due to the limited availability of XVA displays with sufficient resolution for larger documents, as only recently HMDs have become available. In addition, there seems to be no clear path yet of how additional dimensions and interaction concepts may provide benefits for this type of data.

### 7.2. Visualization techniques

We also classified the 37 qualified papers and respective XVA systems regarding the visual metaphors and the visualization techniques they employ. Table 5 summarizes our findings regarding visual metaphors. We decided to do the classification according to the categorization introduced by Ward et al. [WKG15], which differentiates visualization techniques firstly by the kind of input data, and secondly by the composition of the output. The remaining subsection describes these different visualization techniques, and the respective XVA systems which employ them.

#### Trees, graphs and networks visualization

Many research domains, such as social [KKM*20] or biological networks [KKTD17], require the analysis of trees, graphs and network data. Typical visualizations of such data use nodes to represent entities in the analysed system and edges to represent relationships between them. Generally, the selected papers apply force-based or hierarchical layouts to draw the visual representations for graphs and networks. These standard layout techniques optimize node coordinates in such a way that it minimizes edge crossings. However, the work by García-Hernández et al. [GAWK16] visualizes graphs in 3D space by placing nodes so that the distances between connected nodes are proportional to the measured similarity between them. In the papers that qualified for this section, we did not find specific visualization and exploration methods for trees. We obtained nine papers, which discuss network and graph visualization as well as respective interaction techniques. A recent study by Kotlarek et al. [KKM*20] evaluates networks with standard 2D and 3D immersive techniques of graph visualization and respective interaction. The study highlights that users perform better in various network exploration tasks such as structural interpretation when using 3D visualization. Various studies discuss the challenges of network visualizations (layouts) [KMLM16], exploration [SWKA19] and navigation [DCW*17, DCW*18, DCW*20] in immersive environments. A sense of orientation is required to be maintained during graph visualization in immersive platforms. The major challenges for navigation techniques are that they should be faster, less physically demanding and accepted by the users. Visualization layouts that highlight key structural properties and provide higher-level information are another challenge. Studies of such visualizations by Emmert et al. [ESTYHD18] and Tripathi et al. [TDE14] categorize visualization of complex networks into global, modular and hierarchical layouts. Kwon et al. [KMLM16] discuss graph layouts on a 3D platform showing graphs on a spherical surface, combined with an edge bundling technique for consistent viewpoints and graph structure comprehension which avoids complicated navigation.

