PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

The mixed reality of things: emerging challenges for human-information interaction

Ryan P. Spicer
Stephen M. Russell
Evan Suma Rosenberg

SPIE.

Invited Paper

The Mixed Reality of Things: Emerging Challenges for Human-Information Interaction

Ryan P. Spicer[1], Stephen M. Russell[2], and Evan Suma Rosenberg[1]

[1]University of Southern California (United States)
[2]U.S. Army Research Lab (United States)

ABSTRACT

Virtual and mixed reality technology has advanced tremendously over the past several years. This nascent medium has the potential to transform how people communicate over distance, train for unfamiliar tasks, operate in challenging environments, and how they visualize, interact, and make decisions based on complex data. At the same time, the marketplace has experienced a proliferation of network-connected devices and generalized sensors that are becoming increasingly accessible and ubiquitous. As the "Internet of Things" expands to encompass a predicted 50 billion connected devices by 2020, the volume and complexity of information generated in pervasive and virtualized environments will continue to grow exponentially. The convergence of these trends demands a theoretically grounded research agenda that can address emerging challenges for human-information interaction (HII). Virtual and mixed reality environments can provide controlled settings where HII phenomena can be observed and measured, new theories developed, and novel algorithms and interaction techniques evaluated. In this paper, we describe the intersection of pervasive computing with virtual and mixed reality, identify current research gaps and opportunities to advance the fundamental understanding of HII, and discuss implications for the design and development of cyber-human systems for both military and civilian use.

Keywords: Mixed Reality, Virtual Reality, Human-Information Interaction

1. INTRODUCTION

Virtual Reality (VR) and Augmented Reality (AR) offer compelling advantages for completing tasks that require interacting with particular types of information. Due to the perceptual richness of spatial interfaces available through AR and VR, it is inevitable that some knowledge workers will interact with information-spaces through these devices in the future. Once these tasks are identified, the people performing them will require effective ways of working and collaborating in Mixed Reality (MR) spaces including some elements of the physical world and some elements of a virtual environment (VE). Technical hurdles remain to be cleared before MR becomes technically feasible at similar price-points to contemporary commercial, off-the-shelf fully-occluded VR Head Mounted Displays (HMDs). Even so, requirements are pushing technology to converge on MR since knowledge workers cannot afford to operate for longer than brief sessions in fully-occluded VR, cut off from their bodies, their collaborators, and their devices. MR working environments will likely involve some combination of users' physical environment and virtual elements, with the precise combination driven by specific use cases. The virtual and physical may include collaborators, either virtually or physically present in the shared information-space. These

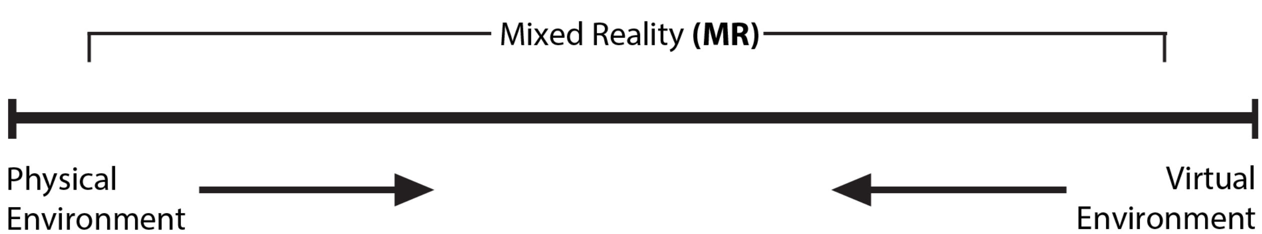
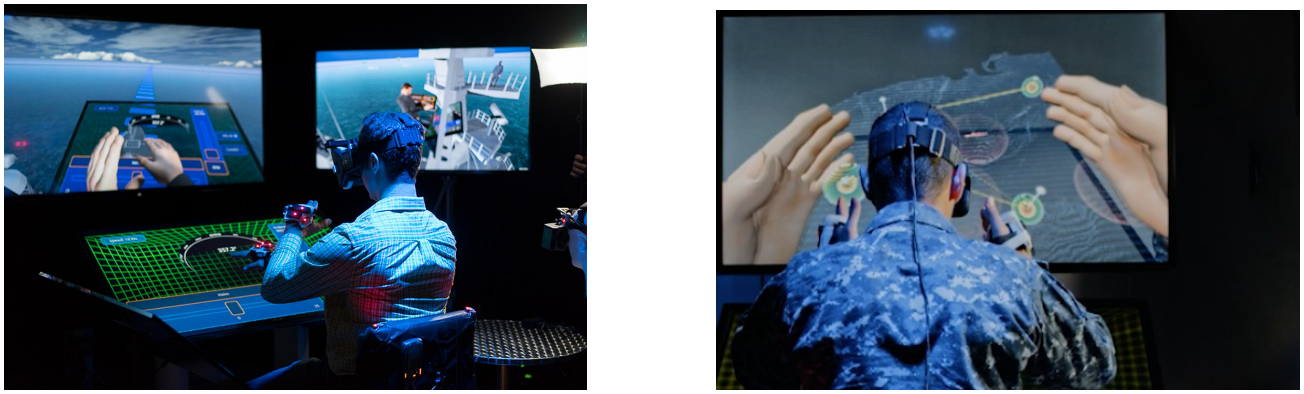


