Stepping into the Unknown: Immersive Spatial Hypertext

Lisa Eidloth*
Hof University
Institute for Information Systems
Hof, Germany
lisa.eidloth@iisys.de

Claus Atzenbeck
Hof University
Institute for Information Systems
Hof, Germany
claus.atzenbeck@iisys.de

Thies Pfeiffer
University of Applied Sciences
Emden/Leer
Electrical Engineering and
Informatics
Emden, Germany
thies.pfeiffer@hs-emden-leer.de

ABSTRACT

Traditional spatial hypertext systems, predominantly limited to two-dimensional (2D) interfaces, offer limited support for addressing long debated inherent problems such as orientation difficulties and navigation in large information spaces. In this context, we present opportunities from interdisciplinary fields such as immersive analytics (IA) and embodied cognition that may mitigate some of these challenges. However, while some research has explored the extension of spatial hypertext to three dimensions, there is a lack of discussion on recent advances in virtual reality technologies and related fields, and their potential impact on immersive spatial hypertext systems. This paper addresses this gap by exploring the integration of immersive technologies into spatial hypertext systems, proposing a novel approach to enhance user engagement and comprehension through three-dimensional (3D) environments and multisensory interaction.

CCS CONCEPTS

• **Human-centered computing → Hypertext / hypermedia**; **Virtual reality**; **Mixed / augmented reality**.

KEYWORDS

hypertext, spatial hypertext, knowledge, extended reality, virtual reality, immersiveness, information exploration

ACM Reference Format:
Lisa Eidloth, Claus Atzenbeck, and Thies Pfeiffer. 2024. Stepping into the Unknown: Immersive Spatial Hypertext. In *7th Workshop on Human Factors in Hypertext (HUMAN '24), September 10–13, 2024, Poznan, Poland*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3679058.3688632

1 INTRODUCTION

In 1996, Rosemary Simpson described a hypothetical approach to spatial hypertext in a fully immersive virtual environment [49]. To date, we have not seen an implementation of this paradigm. Nonetheless, the envisioned system incorporates ideas that remain highly relevant today, particularly in light of ongoing technical advancements and emerging fields such as *immersive analytics* (IA) and *embodied cognition*.

IA is an interdisciplinary subfield situated at the convergence of *information visualization, human-computer interaction* (HCI), and *extended reality* (XR). Similar to *visual data mining* (VDM), IA aims to improve the quality and efficiency of analyzing large datasets, deepen comprehension of data and support decision-making by transforming data into graphical attributes [15, 38]. In contrast to VDM, which adopts an allocentric view of data and a focus on object-to-object relationships from a stationary perspective, immersive analytics employs an egocentric spatial mode of analysis through the use of *virtual reality* (VR) or *augmented reality* (AR) [38].

Increasing presence and natural interaction are the focus of VR. Therefore, much effort is put into the accurate mapping of sensory stimulation, avatar representation, and motion control. Embodied cognition, which emphasizes the human body as a significant part of cognitive processes, is also a crucial aspect of virtual reality research today, particularly in the study of presence and engagement [8, 13, 47].

We want to discuss another paradigm in light of these emerging technologies: spatial hypertext. Before we do so, let us consider contemporary software products designed for 2D interfaces, such as those found on desktop or tablet devices. When these products are adapted for immersive VR or AR environments, many companies take a rather naive approach: transferring systems from a 2D to a 3D environment is often seen as a simple projection of what users are accustomed to. New applications designed for VR or AR often fail to impose semantics onto the third dimension, apart from this dimension being the expression of spatiality.

Similarly, we could aim to simply place spatial hypertext on 2D planes within a 3D scene, such as virtual whiteboards or walls. However, this approach is akin to creating a movie using a static camera position to capture a scene on a theater stage, rather than employing a dynamic camera capable of movement and zooming, as we are accustomed to in modern films. With a dynamic camera, filmmakers can achieve much more than what was conceivable in the early days of moving pictures. Now, we face a similar situation with spatial hypertext, a concept well-established in the hypertext community, in light of emerging VR or AR technologies.

The question is not how we could mimic spatial hypertext in VR or AR environments. Instead, we ask how a *native immersive spatial hypertext*—that is, a hypertext that natively utilizes VR or AR attributes as essential features for expressing associations—could be imagined. This is a step into the unknown, as there are no reference implementations that are based on such a new structural domain rethought from the ground up.

*Corresponding author

HUMAN '24, September 10–13, 2024, Poznan, Poland

ACM ISBN 979-8-4007-1120-6/24/09
https://doi.org/10.1145/3679058.3688632

In this paper, we develop criteria that allow us to assess features with respect to their added value in the context of hypertext. This is a first step towards designing immersive spatial hypertext as a native hypertext structural domain, alongside others such as navigational hypertext [11], taxonomic hypertext [40], argumentation supporting structures [12], or 2D spatial hypertext [48].

The rest of the paper is structured as follows: In Section 2, we present a user scenario, followed by a discussion of relevant work in Section 3. We then delve into immersive analytics in Section 4, embodied cognition and natural interaction in Section 5, and immersive spatial hypertext in Section 6. Our work is concluded in Section 7.

2 USER SCENARIO

In the following, we will describe a scenario of an envisioned immersive spatial hypertext system. The system incorporates some aspects found in the field of VR, IA, artificial intelligence (AI) and embodiment, as well as spatial hypertext systems with recommender functionality, akin to the system described in [43].