**Table 5:** Data and visualization techniques in relevant VA systems.

<table>
<thead>
<tr>
<th>Sub-class</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees, graphs, networks (n = 9)</td>
<td>[CDK<em>17, DCW</em>20, EMP18, GAWK16, HFL<em>17, KKTD17, KKM</em>20, MAG<em>16, PLE</em>19]</td>
</tr>
<tr>
<td>Arbitrary (n = 9)</td>
<td>[HFL<em>17, KKTD17, KKM</em>20]</td>
</tr>
<tr>
<td>Communities (n = 3):</td>
<td>[ACC99, BSB<em>18, BJR20, BHM</em>18, CDH<em>19b, CCD</em>17, DDC<em>14, WFFN19, FVP</em>18, GAWK16, KWO<em>20, LCPD19, LPED20, NMT</em>19, PLE*19, RBLN04, WFRFN18]</td>
</tr>
<tr>
<td>Multivariate data (n = 19)</td>
<td>[ACC99, BSB<em>18, BJR20, BHM</em>18, CDH<em>19b, CCD</em>17, DDC<em>14, GAWK16, KWO</em>20, LCPD19, LPED20, NMT<em>19, PLE</em>19, RBLN04, WFRFN18]</td>
</tr>
<tr>
<td>Points (n = 17):</td>
<td>[ACC99, BSB<em>18, BJR20, BHM</em>18, CDH<em>19b, CCD</em>17, DDC<em>14, WFFN19, FVP</em>18, GAWK16, KWO<em>20, LCPD19, LPED20, NMT</em>19, PLE*19, RBLN04, WFRFN18]</td>
</tr>
<tr>
<td>Lines (n = 7):</td>
<td>[ACC99, BHM<em>18, CDH</em>19b, CCD<em>17, GAWK16, PLE</em>19, RH17a]</td>
</tr>
<tr>
<td>Region (n = 9):</td>
<td>[ACC99, BHM<em>18, CDH</em>19b, DDC<em>14, GAWK16, LCPD19, LPED20, NMT</em>19, RH17a, WWS19]</td>
</tr>
<tr>
<td>Geospatial data (n = 7)</td>
<td>[ACC99, BHM<em>18, CDH</em>19b, DDC<em>14, GAWK16, LCPD19, LPED20, NMT</em>19, RH17a, WWS19]</td>
</tr>
<tr>
<td>Point (n = 4):</td>
<td>[BJR20, NMT<em>19, PLE</em>19, SWS*19]</td>
</tr>
<tr>
<td>Line (n = 5):</td>
<td>[HRD<em>19, NMT</em>19, PLE<em>19, SWS</em>19, WFFN19]</td>
</tr>
<tr>
<td>Area (n = 1):</td>
<td>[WWS19]</td>
</tr>
<tr>
<td>Spatial data (n = 16)</td>
<td>[FP00, PKP*18]</td>
</tr>
<tr>
<td>2D (n = 2):</td>
<td>[ACC99, BSB<em>18, BJR20, FP00, GK19, GAWK16, JTS</em>13, KWWF20, LBS14, LBS13, NMT<em>19, PKP</em>18, RH17a, SB119, SSFG98, WWS19]</td>
</tr>
<tr>
<td>3D (n = 16):</td>
<td>[ACC99, BSB<em>18, BJR20, FP00, GK19, GAWK16, JTS</em>13, KWWF20, LBS14, LBS13, NMT<em>19, PKP</em>18, RH17a, SB119, SSFG98, WWS19]</td>
</tr>
</tbody>
</table>
Multivariate data visualization: Multivariate data are typically represented in a table, consisting of $n$ samples (rows) and $p$ attributes (columns). There are 19 papers in our review shown in Table 5, which specifically use multivariate visual metaphors in immersive platform settings. These techniques mainly build upon point or line-based visualizations. One of the earliest multivariate visual metaphors in an immersive environment by Arms et al. [ACC99] is based on scatterplots, which are used to evaluate user performance in terms of cluster identification. Using scatterplots for cluster identification on 3D immersive visualization platforms results in a better understanding of groups, which are spatially closer to each other. Even if the evaluation is only based on selection as interaction method, it highlights for the first time the importance of immersive visualization for understanding and exploring data. Cavallo et al. [CDH*19b] present a collaborative, hybrid analytic system for exploratory data analysis using high-resolution displays, table projections and AR representations of table data with multivariate visualizations. They study the users’ performance, investigate limitation of immersive technologies, and compare the insights gained through hybrid visualization versus an analysis fully immersed in VR. The ImAxes system by Cordell et al. [CCD*17], a wide range of visualizations can be created by adjusting the arrangements of the data axes. The flexibility of adjusting and manipulating axes allows the creation of different types of visualizations. Liu et al. [LPED20] explore ways of how to translate the small multiples paradigm, commonly used on desktop systems for the visualization of multivariate data, to an immersive environment (see Figure 15). They performed two user studies: The first explores the influence of different layouts on user performance, and the second investigates how well their shelves metaphor scales to more multiples. Butcher et al. [BJR20] present a web-based VR framework for IA. The framework enables platform-independent, multivariate data visualization based on declarative specifications and provides interaction and collaboration methods. They evaluate the performance of the system for various use cases. In addition, they point out common problems in web-based VR analytics, such as dropping frame rates with many data points, decreased performance when rendering text during labelling for 3D plots or stuttering when the rendered visualization is complex.

Spatial data visualization requires data to have an implicit or explicit spatial aspect. This often implies a direct mapping of the spatial attributes of the data to positions in the visualization [WGK15]. In all 16 qualified papers dealing with spatial data, the data were identified to be three-dimensional in nature. Most systems therefore employed 3D visualizations. We only found two contributions using 2D visualizations of spatial data, by providing slice views through 3D data [FP00, PKP*18]. Many employed a device close to the virtuality side of the RVC [BJR20, GK19, HRD*19] for the 3D visualization. The nature of volumetric data would seem to make it well suited for exploration in VR. However, few of the systems we investigated, which were using 3D volumetric data, actually used direct volume rendering techniques [LBS14, PKP*18]. Most use iso-surface mesh visualizations, allowing for fast rendering [GK19, PKP*18, RH17a]. Volume rendering on devices towards the virtuality side of the RVC has to meet even higher demands, as interactive frame rates are much more crucial here to avoid motion sickness. This could be one reason why volume rendering has not yet been widely adopted in this area. Szalavari et al. presented an early system analysing spatial data in a collaborative way [SSFG98]. Their system exclusively uses AR devices, but also enables the inspection of purely virtual objects. They focus on analysing simulation outcomes, even though most of their concepts could also be applied more generally to other spatial data scenarios. Regarding collaboration, they introduce a personal interaction panel concept, which enables the clear distinction between a collective viewpoint and some information only shown privately. In recent years, several cross-device analysis frameworks have been proposed, such as the One Reality framework by Roo and Hatchet [RH17a]. They present their six-layer framework ranging from completely physical to completely virtual and showcase it on spatial data from a volcano mock-up and a car engine. Their prototype supports the collaborative, cross-device analysis of these spatial datasets utilizing spatial augmentation, see-through displays and opaque HMDs. Kunert et al. present a system for collaborative analysis utilizing large wall and tabletop screens combined with HMDs [KWFK20]. They evaluate their system with data from prehistoric rock engravings as well as a virtual 3D city model and focused on evaluating the benefits of collaboration on large screens. Future work is likely to
go into directions as laid out by Gall et al. [GFH21]. They propose to use transitional interfaces in material science to explore spatial data along with derived non-spatial data.