Figure 1: The Reality-Virtuality continuum, as defined by Milgram.[1] The term 'physical environment' is used here in place of 'real environment' to imply that the information is also real.

Next-Generation Analyst V, edited by Timothy P. Hanratty, James Llinas, Proc. of SPIE Vol. 10207, 102070A · © 2017 SPIE · CCC code: 0277-786X/17/$18 · doi: 10.1117/12.2268004



(a) A user interacts with a VE through a combination of gestural and 2D screen interactions.

(b) Interacting with an abstract representation of data through free-space gestures.

Figure 2: Mixed Reality (MR) interactions.

systems fall somewhere on the midpoint of the Real-Virtuality (RV) Continuum,[1] which describes a continuum of display solutions from see-through Augmented Reality (AR) systems that superimpose virtual elements onto the real world[2] to fully-occluded virtual reality displays (Figure 1).

The emerging field of Human-Information Interaction (HII) presents both research challenges and opportunities in the realm of MR systems. HII is the study of relationships between people and information, with an emphasis on decision-making outcomes.[3] HII differs from Human-Computer Interaction (HCI) because it foregrounds specifically the human system interacting with information, rather than the computer User Interface (UI) system, and process of querying data, modifying or annotating it, and using it to drive actionable outcomes. HII systems include aspects of information visualization, but extend beyond it, since users must interact with the presentation of data and use insight derived from the data to drive actionable decisions.

Computational power has become ubiquitous. "Cloud computing," the philosophy of off-loading computing power to machines located off-site and connecting diverse devices, including desktop and laptop computers, mobile devices such as mobile phones and tablets, wearables, and Internet of Things (IoT) devices, pushes systems designers to to consider a variety of lenses into the same information-space suitable to the affordances of each device. Furthermore, inexpensive, low-power network connectivity enables Internet-connected communication with everything from lightbulbs to jetliners. These connected objects provide rich sources of information and potential fulcrums to impact the world. From these IoT devices, to the social graph, to network traffic, data and information are accumulating in storage far faster than individual humans can make sense of them. Additional tools are required to filter this fire hose of data into actionable intelligence, and to help humans leverage their particular skills and advantages over fully-automated systems to achieve profound insights.[4]

The depth and breadth of potential MR solutions presents a rich set of possibilities for novel human information interaction systems. MR environments are inherently spatial, and offers the potential to leverage users' human perception and existing intuition about the physical world. Although projection and large-screen mixed reality systems exist, in this paper we will primarily focus on head mounted displays (HMDs). These displays are stereoscopic, and typically provide six-degrees-of-freedom tracking so that the rendering of the Virtual Environment (VE) can respond in appropriate ways to the user's motion in the physical world (Figure 2). MR systems will, at least initially, be novel for most users. They may also be used intermittently. For this reason, system interactions should have a low floor and high ceiling.[5] That is, they should be easy for a new user to pick up and use, but provide access to enough complexity to provide value to expert users.