2.1 Harnessing Spatial and Episodic Memory

Alice enters her personal virtual information space to continue her research. She walks down the VR hallway and enters the room she has named 'political discourse journey'. It is filled with objects, including concrete ones like plants, a chair, and a cup, as well as many abstract rectangles with text representing information nodes. Some rectangles float above a table, some are affixed to a wall, and others appear scattered around the room. Alice arranged the room herself over the past few days while exploring the topic of political discourses. As she explores, Alice hears sounds emanating from some of these objects and recalls the information she marked as most relevant yesterday for her current interests.

2.2 Embodied AI-supported Exploration

With a new idea in mind, Alice decides she would like to see more information about these objects. She puts on a coat hanging on a coat hanger and writes some keywords on a piece of paper. Immediately, the recommender system is triggered, and the surrounding space begins to change. New information nodes appear around her. She decides to dive deeper by attaching some of them to her avatar. The room changes again. She explores her information space by walking around. Some pieces of information float towards her, while others move away, depending on her movement in the space. She stops and decides to freeze the current state by snapping her fingers. The room stands still. Alice moves on. The information nodes she had attached to herself reappear in their previous form and remain where she was when she snapped her fingers. Alice decides to lie down on the floor. As she explores her space from this perspective, she recognizes possible associations of information snippets that she hadn't noticed before: some of the information nodes now seem closer together than they did when she was standing. She rotates the room by moving her hand to get a different perspective. To get a better view of the data, she makes a pinch gesture, causing the room and all the information entities to shrink. Alice notices a cluster of information that appears to be an outlier from the rest of her information space. As she zooms in and explores the content of this cluster, she remembers a conversation she had with her friend Bob yesterday. They were talking about politics and trying to understand why they disagreed. She cannot remember the details of this conversation, but she now feels that Bob should look at what she found in her room, so she calls him in.

2.3 Co-Presence in Shared and Private Spaces

Bob enters her room as his avatar. She directs him to the cluster of information and explains to him how this relates to her opinions on certain political issues. Bob realizes that he had never thought about the issue from this perspective. "I'd like to show you some of my information", he says, gesturing with his hands. Bob activates his personal information space about politics. New rectangles float into the room. As they both explore the newly created space, they realize that there is some agreement in their opinions. Apparently, the two information spaces overlap and complement each other in ways that neither of them had realized during their conversation yesterday.

3 RELATED WORK

Research on applying opportunities stemming from a 3D environment and (immersive) virtual reality technologies in the field of spatial hypertext has been discussed in the past, although there has not been a concept or implementation of a 'native' immersive spatial hypertext system. A few recent examples have studied linking and annotating entities in virtual environments (cf. VAnnotatoR [33], Va.Si.Li-Lab [1]). Earlier work discussed the spatial metaphor in the scope of a "virtual project room", where nodes of a spatial hypertext are arranged in a 2D virtual workspace (i.e., on a table) [37]. The *Information Visualizer* is an experimental system that uses interactive 3D rooms as a concept for hypertext workspaces to increase and densify the visible space where information is presented to the user [9]. A crucial point of argumentation in this context is that information in a workspace does not necessarily have to be captured at a glance but rather explored. This will be laid out in more detail in Section 6.1.1 and Section 6.2.1. Viki LibraRy is a project that explores the generation and exploration of online information in VR by utilizing spatial hypertext metaphors [2]. Finally, although hypothetical, we recognize Rosemary Simpson's deliberations (cf. Section 6) as close to what we would today recognize as 'native' immersive spatial hypertext system [49]. It is also worth considering other fields, such as Immersive Analytics and embodied cognition, as beneficial for gaining a better understanding and more detailed insights into design choices, challenges, and possible benefits stemming from VR and immersion. The following sections will provide an overview of these topics.

4 IMMERSIVE ANALYTICS

As IA is a highly interdisciplinary domain, it integrates approaches from different fields. In this section, we introduce some of the opportunities arising from IA that are relevant for potential immersive spatial hypertext systems.

An obvious opportunity of the IA approach is subsumed under the term *spatial immersion*, which refers to using a third spatial dimension to either expand the information space or as an additional encoding channel for data variables [44]. This includes viewing data from multiple perspectives and arranging objects in three dimensions [25]. In addition to the use of a third dimension, immersive attributes such as multisensory representations can be used to encode data variables. Embodied exploration, facilitated by auditory and haptic engagement, is another key aspect of this approach [7, 25]. The goal here is not necessarily to mimic real-world experiences, but to use the user's body as an interaction device and the user's senses as receivers of multichannel encoded data.

Situated analytics—another well established research focus in IA—describes the display of visualizations in direct reference to (real-world) objects [16, 27]. Work in this field includes displaying visualizations outdoors [45], enhancing real-world objects with real-time data [51], or supplementing 2D visualizations [10, 44]. Also, the aspect of facilitating spatially distributed collaboration through virtual representation of data and agents (*tele-immersion*) has been discussed for several decades, and its versatile applications also include IA [42, 44, 54].

The versatility and multidisciplinary nature of the field of IA and its applications can provide important insights into issues that arise when considering potential implementations of spatial hypertext in a virtual immersive environment. However, as IA is a relatively young field of research, many questions remain unanswered. Despite the idea that multimodality is beneficial in IA, there are still numerous unknown variables and possible challenges, such as conflicting stimuli, sensory crosstalk, perceptual overload, and cognitive overload [32]. Methodological recommendations to address potential challenges will be developed [17]. However, the multitude of possibilities and potential design factors for IA systems leads to high complexity due to the interdependence of different design choices [44].