**Geospatial data visualization** differs from data visualization by the fact that the spatial attributes of the data describe specific locations in the real world [WGK15]. Seven of the qualified papers describe systems for the analysis of geospatial data. All of these systems provide some kind of map visualization and combine it with additional visualization elements such as line, point and area markers. Ssin et al. [SWS*19] analyse ship route data, which is geospatial data combined with temporal information, as shown in Figure 15. The ship routes are gathered from multiple sources and are visualized as line strips. Additional detail in the form of space-time cubes as well as views from different perspectives on the same route can be overlaid via AR. They evaluate their system against a desktop-only system and even though they do not find an improvement of the task completion time, they are able to demonstrate reduced failure rates.

The FiberClay system by Hurter et al. [HRD*19] explores similar data, namely trajectory data from multiple domains. They collected qualitative feedback mainly from the analysis of air traffic and wind data, and they could show that their system could generate valuable insights and fostered engagement with the data. The system itself is only employing immersive visualization, yet provides interesting reflections on the comparison to 2D systems. The work on visual link routing by Prouzeau et al. [PLE*19] shows the potential for linking geospatial visualizations with multivariate visualizations. They present this mainly in an immersive context, but point out the importance of such linking in collaborative scenarios, and they show view-point dependent methods to achieve this. Whitlock et al. [WWS19] explore concepts for ‘immersive VA in the field’. They conducted interviews with experts in various fields performing such field data collection tasks. They implemented a design probe for utilizing handheld AR, combined with the visualization of abstract information linked to a location, to support the combined data collection and analysis. This promises more rapid analysis and decision cycles over the current practice of prior data collection followed by a separate analysis and decision making stage.

### 7.3. Interaction techniques

A further aspect we investigated in our survey were interactions of the user with the respective visualization techniques. We therefore looked for an interaction taxonomy which reflects both the intention of the user as well as the interaction with the visualization. Shneiderman et al. [Shn96] introduced a taxonomy of high-level tasks for visualization (overview, zoom and filter, details-on-demand, relate, history and extract). A more recent taxonomy of interaction techniques was proposed by Yi et al. [YaKJSJ07], which more closely targets our intentions. This taxonomy was originally introduced for information visualization and distinguishes seven interaction techniques: Select (mark something as interesting), Explore (show me something else), Reconfigure (show me a different arrangement), Encode (show me a different representation), Abstract/Elaborate (show me more or less detail), Filter (show me something conditionally) and Connect (show me related items). As this taxonomy covers the main interaction techniques in XVA, we considered the taxonomy of Yi et al. as optimal for our classification. Besançon et al. in their recent state of the art report on spatial interfaces also provide a classification of interaction techniques [BYK*21]. Their taxonomy is very interesting, but tailored for spatial data, which is why we did not use it here.