In this paper, we consider the disciplines of Mixed Reality and Human Information Interaction to identify open research questions and promising future directions. Specifically, we discuss information analysis, mapping data in MR environments, user input for MR HII, searching and browsing for information in MR environments, and sharing MR environments with multiple collaborators for HII.

2. DATA SOURCES

Significant scholarship has been devoted to the raw quantity of information being produced. Lyman and Varian estimated that in just the year 2002, the world produced 5 exabytes (5×10^{18} bytes) of new information, and that the amount of information produced grew by about 30% per anum.[6] This estimate was produced before the first mass-available smart phones, before the advent of popular media sharing services like YouTube, and before the wide popularity of blogs, never mind microblogs. One source puts the volume of data worldwide at 4.4 zettabytes (4.4×10^{21} bytes) as of 2014, and projects that the total data will reach 44 zettabytes by 2020.[7]

2.1 Internet of Things

The so-called Internet of Things consists of many individual devices that have been connected to network systems. The growth in this sector has been driven by the proliferation of low-cost, low-power wireless networking chipsets, and decentralized "cloud computing." Devices ranging from light bulbs to thermostats have been connected to networks to facilitate convenient remote control and monitoring. Information may be available from witting, unwitting, and hostile devices. Devices you own may share information with you intentionally. Devices you do not own may nonetheless share information with you, as a result of a successful cyber-attack. Both classes of device may be suborned by opposing forces to provide inaccurate data. HII must assist analysts in making sense of all three types of data, to make actionable decisions.

2.2 User-Generated Content

Since the "Web 2.0" era, user-generated content has been perceived as a major theme of Internet business. User-generated content can include textual output blog posts or image, audio, or video sharing. Each of these sources of data provides both explicit and implicit data and metadata about content creators, and their surrounding context. Much of this content is irrelevant to any given context or query, but methods to effectively filter and query this corpus of information must be developed to allow analysts to identify and extract relevant information.

3. HUMAN INTERACTION WITH INFORMATION

The work of the analyst is to assess the world around them, forecast future capabilities, threats, vulnerabilities and opportunities, and develop options.[4] Analysts build a set of information and then connect it to defend their reasoning.[4,8,9] In the context of the analyst, the role of HII is to study how analysts work with information and make decisions, and to determine how to best assist them in obtaining the best information to make defensible decisions.

The deluge of new information is adding new data to the world faster than any human alone can address it. Nevertheless, the requirement exists for people to make actionable decisions quickly in light of available information. This presents the need for tools to summarize data, identify "interesting" trends, and allow users to search data for meaningful information. Analysts must make sense of this information, whether by foraging for items of interest[10] or using other search strategies.

Human Information Interaction (HII) is the study of the relationship between people and information, with an emphasis on decision-making outcomes.[3] HII is complimentary to Human-Computer Interaction (HCI), but focuses on data and actionable outcomes rather than user interface (UI). HII tends to be highly context-specific, since interactions with information are, by definition, centered around specific information problems and domains. HII research has involved qualitative and quantitative studies. Research has explored both context-independent HII principles, and in-context HII for specific user communities or work contexts. General HII offers the prospect of broader applicability, since results are not scoped to the community and context under study, but also resists drawing deeper conclusions about specific communities. There is some question about the locus of context for HII. For the purpose of this discussion, we focus on task-centric context.

Significant prior work in Information Retrieval (IR) and Human Information Behaviour (HIB) has considered how to provide the best document in response to a query. The context under study here considers information aggregated from many sources instead of discrete documents. It may be possible to apply IR and HIB concepts to this context by considering the document as a view onto a data set that contains information from many sources in a specific presentation rather than a discrete element.

One aspect of information presentation of particular concern within the context of analytical work is Value of Information (VOI) and relevance to analysis. HII systems must present VOI as specified by analysts or systems inputting information, in such a way that analysts later retrieving or viewing the information can make appropriate decisions about its relevance and certainty to their analysis. This information can help analysts understand the trade-space in deciding whether to widen or narrow their search, or drill down into specific documents.[11]

4. MAPPING DATA TO MIXED ENVIRONMENTS

MR can provide potential advantages for analysts performing HII in two broad areas: First, a MR display allows an analyst to create and use an arbitrary number of virtual screens in their physical environment. Second, MR allows a window into an unbounded three-dimensional space in which visual analytics tasks can be performed. These two modes of operation have slightly different affordances.