5 EMBODIED COGNITION AND NATURAL INTERACTION

An underlying proposition of embodied cognition and the aim to enable natural interaction in virtual reality applications is that the human body plays a constitutive role in cognition, not just a causal one [47]. Based on this premise, Costa (2013) concludes that an avatar with limited functionality hinders the exploration of a virtual environment [13]. On the other hand, an avatar that is not congruent with the functionalities of the human body is not necessarily an impediment, but can provide a greater range of interactions and enhance the user experience. Inhabiting an avatar that is significantly different from a human body (e.g., an animal) is of interest in the field of homuncular flexibility and is partly based on the assumption that humans are *tool users* [56]. For example, Berti and Frassinetti proposed a cognitive difference in mapping near and far with and without using a stick as a tool [6].

Combining the premises that embodiment is a constitutive part of interaction in VR and that humans are able to go beyond what is naturally given to some extent (tool use), has significant implications for immersive spatial hypertext systems. In VR, a user can thus be part of a spatial hypertext, embodying an information entity (i.e., a hypertext node) and potentially switching between the roles of constituent and observer. This can open up new perspectives on the content of spatial hypertext and also promote a deeper understanding of the information space.

6 IMMERSIVE SPATIAL HYPERTEXT

In 1996 Rosemary Simpson envisioned a virtual "personal working information farm" that is operated by body and eye movements, speech, and touch gestures [49]. The system would also be used for virtual synchronous and asynchronous remote collaboration. It would use a multisensory representation of data: recordings would complement the information that the user sees in his personal hypertext. The workspace would include a virtual drawing board and a full-size virtual room that could be transferred to different locations (e.g., different virtual buildings, personal rooms of colleagues), depending on the information needed. Also, virtual objects that could be placed around the room, would be used for mnemonic purposes or as gateways to other virtual locations. Furthermore, multiple agents would be part of the system, gathering information from the Web and from global or personal knowledge bases. The retrieved information would then be arranged in a dynamically determined structure that could be modified by the user. Different views (i.e., visualizations/arrangements of information nodes) would also be provided on demand. This envisioned approach to a system that may be called immersive spatial hypertext includes several aspects that reflect opportunities currently being discussed in research fields such as immersive analytics (cf. Section 4). We argue that these opportunities also need to be discussed in the field of spatial hypertext, as their implementation has the potential to enhance this field by increasing user engagement, opening new perspectives on topics, such as orientation, spatial parsing, and using a recommender functionality not only for proposing additional information but also suggesting a user's movement in space, and enabling new forms of collaboration.

However, the implementation of immersive spatial hypertext is still approached with caution. This caution may be due to the wide variety and interdependence of design choices and the lack of a methodological framework. Additionally, it is uncertain how inherent characteristics of spatial hypertext systems correlate with new design choices stemming from a virtual and immersive environment. What we envision as an immersive spatial hypertext may increase complexity on the system side, the user side, or both, which must be considered when designing such a system. In the following section, the characteristics of traditional spatial hypertext and their impact on the concept of a potential immersive spatial hypertext will be laid out and discussed. Furthermore, new possibilities emerging from an immersive virtual environment will be explored.

6.1 Spatial Hypertext Design Choices

A foundation for the implementation of a third dimension and potential immersion in spatial hypertext lies in the careful consideration of the transfer of design choices inherent to spatial hypertext.

Discussion of potential intricacies must be an essential part of this consideration. The following is an outline of these issues.

6.1.1 Perspective. Besides parsing structures based on attributes such as color or the arrangement of nodes in predefined structures such as lists, an essential aspect of spatial hypertext is the assignment of semantics to the proximity of information entities placed in a 2D workspace. However, a 3D representation of this information space results in multiple possibilities for interpreting proximity, as

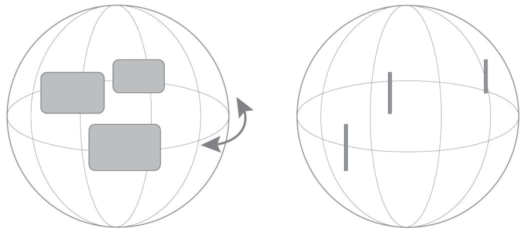

Figure 1: Three-dimensional representation of spatial hypertext nodes viewed from a 0° and 90° angle

the user alters the view point through head or body movement. The proximity assessment from one view angle must not be concordant with its assessment from another angle (cf. Figure 1). In part, the interpretation of proximity in a third dimension might be mitigated by employing depth cues, such as relative size or occlusion. Nevertheless, an exact perception of the proximity of information entities will only be achieved by the user altering his position.

Coming from a 2D spatial hypertext with an 'at a glance' perceivable proximity estimation, the additional cognitive effort required for 3D spatial hypertext might seem challenging at first. However, from a human perspective, this issue does not contradict the usability of spatial hypertext. Humans are accustomed to retrieving and allocating information in a 3D environment. A widely accepted notion of storing information in our memory involves mental models, which are dynamic representations of 'concepts' that are gradually built and altered as humans navigate their environment. Gibson's theory of perception supports this idea, suggesting that perception arises from identifying invariants within a continuous stream of stimuli [20]. The notion of perception being an active process has been a long debated topic, but may be generally accepted as a concept of perception [3, 4, 41].

For a 3D spatial hypertext, this means that the perception of three-dimensional proximity of information entities can be explored and understood through the user's movement. For example, the relative proximity of an object to its observer may be estimated by perceiving its speed of movement within the field of view when either the observer or the object is moving (*motion perspective* [29]). Thus, having a 'not at a glance' perceivable proximity variant of nodes in a spatial hypertext does not preclude the possibility of achieving a clear understanding of the overall proximity of these nodes. Nevertheless, this still needs to be further determined through the testing of future implementations of immersive spatial hypertext.