Our findings are summarized in Table 6 (García-Hernández et al. [GAWK16] is not listed in this table, since it describes multiple systems and no specific interaction techniques): Most systems (27 out of 37) provide a selection mechanism to pick and highlight interesting data items. An intuitive way to enable selection are tangible input devices such as in the GeoGate system [SWS*19] (see Figure 15). Similarly, a large portion of systems (28) also provide ways to explore the datasets of interest, allowing the user to navigate to different subsets of the data. An option to reconfigure was provided by 16 systems. Filtering capabilities were provided by roughly a third of the qualified contributions (13), and nearly as many provided a way to abstract and elaborate (11). With Fröhlich’s Cubic Mouse [FP00] for example, filtering is achieved by using their tangible input device to place a slice plane, which subsequently filters out data outside of the area of the plane. Similarly, Jackson et al. [JTS*13] utilize a paper prop in the shape of a tube to filter fibre data by the current alignment of the tube. Eight systems had an option to show different representations of the same data, denoted by the encode technique, for example, through cross-device combinations such as VR and tablets [SBI19] (see Figure 1). Only seven systems provided a way to show related items via a connect operation. This is surprising, as we would have expected that one advantage for cross-device analysis would be that the different devices can be used for different views on the data, and that the systems would provide ways to meaningfully connect between the different views. A prime example of how connect operations could be supported by systems is the visual link routing method by Prouzeau et al. [PLE*19], which links elements of different visualizations in 3D, and was evaluated on VR and AR.

### 7.4. Visualization and VA challenges and recommendations

As the analysis above indicates, employing visualization techniques in a collaborative, cross-device setting shows promise to solve many visualization tasks better than conventional systems. Visualization and interaction in XVA systems also come with new challenges, however:

1. **Technical challenges**: XVA depends on adequate hardware and software support for interactive and efficient analysis. Volume rendering on the virtuality side of the RVC is one example where optimization of hardware and software is required for enabling a more widespread usage. We also see a need for web-based frameworks and libraries, which can be accessed from scripting languages such as Python or javascript, for an easy entry into developing XVA systems. For IA, similar frameworks have recently started to appear [BJR20]. Much work is however still to be done in extending these for an analysis across multiple stages in the RVC.

2. **Missing guidance on visualization and interaction methods**: The effectiveness of visualization and interaction techniques across different stages is not yet well evaluated. Some work has been started in this direction, for example, investigations...
regarding 2D scatter plots on desktops versus 3D scatter plots in VR [WFRFN18], or comparing scatter plots in a desktop setting versus an AR setting [BSB*18]. Such literature, however, is currently sparse, and guidelines regarding which data best to visualize using which visualization and in which stage are not available. Also, most qualified papers only utilize a small subset of the seven types of interaction techniques from our classification. Therefore, we suggest systematic studies regarding the effectiveness of interaction and visualization techniques on and across different stages to establish a set of general guidelines.

3. Space utilization across the RVC: Many visualization techniques have constraints on space utilization for visualizing a large dataset or multiple views, which directly affects the users’ performance on data exploration tasks. One way to address this issue are specialized visualization techniques adapted for the respective stage in the RVC, such as the spherical graph layouts by Kwon et al. [KMLM16]. Additionally, visual representations of summarized data, such as miniature models as in the Worlds-in-Wedges by Nam et al. [NMT*19], redundancy removal, and data abstraction by interaction and space layout design (such as in the work by Liu et al. [LPED20]) can be useful approaches in this respect. Not much work has however yet been done on reconciling the different space constraints of devices across the RVC.

4. Consistent visualization and interaction techniques across devices: Visualizing multi-modal, multivariate and multi-channel data in XV platforms requires choosing appropriate visualization encodings and interaction techniques for each device, as the platforms differ in capabilities (for example, a screen only has two native dimensions, while VR comes with three; colours in AR can vary when merged with underlying colours from the background). For the same data viewed on different stages in the RVC or in a collaborative setting, visualization and interaction techniques should match as closely as possible between the different scenarios to provide a consistent user experience. Such translations of encodings and interaction techniques between devices on different stages in the RVC are an open challenge.

8. Evaluation for XVA

Our analysis of evaluation methods and user studies is based on 19 papers that provided sufficient detail on some form of evaluation and were concerned with XR [BKP01, CDH*19b, ESE06, RBCH18], XV [BSYB20, GDM19, GLB05, KWO*20, KRD16, LBS14, LBS13, LCPD19, MAG*16, NMT*19, RH17a, WWS19], transitional interfaces [MBHM17, MBH171, MBR20, SWOG17]. From this, three evaluations were excluded as they took place with too informal study designs or settings, or did not provide a structured evaluation involving users [BKP01, CRHG17, RH17b]. The remaining 16 papers are discussed in this section.