In the case of virtual screens, the screens can be positioned and angled arbitrarily, and can display any information that could be displayed on a 2D display screen (eg: LCD or similar). Some research considers virtual screens anchored to existing walls or other structures the physical environment, as with Microsoft's Hololens. It may be useful to consider a shell of rectangular screens describing a sphere around the user. The utility of these virtual screens is constrained by the angular resolution of the MR display device. With industry taking an active interest in MR display technology (eg: Microsoft Hololens, Meta AR, Avegant, etc, as of this writing), research into future MR HII capabilities can assume a steady increase in the angular resolution and other qualities of MR displays.

MR can also provide a head-coupled stereo viewpoint onto the three dimensional (3D) world that is not bounded by the frame of a screen. In this domain, MR HII can leverage the strengths of the visual analytics community and the virtual and mixed reality research community. The visual analytics community defines the goal of support tools as to assist the analyst in executing sound analytic technique routinely, which includes searching the possibility-space for viable explanations, which must then be evaluated and supported or refuted with evidence.[4] Information visualization as a discipline seeks to amplify human cognition by leveraging visual processing. Card et al. define six major contributions of visual analytics: (1) to increase the memory and processing resources available to users; (2) to reduce the search for information. (3) using visual representations of information to enhance detection of patterns. (4) by supporting perceptual inference; (5) by using perceptual attention mechanisms for monitoring; (6) By encoding information in a manipulable medium.[12] MR environments have the potential to combine the perceptual advantages of visual analytics with the perceptual advantages of head-coupled perspective, broadening the information channel available to the user.

VR and MR are inherently spatial and embodied, and may offer benefits for viewing and interacting with information while conducting HII tasks. The key difference between VR and MR displays and other screens is that the user's body is key to the experience. The viewpoint into virtual elements is tied to the user's head pose. The virtual elements of a MR display are effectively bounded only by the limits of the user's perception. Mixed realities are multi-sensory, and involve the user's body and proprioception.

To be acted upon in the context of HII, data must be represented computationally, and transformed into a suitable visual presentation.[4] Data is available in an ever-increasing number of forms. Time-series numeric data from sensors is perhaps the easiest to envision, but data can also include texts such as articles, blog posts, e-mails, and the like. Much of the data currently available is in forms such as still images and audio or video recordings. These media assets can be be analyzed to produce, for instance, text transcripts or other features. Contemporary work in machine-learning for classification tasks makes a variety of features more easily available than ever before. Much of this data is also surrounded by metadata such as authorship, intended recipient lists, geospatial characteristics, temporal characterisitics, comments, and the like. This data can be correlated with other data such as the social graph. The challenge for analysts is not just getting enough data, but filtering meaningful data from the available information. Once the data has been selected, it must be presented in such a way as to support decision making.

Data visualization is almost always a combination of dimension reduction and selecting salient representations of the remaining dimensions. Many virtual and mixed environments to date deal with explicitly spatial data, such as mapping/GIS, or 3D representations of physical objects. Many research trajectories at present investigate the use of virtual replications of physical environments for simulation and training. In these cases, mapping from the physical to the virtual in the simulation context is unambiguous. An environment is successful if the user is able to suspend disbelief and react to the virtual world as if it was the real world. MR systems can be used in this context, if the user is out in the physical office: MR displays can superimpose geo-located information on the physical world, and those that allow multiple users to share a visualization of some information. In the first case, the primary value of AR is the combination of virtual elements aligned in space with physical elements.

Knowledge-workers, however, will mostly interact with data about other places in their own office spaces. For this reason, research into optimal mappings of data into MR environments will be essential. Not all HII tasks will be best accomplished in MR environments, so determining the suitability of task to MR environments will be critical. Existing research has engaged some aspects of HII in VR,[13] but additional investigation is required.