6.1.2 Visual Symbols [35] & Visual Cues. The representation of information entities in a spatial hypertext is a topic that needs to be addressed in both 2D and 3D environments, as size, shape, color, and other attributes can explicitly carry meaning and inherently contribute to the perception of an entity's position or prominence in the information space (*visual cues*). However, a virtual environment opens up a greater variety of possibilities with respect to visualization, level of abstraction, positioning (e.g., absolute or relative to the user), and dynamic behavior of spatial hypertext nodes, such as dynamic rotation.

For example, dynamically rotating nodes to face the user based on their position in space could distort their relative proximity if no corrective mechanism (e.g., recalculating relative proximity based on the size of the nodes) is employed. Additionally, in a 3D space, the size of the nodes may be perceived as a depth cue, which is critical, especially when the nodes are not sized uniformly but sized based on their content.

Although this is not exclusive to immersive spatial hypertext, a node in a virtual environment might not necessarily be an abstract information entity (e.g., text in a box), but a 3D model of an object. If rendered realistically, this affects its size and depth, which again has implications on its relative positioning to other nodes or the user.

These examples may illustrate the strong interdependence of features, but their impact on usage, usability and user experience remains to be determined.

6.1.3 Spatial Parsing. Spatial parsing was introduced by several early spatial hypertext systems [30] and was applied in later systems, as well as extended by time as a factor for inferring users' intentions [46]. Parsing in three dimensions might not only entail incorporating a third dimension into calculation, but will have a lot more implications: Apart from already incorporated factors for spatial parsing in two dimensions (e.g., time, relative distance of nodes, recognizing structures, such as lists), three-dimensional parsing might also include the distance from the user to nodes or the current field of view of a user, a these might be an expression of a user's intentions.

6.1.4 Recommender Functionality. Spatial hypertext systems with recommender functionality, such as the one described in [43], use spatial parsing to analyze and query a knowledge base in order to provide the user with additional information entities. This additional information is placed in a workspace based on its affiliation with information entities already located in the space. This is intended to promote creative thinking.

In an environment where information is encoded in three dimensions, parsing might look different from what we see in current systems. Recommending information might also differ: a workspace can be enriched not only by additional information that may be useful to the user, but also by recommending a movement or point of view. By moving through the virtual space, the user might then gain a different perspective on the information presented and derive new insights from this perspective.

6.1.5 Egocentric Distance. The shift from an allocentric to an egocentric view of nodes in an immersive spatial hypertext has implications for other factors, depending on the design of the system. Nodes may grow or shrink in size based on the user's position or current focus. Additionally, it may be considered to impose semantics on the distance from a node to the user. For example, a node that is farther away from the user's current position might be considered, at least temporarily, to be of less importance. As a consequence, spatial parsing and querying recommendations might be adjusted, and/or its visuals might fade out to make it less prominent in the overall space. Since there is a lack of implementations and testing of features like those described, potential implications will need to be defined in the future.

6.1.6 Orientation & Navigation. Orientation and navigation are longstanding topics in hypertext in general. Expanding the workspace to accommodate more information may lead to negative consequences [23] (cf. Section 6.2.1). At the same time, a virtual environment offers opportunities to display objects in direct reference to virtual items for mnemonic purposes, potentially enhancing orientation. However, further research is necessary to gain deeper insights into this topic.

6.1.7 Author versus Reader. Hypertext may be regarded from two different perspectives: author and reader [24, 34]. Although mainly discussed in the narrative field of hypertext research, this view can also be applied to spatial hypertext, as the two roles imply different intentions and target audiences, thus making different demands on the system. When transferring traditional 2D or 2.5D spatial hypertext into a 3D, and especially immersive, environment, this needs to be addressed. Spatial parsing may need to capture an author's intentions differently from a reader's. Suggested movements and viewpoints may play a greater role in supporting a reader than during the construction of a spatial hypertext by an author. This needs to be considered when constructing an immersive spatial hypertext system, but it should also be seen as an opportunity to utilize the characteristics of immersion as essential parts of spatial hypertext.

6.2 3D/Immersion Design Choices And Opportunities

Besides considering factors stemming from the characteristics of spatial hypertext itself, it is vital to account for the characteristics of a 3D environment and immersion. This leads to important insights into potential advancements and opportunities when combining hypertext with these characteristics

6.2.1 Expanding the Workspace. According to Card, "each piece of information has a cost associated with finding and accessing it" [9]. Thus, the workspace of a user has a cost structure for this information at each moment. Information that is frequently or currently needed should be accessible at a low cost to the user. Further, the capacity of a storage system, whether digital or human cognitive, is usually negatively correlated with the cost of accessing information in it. Expanding a workspace enables the system to present more information immediately to the user's perception. However, this increase in accessible information can raise the cost of searching for information [9]. Therefore, expanding the visible workspace to the user may have detrimental effects on orientation, especially when densely populated with information. This challenge of representing a large amount of information while mitigating disorientation has been extensively discussed in previous research [19, 23, 28, 36].

Maintaining orientation in a densely populated workspace requires providing context for the currently focused information. This context can be facilitated by specific workspace layouts, which have been thoroughly examined in the literature. When using binocular presentation of abstract data, the user's perceptional context is influenced by their field of view, dynamically altered by body movements [29]. As discussed earlier in Section 3, perception is an active process facilitated by body movement, requiring active exploration of a space for understanding and memorization. In this context, situated presentation, which will be discussed in Section 6.2.3, may support orientation in large-scale spaces.

6.2.2 Third Dimension as Encoding Channel. A third spatial dimension entails an additional channel for encoding data. While this allows for a greater number of variables to be represented, it also introduces challenges such as foreshortening, occlusion, and depth disparity [29].