8.1. Study designs

With regard to their study designs, 15 papers used controlled experiments in a laboratory setting to compare different designs, systems or technologies using quantitative data (see Table 7 for an overview). For these quantitative comparisons, ten papers used objective measures, such as task times and completion rates, and all 15 included subjective data from self-assessment questionnaires, such as standardized or custom questionnaires on task load, presence, immersion or similar (more details in Section 8.3). Unlike the 15 papers mentioned above, Sisto et al. [SWOG17] did not compare different conditions but rather evaluated whether participants noticed changes in the environment while being involved in a task unrelated to the environmental changes. In addition to a quantitative approach, five papers used qualitative methods, such as semi-structured [NJ19, PDE*19] or unstructured interviews [ESE06, GLB05] as well as the think-aloud method [CDH*19b, ESE06], to gather feedback on the participant’s experience. However, no papers used formal qualitative research and analysis approaches such as thematic coding.

Given this dominance of controlled and quantitative lab studies, future XVA evaluation could greatly benefit from a greater breadth of evaluation methods, starting with qualitative research as a rich source for a deeper understanding of users’ needs and concerns in early phases of the design (as used by Eissele et al. [ESE06]) and

<table>
<thead>
<tr>
<th>Interaction technique</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Select ( (n = 27) )</td>
<td>[ACC99, BJR20, BHM<em>18, CDH</em>19b, CCD<em>17, CDK</em>17, DCW<em>20, EMP18, HFL</em>17, HRD<em>19, KKD17, KRM</em>20, KWO<em>20, KWO</em>20, KWF<em>20, LCPD19, LPED20, MAG</em>16, NMT<em>19, PKP</em>18, PLE<em>19, RBNL04, RH17a, SWS</em>19, WFN19, WFRFN18, WWS19]</td>
</tr>
<tr>
<td>Explore ( (n = 28) )</td>
<td>[BJR20, BHM<em>18, CDH</em>19b, CCD<em>17, CDK</em>17, DCW<em>20, EMP18, EVP</em>18, FP00, GK19, HFL<em>17, HRD</em>19, JTS<em>13, KRD16, LBS14, LBS13, LPED20, NMT</em>19, PKP<em>18, PLE</em>19, RBNL04, RH17a, SBI19, SWS*19, SSFG98, WFN19, WFRFN18, WWS19]</td>
</tr>
<tr>
<td>Reconfigure ( (n = 16) )</td>
<td>[BHM<em>18, CDH</em>19b, CCD<em>17, CDK</em>17, DCW<em>20, EMP18, HFL</em>17, HRD<em>19, JTS</em>13, KKD17, KRM<em>20, KWO</em>20, LCPD19, SBI19, SWS*19, SSFG98, WFN19, WFRFN18, WWS19]</td>
</tr>
<tr>
<td>Encode ( (n = 8) )</td>
<td>[CDH<em>19b, DDC</em>14, HRD<em>19, JTS</em>13, KKD17, PLE*19, SBI19, SSFG98]</td>
</tr>
<tr>
<td>Abstract/elaborate ( (n = 11) )</td>
<td>[BJR20, CDH<em>19b, DDC</em>14, EMP18, HFL<em>17, KKD17, LCPD19, MAG</em>16, NMT*19, RH17a, WWS19]</td>
</tr>
<tr>
<td>Filter ( (n = 13) )</td>
<td>[BSB<em>18, BJR20, BHM</em>18, CCD<em>17, DCW</em>20, EMP18, FP00, HRD<em>19, JTS</em>13, KKD17, LPED20, SWS*19, WFN19]</td>
</tr>
<tr>
<td>Connect ( (n = 7) )</td>
<td>[BJR20, BHM<em>18, CDH</em>19b, CCD<em>17, KKD17, MAG</em>16, PLE*19]</td>
</tr>
</tbody>
</table>
by increasing external and ecological validity in later stages through studying real-world applications and domains with 'in-the-wild' deployments [RM17]. This also resonates with the findings of Besançon et al. who mention that the degree to which domain experts can include a tool into their workflow is much more important than error rate or task time [BYK*21].

8.2. Study participants

The mean number of participants in the analysed user studies was 24 (SD = 15, min = 10, max = 80). Only 14 of 16 papers reported gender distribution with an average percentage of female participants of 31% (SD = 13%, min = 11%, max = 60%). Of the 11 papers that reported recruitment strategies or the background of their participants, nine papers mostly or exclusively used members of their university as participants. Only Cavallo et al. [CDH*19b] recruited experts from the target domain of data science for their study, and Piumsomboon et al. [PDE*19] recruited participants from the general public. Only eight papers reported mean age with an overall mean of 27.1 years (SD = 3.4, min = 21.6, max = 32.6).