One of the first disciplines to explore a third encoding channel was Scientific Visualization (SciVis). Data visualized in this context such as geospatial data and particle flow data have an inherent spatial embedding in three dimensions. The visual representation of abstract data was explored in the Information Visualization (InfoVis) field by early pioneers (cf. [9]) and in research on visual querying (cf. 2D [50], 3D [26]). Today, studies indicate that representing abstract data, such as graphs or graph-like structures, can benefit significantly in terms of comprehensibility, ease of use, manipulation, and understanding from 3D or binocular presentations. These findings should be considered in the context of transferring spatial hypertext into three dimensions. For example, the results of a user study reported by Ware and Franck indicate that the visualization of a comprehensible graph structure on a 3D display can be 1.6 times larger than on a 2D display. With head-coupled motion enabled, the 3D graph can be three times larger [55]. These results were partly replicated by another study by Belcher et al.. Additionally, in this study, ease of use, information display preference, ease of manipulation, and perceived performance were ranked higher than in the compared 2D display version [5]. More recently, Greffard et al. reported that in their study, binocular presentation of graphs outperformed 2D presentation in community detection tasks in both response time and quality of results, especially for complex graph structures [22].

6.2.3 Situated Presentation. In the field of IA, situated visualizations are used to support analytic reasoning by placing information in direct reference to real-life or virtual objects [53]. These objects are semantically bound to the information placed in proximity to them. However, this semantic relationship is typically established by the system to display this connection. In the realm of spatial hypertext, a situated presentation may involve integrating these (virtual) objects as nodes that carry meaning. They may also be used as landmarks for mnemonic purposes. As suggested by [9], especially when handling a single large workspace, parts of this workspace carrying meaning can compensate for the otherwise increasing cost of searching for information. Drawing from insights in the field of IA, a situated presentation of information can be used to mitigate possible intricacies stemming from the expansion of a workspace, such as the cognitive cost of orientation or finding information.

6.2.4 Embodied Cognition. As described in Section 5 embodied cognition may be a central aspect to enable presence and enhance user engagement with a spatial hypertext. From the user's perspective, being embodied in a virtual immersive environment can impact presence and engagement. For the system, the fact that a user can perform complex body movements (in contrast to limited mouse movements) holds the possibility of assigning meaning to those actions. As described earlier, movement in space may, for example, be part of a recommender functionality.

6.2.5 Multisensory Presentation. Research shows that in the absence of *intermodal correlation* (i.e., conflicting sensory input), the human mind (at least in the initial state of stimulation) will give preference to visual stimuli [13]. In general, for non-disabled people, the visual channel is the primary source of information for navigating an environment. Other stimuli, such as auditory or haptic stimuli, are nevertheless important complementary sources of information. Marucci et al. concluded from their study that complementing visual with auditory signals in a virtual environment has a beneficial effect on presence and the processing of stimuli [31]. Their results also suggest that bimodal (i.e., visual-audio and visual-vibrotactile) and trimodal (i.e., visual-audio-vibrotactile) stimulation, especially under conditions of high perceptual load, significantly improve performance compared to visual stimulation alone. The use of stimuli that are perceived as unrealistic, on the other hand, has been found to be detrimental to performance and can cause frustration in users [57]. Research also suggests that multisensory stimulation can affect learning abilities and enhance the recall of memories, such as object location [14, 52]. The multisensory presentation of information in a virtual environment may therefore be an appropriate support for the memorization and identification of objects. Hence, although implementations and literature on full immersion (i.e., additional stimulation of senses such as haptic or oral) are still scarce, the factor of a multisensory presentation should be considered for a potential immersive spatial hypertext system [7, 18, 21, 39].

7 CONCLUSION

In this paper, we present an overview of possible factors of spatial hypertext that need to be considered when conceptualizing and implementing an immersive spatial hypertext system. In addition, we identified opportunities from the fields of virtual reality, immersive analytics, and embodied cognition that may influence new concepts of immersive spatial hypertext systems. However, the factors inherent in the current understanding of spatial hypertext and the emerging opportunities are interdependent, and further insight is needed to understand their impact on user, usage, and the requirements on such a system. Possible next steps may include a thorough conceptualization, implementation and testing of novel concepts of immersive spatial hypertext. Furthermore, we would like to broaden the discussion about spatial hypertext to include concepts from overlapping domains to enrich the development of future systems.

REFERENCES

[1] Giuseppe Abrami, Alexander Mehler, Mevlüt Bagci, Patrick Schrottenbacher, Alexander Henlein, Christian Spiekermann, Juliane Engel, and Jakob Schreiber. 2023. Va.Si.Li-Lab as a Collaborative Multi-User Annotation Tool in Virtual Reality and Its Potential Fields of Application. In *Proceedings of the 34th ACM Conference on Hypertext and Social Media* (Rome, Italy) *(HT '23)*. Association for Computing Machinery, New York, NY, USA, Article 22, 9 pages. https://doi.org/10.1145/3603163.3609076

[2] Shaduan Babbili, Kevin Bönisch, Yannick Heinrich, Philipp Stephan, Giuseppe Abrami, and Alexander Mehler. 2023. Viki LibraRy: A Virtual Reality Library for Collaborative Browsing and Navigation Through Hypertext. In *Proceedings of the 34th ACM Conference on Hypertext and Social Media* (Rome, Italy) *(HT '23)*. Association for Computing Machinery, New York, NY, USA, Article 6, 3 pages. https://doi.org/10.1145/3603163.3609079

[3] R. Bajcsy. 1988. Active Perception. *Proc. IEEE* 76, 8 (1988), 966–1005. https://doi.org/10.1109/5.5968

[4] Ruzena Bajcsy and Mario Campos. 1992. Active and Exploratory Perception. *CVGIP: Image Understanding* 56, 1 (1992), 31–40. https://doi.org/10.1016/1049-9660(92)90083-F

[5] D. Belcher, M. Billinghurst, S.E. Hayes, and R. Stiles. 2003. Using Augmented Reality for Visualizing Complex Graphs in Three Dimensions. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* 84–93. https://doi.org/10.1109/ISMAR.2003.1240691

[6] Anna Berti and Francesca Frassinetti. 2000. When Far Becomes Near: Remapping of Space by Tool Use. *Journal of Cognitive Neuroscience* 12, 3 (05 2000), 415–420. https://doi.org/10.1162/089892900562237

[7] Ayush Bhardwaj, Junghoon Chae, Richard Huynh Noeske, and Jin Ryong Kim. 2021. TangibleData: Interactive Data Visualization with Mid-Air Haptics. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology* (Osaka, Japan) *(VRST '21)*. Association for Computing Machinery, New York, NY, USA, Article 29, 11 pages. https://doi.org/10.1145/3489849.3489890

[8] Frank Biocca. 1997. The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments [1]. *Journal of Computer-Mediated Communication* 3, 2 (09 1997), JCMC324. https://doi.org/10.1111/j.1083-6101.1997.tb00070.x

[9] Stuart K. Card, George G. Robertson, and Jock D. Mackinlay. 1991. The Information Visualizer, An Information Workspace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New Orleans, Louisiana, USA) *(CHI '91)*. Association for Computing Machinery, New York, NY, USA, 181–186. https://doi.org/10.1145/108844.108874

[10] Zhutian Chen, Wai Tong, Qianwen Wang, Benjamin Bach, and Huamin Qu. 2020. Augmenting Static Visualizations with PapARVis Designer. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376436

[11] Jeff Conklin. 1987. Hypertext: An Introduction and Survey. *Computer* 20, 9 (1987), 17–41. http://doi.org/10.1109/MC.1987.1663693

[12] Jeff Conklin and Michael L. Begeman. 1987. gIBIS: A Hypertext Tool for Team Design Deliberation. In *Proceedings of the ACM Conference on Hypertext*. ACM, New York, NY, USA, 247–251. http://doi.acm.org/10.1145/317426.317444

[13] Mark R. Costa, Sung Yeun Kim, and Frank Biocca. 2013. Embodiment and Embodied Cognition. In *Virtual Augmented and Mixed Reality. Designing and Developing Augmented and Virtual Environments*, Randall Shumaker (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 333–342. https://doi.org/10.1007/978-3-642-39405-8_37

[14] Shea Duarte, Andrew Yonelinas, Simona Ghetti, and Joy Geng. 2024. Multisensory Processing Impacts Memory for Objects and their Sources. *Memory & Cognition* (06 2024). https://doi.org/10.3758/s13421-024-01592-x

[15] Tim Dwyer, Kim Marriott, Tobias Isenberg, Karsten Klein, Nathalie Riche, Falk Schreiber, Wolfgang Stuerzlinger, and Bruce H. Thomas. 2018. *Immersive Analytics: An Introduction*. Springer International Publishing, 1–23. https://doi.org/10.1007/978-3-030-01388-2_1

[16] Neven ElSayed, Bruce Thomas, Kim Marriott, Julia Piantadosi, and Ross Smith. 2015. Situated Analytics. In *2015 Big Data Visual Analytics (BDVA)*. 1–8. https://doi.org/10.1109/BDVA.2015.7314302

[17] Barrett Ens, Benjamin Bach, Maxime Cordeil, Ulrich Engelke, Marcos Serrano, Wesley Willett, Arnaud Prouzeau, Christoph Anthes, Wolfgang Büschel, Cody Dunne, Tim Dwyer, Jens Grubert, Jason H. Haga, Nurit Kirshenbaum, Dylan Kobayashi, Tica Lin, Monsurat Olaosebikan, Fabian Pointecker, David Saffo, Nazmus Saquib, Dieter Schmalstieg, Danielle Albers Szafir, Matt Whitlock, and Yalong Yang. 2021. Grand Challenges in Immersive Analytics. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 459, 17 pages. https://doi.org/10.1145/3411764.3446866

[18] Martin Feick, Kora Persephone Regitz, Anthony Tang, and Antonio Krüger. 2022. Designing Visuo-Haptic Illusions with Proxies in Virtual Reality: Exploration of Grasp, Movement Trajectory and Object Mass. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 635, 15 pages. https://doi.org/10.1145/3491102.3517671

[19] G. W. Furnas. 1986. Generalized Fisheye Views. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, Massachusetts, USA) *(CHI '86)*. Association for Computing Machinery, New York, NY, USA, 16–23. https://doi.org/10.1145/22627.22342

[20] E. Bruce Goldstein. 1981. The Ecology of J. J. Gibson's Perception. *Leonardo* 14, 3 (1981), 191–195. https://doi.org/10.2307/1574269

[21] Alexandre Gomes de Siqueira, Rohith Venkatakrishnan, Roshan Venkatakrishnan, Ayush Bhargava, Kathryn Lucaites, Hannah Solini, Moloud Nasiri, Andrew Robb, Christopher Pagano, Brygg Ullmer, and Sabarish Babu. 2021. In *Empirically Evaluating the Effects of Perceptual Information Channels on the Size Perception of Tangibles in Near-Field Virtual Reality*. 1–10. https://doi.org/10.1109/VR50410.2021.00086