In light of the close ties of XVA evaluation to human psychology and behaviour, future studies should be careful about not underestimating variations across human populations, which have been shown to affect even low-level visual perception [HHN10]. They should avoid generalizing findings from WEIRD ('Western, Educated, Industrialized, Rich and Democratic') samples [HHN10] and work towards equal gender distribution, representative age distribution, and focus either on participants from the general public for better generalizability, or on domain experts for external and ecological validity.

8.3. Measurement of MR-specific constructs

In addition to their usability, MR systems were typically evaluated in terms of MR-specific constructs such as (i) simulator sickness, (ii) cybersickness and (iii) presence using questionnaires.

1. Three out of 16 papers [KWFK20, MBR20, VF17] reported an evaluation of simulator sickness using the Simulator Sickness Questionnaire (SSQ) by Kennedy et al. [KLBL93].

2. One paper reported evaluating cybersickness by using custom questions in a feedback questionnaire [PC17], and one paper reported questions on comfort and sickness, where sickness was not specified further, also using a custom questionnaire [NJ19].

3. Two papers [MBHM17, NJ19] used the Igroup Presence Questionnaire [SFR01], two other papers [HL19, SBH*09] used the Slater–Usoh–Steed Presence Questionnaire [SUS94] and one [BSYB20] used the MEC Spatial Presence Questionnaire [VWG*04].

Furthermore, one paper [VF17] reported evaluating the factors attention/absorption, internal/external correspondence and reality judgement using the Reality Judgement and Presence Questionnaire by Banos et al. [BBGP*00].

It is noteworthy that measuring MR-specific constructs was based almost exclusively on questionnaires. Given the many practicalities of using questionnaires in VR [PAP*20], future XVA could benefit from exploring physiological measurements as potential enhancements or alternatives to current practices. This also resonates with a recent analysis of cybersickness research by Caserman et al., who identify the SSQ as the golden standard but also mention the possibility of using biosignals to detect cybersickness [CGAZG21]. For example, Steinicke et al. used physiological measurements of skin temperature, breathing rate, heart rate and skin conductance to assess the participant’s level of immersion in a flight simulation [SBH*09].

8.4. Measurement of cognitive load

VA are thought to amplify cognition by exploiting the human eye’s broad bandwidth pathway into the mind to let users see, explore and understand large amounts of information [TC05]. Therefore, it is important to consider and measure users’ cognitive load during the use of XVA systems. Consequently, five papers [BSYB20, GDM19, NJ19, PC17, RBCH18] reported results from the NASA Task Load Index (NASA TLX) [HS88] with two of them [PC17, RBCH18] stating that they used the NASA TLX specifically to measure cognitive load. However, the NASA TLX is not designed to assess cognitive load but rather indicates a general workload with several subscales, including mental demand, which is not equivalent to cognitive load. Moreover, the NASA TLX is not perfectly equipped to specifically analyse mental workload either, as McKendrick and Cherry conclude in their in-depth analysis of the NASA TLX [MC18].
An alternative could be using thermal imaging cameras to measure cognitive load by recording temperature differences between a person’s forehead and nose [AVD*17]. Cho and Bianchi-Berthouze provide a comprehensive overview of thermal imaging technology for measuring physiological and affective needs [CBB19]. However, HMDs for MR usually occlude the forehead and nose from thermal cameras, limiting their use for XVA. Other approaches could include the analysis of eye tracking data for physiological measurement of cognitive load [PKSH10, ZPR16b, PL19, FLGL15]. This approach could especially be useful in XR scenarios since eye tracking is available in commercial HMDs such as the HTC Vive Pro Eye [HTC21], FOVE [FOV21] and Varjo HMDs [Var21]. More recent research uses functional near-infrared spectroscopy to examine the sensemaking process [GSL21].

8.5. Statistical data analysis

Three out of the 16 papers, which chose a quantitative approach for their user study, only reported descriptive statistics of their findings [CDH*19b, ESE06, SWOG17]. The other 13 papers used popular methods of inferential statistics such as ANOVA or t-test. This is noteworthy, since many of the psychological phenomena studied in human–computer interaction cannot be measured directly and are therefore considered latent variables. Researchers have called for utilizing more advanced statistical methods that actually treat these phenomena as such indirectly measured variables [BM16]. One method that enables more complex analytical constructs is structured equation modelling (SEM). This method enables researchers to look at latent connections and causal relations in a greater theoretical framework [McI12]. For example, SEM models could provide a more extensive method to investigate cognitive load combining questionnaire data, physiological measurements of the skin with optical sensors and eye tracking data [FLGL15, PL19].