[22] Nicolas Greffard, Fabien Picarougne, and Pascale Kuntz. 2011. Visual Community Detection: An Evaluation of 2D, 3D Perspective and 3D Stereoscopic Displays. 215–225. https://doi.org/10.1007/978-3-642-25878-7_21

[23] D. Austin Henderson and Stuart Card. 1986. Rooms: The Use of Multiple Virtual Workspaces to Reduce Space Contention in a Window-Based Graphical User Interface. *ACM Trans. Graph.* 5, 3 (jul 1986), 211–243. https://doi.org/10.1145/24054.24056

[24] Sofia Kitromili, James Jordan, and David E. Millard. 2020. What Authors Think about Hypertext Authoring. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media* (Virtual Event, USA) *(HT '20)*. Association for Computing Machinery, New York, NY, USA, 9–16. https://doi.org/10.1145/3372923.3404798

[25] Karsten Klein, Michael Sedlmair, and Falk Schreiber. 2022. Immersive Analytics: An Overview. *it - Information Technology* 64, 4-5 (2022), 155–168. https://doi.org/10.1515/itit-2022-0037

[26] Uwe Krohn. 1996. VINETA: Navigation Through Virtual Information Spaces. In *Proceedings of the Workshop on Advanced Visual Interfaces* (Gubbio, Italy) *(AVI '96)*. Association for Computing Machinery, New York, NY, USA, 49–58. https://doi.org/10.1145/948449.948458

[27] Christoph Leuze, Grant Yang, Brian Hargreaves, Bruce Daniel, and Jennifer A McNab. 2018. Mixed-Reality Guidance for Brain Stimulation Treatment of Depression. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 377–380. https://doi.org/10.1109/ISMAR-Adjunct.2018.00109

[28] Jock Mackinlay, George Robertson, and Stuart Card. 1991. The Perspective Wall: Detail and Context Smoothly Integrated. *Conference on Human Factors in Computing Systems - Proceedings*, 173–176. https://doi.org/10.1145/108844.108870

[29] Kim Marriott, Jian Chen, Marcel Hlawatsch, Takayuki Itoh, Miguel Nacenta, Guido Reina, and Wolfgang Stuerzlinger. 2018. *Immersive Analytics: Time to Reconsider the Value of 3D for Information Visualisation*. 25–55. https://doi.org/10.1007/978-3-030-01388-2_2

[30] Catherine C. Marshall and Frank M. Shipman. 1993. Searching for the Missing Link: Discovering Implicit Structure in Spatial Hypertext. In *Proceedings of the Fifth ACM Conference on Hypertext* (Seattle, Washington, USA) *(HYPERTEXT '93)*. Association for Computing Machinery, New York, NY, USA, 217–230. https://doi.org/10.1145/168750.168826

[31] Matteo Marucci, Gianluca Di Flumeri, Gianluca Borghini, Nicolina Sciaraffa, Michele Scandola, Enea Pavone, Fabio Babiloni, Viviana Betti, and Pietro Aricò. 2021. The Impact of Multisensory Integration and Perceptual Load in Virtual Reality Settings on Performance, Workload and Presence. *Scientific Reports* 11 (03 2021). https://doi.org/10.1038/s41598-021-84196-8

[32] Jon McCormack, Jonathan C. Roberts, Benjamin Bach, Carla Dal Sasso Freitas, Takayuki Itoh, Christophe Hurter, and Kim Marriott. 2018. *Multisensory Immersive Analytics*. Springer International Publishing, Cham, 57–94. https://doi.org/10.1007/978-3-030-01388-2_3

[33] Alexander Mehler, Giuseppe Abrami, Christian Spiekermann, and Matthias Jostock. 2018. VAnnotatoR: A Framework for Generating Multimodal Hypertexts. In *Proceedings of the 29th on Hypertext and Social Media* (Baltimore, MD, USA) *(HT '18)*. Association for Computing Machinery, New York, NY, USA, 150–154. https://doi.org/10.1145/3209542.3209572

[34] Susan Michalak and Mary Coney. 1993. Hypertext and the Author/Reader Dialogue. In *Proceedings of the Fifth ACM Conference on Hypertext* (Seattle, Washington, USA) *(HYPERTEXT '93)*. Association for Computing Machinery, New York, NY, USA, 174–182. https://doi.org/10.1145/168750.168820

[35] David Millard, Luc Moreau, Hugh Davis, and Siegfried Reich. 2000. FOHM: A Fundamental Open Hypertext Model for Investigating Interoperability between Hypertext Domains. 93–102. https://doi.org/10.1145/336296.336334

[36] Kenneth Mitchell and Jessie Kennedy. 1997. The Perspective Tunnel: An Inside View on Smoothly Integrating Detail and Context, Wilfrid Lefer and Michel Grave (Eds.). Boulogne Sur Med, France.