8.6. XVA evaluation challenges and recommendations

Our analysis identified many possibilities for improvement that are not specific to XVA evaluation but generally true for user research in information visualization and human–computer interaction (for example, more qualitative and mixed methods, better study designs and user samples). With regard to XVA there are, however, two critical challenges that we propose to be addressed in future work:

1. There is not yet a theoretical construct that could serve as an equivalent to VR’s presence construct and to evaluate the quality, mental demand, and plausibility of interacting, collaborating and transitioning within XV. The development and operationalization of such a construct will be necessary for meaningful XVA evaluation.

2. With regard to the practicalities of using questionnaires in VR or XVA settings, physiological measurement of latent variables such as presence or cognitive load should be explored as enhancements of or alternatives to self-reported data.

9. Application Domains

We were particularly interested in analysing the application domains where XVA techniques are employed. A dedication to a specific domain can be considered as a strong statement about the maturity of XVA in the respective area. However, the majority of core-relevant papers did not point out a specific application domain (92 out of 118), but rather build upon generic data. It was only possible to link a total of 26 of these papers to a particular domain. Many publications do not explicitly mention an application domain, such as the work on visualization techniques for graphs by Erra et al. [EMP19] or the work of Kunert et al. on combinations of different tools [KWFK20]. In our survey, we determined that biology and medicine was mentioned most in terms of the targeted domain aside production and supply chain applications. Less frequently, we found other application areas such as education and training, and material science. Specific examples of these application domains are explained in the following paragraphs.

In the domain of biology and medicine, XVA is used to correlate and visualize patient data before and after surgeries with the overarching goal to optimize the patient’s treatment, as presented by Pfeiffer et al. [PKP*18]. The heterogeneous medical data used in this application (such as textual, image or temporal data) are integrated in a virtual environment framework for surgical applications such as preoperative planning. Here, the authors introduced an approach using HMDs in different phases of the medical treatment. Surgeons are using HMDs to analyse the available patient data for advanced preoperative planning. They can combine information about past surgeries or medication in VR and thus find relations and changes in the data. Another approach for preoperative planning is their 3D visualization of 2D slice images from CT scans. Surgeons profit from an immersive 3D data representation as they can explore complex situations in a natural and intuitive way, such as interaction with data when exploring tumours very close to veins or arteries.

In the production domain, XVA is used in a similar way: For example, Utzig et al. [UKAG19] presented XVA techniques for aircraft maintenance. Service personnel at the airport receive instructions and walk-throughs for upcoming maintenance tasks through AR HMDs. In addition to guidance regarding tasks, information required for each task is displayed in the field of view of the service personnel. The inspection task can be automatically documented through this system. If an operator discovers indications or obvious damage at the test specimen, respective locations can be labelled. The virtual label is linked to the position on the test specimen and may be retrieved for the following inspection. Besides visualization, complex parts can be explored in depth, combined with assembly information, which facilitates inspections and repair processes. It is also possible to connect with other maintenance experts to support the service personnel in a joint session also when located in different sites.

Regarding education and training, first steps are made in the direction of XVA in terms of on-site inspections of pipes and, more specifically, on welds. In this context, Amza et al. [AZPT17] presented an AR-based visualization approach using handheld devices for training operators to use ultrasonic equipment for the non-destructive testing (NDT) of pipework. For this application, it is of primary interest to provide workers with information on the steps they have to perform in terms of testing. Furthermore, the main components of the ultrasonic equipment are explained together with the proper handling and placing of the equipment, in order to ensure...
suitable alignment and movements on the pipe and on the weld. Collaboration across devices in the sense of XVA was not targeted, but may be a step towards future work also in this domain; for example, it could be used for requesting supervisor knowledge in case of unclear testing results.

In material science, understanding and interpreting multidimensional volumetric data are of high interest for the analysis of complex and heterogeneous material systems. Gall [Gal20] presented a VR-based technique for the visual analysis and exploration of fibre reinforced composite materials. Here the focus was put on investigations of novel visual metaphors such as a model in miniature, exploded views for partitioning the data, histograms and node-link diagrams. Using these visualization techniques coupled with intuitive interaction techniques, unique insights into complex materials were generated. Also here remote collaboration over long distances in the sense of XVA is seen as opportunity. In addition, situated analytics in order to augment the physical world with additional abstract data was denoted as interesting new area.