[37] Preben Mogensen and Kaj Grønbæk. 2000. Hypermedia in the Virtual Project Room - Toward Open 3D Spatial Hypermedia. In *Proceedings of the Eleventh ACM on Hypertext and Hypermedia* (San Antonio, Texas, USA) *(HYPERTEXT '00)*. Association for Computing Machinery, New York, NY, USA, 113–122. https://doi.org/10.1145/336296.336340

[38] Jules Moloney, Branka Spehar, Anastasia Globa, and Rui Wang. 2018. The Affordance of Virtual Reality to Enable the Sensory Representation of Multi-Dimensional Data for Immersive Analytics: From Experience to Insight. *Journal of Big Data* 5 (12 2018). Issue 1. https://doi.org/10.1186/s40537-018-0158-z

[39] Thomas Muender, Michael Bonfert, Anke Verena Reinschluessel, Rainer Malaka, and Tanja Döring. 2022. Haptic Fidelity Framework: Defining the Factors of Realistic Haptic Feedback for Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 431, 17 pages. https://doi.org/10.1145/3491102.3501953

[40] H. Van Dyke Parunak. 1991. Don't Link Me In: Set Based Hypermedia for Taxonomic Reasoning. In *Proceedings of the 3rd Annual ACM Conference on Hypertext*. ACM, New York, NY, USA, 233–242. http://doi.acm.org/10.1145/122974.122998

[41] Giovanni Pezzulo, Francesco Donnarumma, Pierpaolo Iodice, Domenico Maisto, and Ivilin Stoianov. 2017. Model-Based Approaches to Active Perception and Control. *Entropy* 19, 6 (2017). https://doi.org/10.3390/e19060266

[42] Catlin Pidel and Philipp Ackermann. 2020. Collaboration in Virtual and Augmented Reality: A Systematic Overview. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso De Paolis and Patrick Bourdot (Eds.). Springer International Publishing, Cham, 141–156.

[43] Daniel Roßner, Claus Atzenbeck, and Tom Gross. 2019. Visualization of the Relevance: Using Physics Simulations for Encoding Context. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media* (Hof, Germany) *(HT '19)*. Association for Computing Machinery, New York, NY, USA, 67–76. https://doi.org/10.1145/3342220.3343659

[44] David Saffo, Sara Di Bartolomeo, Tarik Crnovrsanin, Laura South, Justin Raynor, Caglar Yildirim, and Cody Dunne. 2024. Unraveling the Design Space of Immersive Analytics: A Systematic Review. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 495–506. https://doi.org/10.1109/TVCG.2023.3327368

[45] Marc Satkowski and Raimund Dachselt. 2021. Investigating the Impact of Real-World Environments on the Perception of 2D Visualizations in Augmented Reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 522, 15 pages. https://doi.org/10.1145/3411764.3445330

[46] Thomas Schedel and Claus Atzenbeck. 2016. Spatio-Temporal Parsing in Spatial Hypermedia. In *Proceedings of the 27th ACM Conference on Hypertext and Social Media* (Halifax, Nova Scotia, Canada) *(HT '16)*. Association for Computing Machinery, New York, NY, USA, 149–157. https://doi.org/10.1145/2914586.2914596

[47] Lawrence Shapiro and Shannon Spaulding. 2021. Embodied Cognition. In *The Stanford Encyclopedia of Philosophy* (Winter 2021 ed.), Edward N. Zalta and Uri Nodelman (Eds.). Metaphysics Research Lab, Stanford University.

[48] Frank M. Shipman and Catherine C. Marshall. 1999. Spatial Hypertext: An Alternative to Navigational and Semantic Links. *Comput. Surveys* 31, 4es (dec 1999), 14–es. https://doi.org/10.1145/345966.346001

[49] Rosemary Michelle Simpson. 1996. Beyond the Plane: Spatial Hypertext in a Virtual Reality World. *SIGWEB Newsl.* 5, 3 (oct 1996), 13–16. https://doi.org/10.1145/264838.264841

[50] A. Spoerri. 1993. InfoCrystal: A Visual Tool for Information Retrieval. In *Proceedings Visualization '93*. 150–157. https://doi.org/10.1109/VISUAL.1993.398863

[51] Leonardo Tanzi, Pietro Piazzolla, Francesco Porpiglia, and Enrico Vezzetti. 2021. Real-time Deep Learning Semantic Segmentation During Intra-operative Surgery for 3D Augmented Reality Assistance. *International Journal of Computer Assisted Radiology and Surgery* 16 (06 2021). https://doi.org/10.1007/s11548-021-02432-y

[52] Antonia Thelen and Micah M. Murray. 2013. The Efficacy of Single-Trial Multisensory Memories. *Multisensory Research* 26, 5 (2013), 483 – 502. https://doi.org/10.1163/22134808-00002426

[53] Bruce H. Thomas, Gregory F. Welch, Pierre Dragicevic, Niklas Elmqvist, Pourang Irani, Yvonne Jansen, Dieter Schmalstieg, Aurélien Tabard, Neven A. M. ElSayed, and Wesley Smith, Ross T.and Willett. 2018. *Situated Analytics*. Springer International Publishing, Cham, 185–220. https://doi.org/10.1007/978-3-030-01388-2_7

[54] Andries van Dam, Andrew Forsberg, David Laidlaw, Joseph Jr, and Rosemary Michelle Simpson. 2000. Immersive VR for Scientific Visualization: A Progress Report. *IEEE Computer Graphics and Applications* 20 (12 2000), 26–52. https://doi.org/10.1109/38.888006

[55] C. Ware and G. Franck. 1994. Viewing a Graph in a Virtual Reality Display is Three Times as Good as a 2D Diagram. In *Proceedings of 1994 IEEE Symposium on Visual Languages*. 182–183. https://doi.org/10.1109/VL.1994.363621

[56] Andrea Stevenson Won, Jeremy Bailenson, Jimmy Lee, and Jaron Lanier. 2015. Homuncular Flexibility in Virtual Reality. *Journal of Computer-Mediated Communication* 20, 3 (01 2015), 241–259. https://doi.org/10.1111/jcc4.12107

[57] Dr Ying Zhang and Adrian Travis. 2008. Creation and Evaluation of a Multi-Sensory Virtual Assembly Environment. *International Journal of Automation and Computing* 5 (04 2008), 163–173. https://doi.org/10.1007/s11633-008-0163-3