10. Challenges and Opportunities

In contrast to the individual challenges of aspects of XVA discussed above, we also synthesized our findings to challenges for XVA on a higher level and analysed our results regarding gaps in literature. Summarizing these efforts, we list here what we have identified as main open challenges and opportunities for future XVA research:

1. **Device integration**: The wide variety of devices utilized in collaborative XVA applications requires seamless interaction between those devices to minimize disruptions when interacting between different stages in the RVC. Some initial work in this direction has been done, for example, regarding generic network integration [RS99] or using one input pen across multiple displays [HRG04]. The context of XVA requires specific considerations, however. Ideally, single devices can be used at multiple stages in the RVC, to avoid hard transitions when switching device. In the area of devices, technical issues such as a constrained field of view, restricted mobility and limited hardware capabilities still have to be addressed.

2. **Transition techniques across the RVC**: A critical challenge is the design and implementation of a seamless transition without disorienting and without disrupting the workflow throughout the RVC. Further research is necessary to understand how such a transition needs to be designed to best support the user.

3. **Design of XVA metaphors along the RVC**: The scope of work evaluating specific visualization techniques across different stages in the RVC so far is very limited and leaves large gaps to be explored. What we envision is a set of guidelines and ideally a framework regarding which data best to visualize using which specific visualization techniques in which particular stage. Such frameworks and guidelines also need to address the consistency of visualization and interaction techniques across devices.

4. **Collaboration across the RVC**: Successful collaboration in XVA requires sufficient awareness between all team members, even when they are distributed across different locations and different stages in the RVC. Current research has only proposed specific techniques for awareness cues for collaboration within VR, between VR and AR, or between 2D displays and AR. There is yet no generalizable framework for suitable XV awareness cues across the entire RVC. This will be necessary to enable efficient communication and spatial interaction in XVA and for identifying and dynamically providing suitable multi-modal awareness cues (such as spatial audio, gestures or eye gaze) and user representations.

5. **Application integration**: As current visualization frameworks typically target only specific devices, tailored software frameworks and toolkits are required, especially for the support of visualization and interaction techniques in a collaborative and cross-device analysis. To make experimentation with XVA systems easier, such frameworks need to be easily accessible, for example, via scripting language interfaces.

6. **Quality models**: By letting users frequently move and collaborate between different stages in the RVC, XVA creates entirely new challenges for designers and users alike. Therefore, there are yet no XVA-specific quality models that could help researchers to focus their efforts on the most important aspects during design, technological and implementation-related choices or evaluation. Quality models would help researchers to explicitly prioritize contributing factors for XVA success or failure, such as the users’ sensation of plausibility, coherence and consistency of virtual elements after XV transitions (similar to presence in VR), users’ mental demand or cognitive load for (collaboratively) navigating the work environment, system performance (such as frames per second, latency of visualization and interaction) or other cybersickness-related sources of discomfort.

11. Conclusion

We provided a detailed disambiguation of the novel field of XVA from other related areas. In short, XVA systems enable visual analysis on a combination of devices along the RVC. They do this by employing transitional interfaces, potentially coupled with collaborative methods, and within the RVC stages, along with suitable visualization and interaction metaphors. Our literature survey revealed that not many systems exist so far which cover all these core aspects of XVA. As 14 reveals, only two out of the 37 relevant VA systems actually interconnect users from different stages in the RVC, and only one allows true shifting across the RVC. We also analysed the existing body of work regarding separate aspects of XVA and provided recommendations and challenges on each of these aspects as guidance for the development of future XVA systems: We explored the different levels of XV, that is, the ways how devices at different stages in the RVC can interconnect, techniques for transitioning between different stages in the RVC, as well as for collaboration between them. We analysed visualization and interaction techniques for XVA and looked at evaluation done so far on XVA. Finally, we analysed application domains in which XVA systems are starting to appear. Through our literature survey, we identified gaps in the existing research and condensed these findings into what we see as the current high-level challenges in XVA.
Acknowledgements

We would like to thank the reviewers for their very detailed feedback and constructive criticism, which has helped in shaping the survey into its current form. This work is part of X-PRO. The project X-PRO is financed by research subsidies granted by the government of Upper Austria.

References


[BFR17] Fröhler et al. / Cross-Virtuality Analytics 485

© 2022 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd


[HS88] Hart S., Staveland L.: Development of NASA TLX (task load index): Results of empirical and the-


[KTY99a] Kiyokawa K., Takemura H., Yokoya N.: A collaboration support technique by integrating a shared virtual reality and


© 2022 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd


© 2022 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd


Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information

© 2022 Eurographics - The European Association for Computer Graphics and John Wiley & Sons Ltd