4.1 Position in 3D

MR for HII in this context poses both challenges and opportunities relative to the more prosaic one-to-one mappings of three dimensionality. Since MR is inherently spatial, the three spatial dimensions are always available. These dimensions must be allocated carefully, however, since less information can be reliably extracted from depth and distance from the viewer. Beyond a few meters, parallax motion cues are more salient than stereoscopy,[14] and overlapping entities can potentially obscure more information than they provide.

Research suggests that head-coupled perspective and stereo cues increase the size of an abstract graph that can be understood by users by a factor of three in "fish-tank" VR.[15] Greater involvement of the user's full range of head movement in HMD-based MR may increase this figure further.

The temporal dimension is also available. Representations can be designed to vary with time and the user can be given control over the flow of time in the virtual portions of the mixed environment. Changes in position or representation over time can allow users to maintain context between alternate projections, which may assist in identifying connections between elements in ways that might not be clear when jumping between alternate views without a transition animation. Similarities and differences among the motion of elements also provide a grouping cue for human perception, or a way of calling user attention to specific elements.[16]

Beyond position and time, orientation can again provide a number of cues. However, care must be taken that the mapping remains legible to the user. There is a point beyond which subtlety in mapping, while it allows fine gradations in meaning, loses usefulness because those gradations are not understandable to the user.

4.2 Visual Features

Additional features such as size, shape and form, color, and motion can be manipulated, but just as in two-dimensional information visualization, care must be taken that important differences remain salient to the user.

Many types of information of interest to knowledge workers are not as inherently spatial. Even in the case of information that is tied to a concrete physical location, properties beyond geo-location are often more salient. Information that can be mapped onto connectivity graphs can leverage existing research into graph layout. In the case of data with multiple salient spatial projections, which may or may not include geo-location, designing systems that transform between projections while retaining user awareness of relationships is an important question. Some research has addressed this challenge in the context of multiple 2D screens or windows,[17] but questions about how to exploit the affordances of MR to maintain this continuity of awareness remain open.

Some information can not be reformulated as graph connectivity. Spatializing this data presents an important set of questions. In many cases, information visualization best practices for 2d data should inform the design of MR data presentation. The mapping of values into 3D space involves the position and orientation of data points, as well as the rendering of data points. Many techniques are technically feasible, so the question becomes optimizing presentation for user effectiveness.

Research has investigated the saliency of lighting effects on three dimensional objects for understanding node-link graphs on screen displays.[18] Work has explored the synthesis of visual texture to communicate data in 3D visualizations.[19] This perceptual channel adds additional richness to shape and form of data points in three dimensional space.

The majority of this work has targeted two-dimensional or screen-based 3D applications, which may include perspective projection and navigation, but not stereoscopy or head-coupled perspective. While many of these results may transfer directly to the MR context, care must be taken to ensure that the specific benefits and pitfalls of MR rendering and interaction are accounted for in future HII systems.

5. BOUNDING UNBOUNDED INPUT

MR allows for unbounded input, in contrast to windows, icons, menus, pointer (WIMP) and keyboard computer interfaces. In a screen based interface, interaction typically occurs through a combination of keyboard and mouse commands. These interfaces typically follow conventions – buttons afford actions; checkboxes, dropdowns, or radio buttons afford selection; text-fields afford data entry; scroll-bars afford navigating through large lists. These conventions are the domain of HCI, but can be used to create systems that support HII.

Commercial VR systems increasingly provide some element of embodied user interaction, in addition to head-coupled stereo imagery. Hand-held wands are available for the major commercial systems, and free-space gesture systems are increasingly available and reliable. Free space gesture systems offer extreme flexibility, but also present challenges in making the space of available gestures clear to the user. With any free-space gesture system, user fatigue becomes a significant issue. User interface affordances must be selected to optimize for user performance over long sessions and minimize fatigue.

In either case, the potential interaction space is significantly less-bounded and conventions are less established. Research has investigated best practices to create discoverable free-space gestures.[20] Further development is required to create interaction paradigms that are rich enough to facilitate useful HII outcomes and discoverable enough to be useful to end users. Our philosophy of MR interaction suggests that a successful MR system may involve both unbounded, free-space gestures and more traditional keyboard/mouse pointer interaction. In considering the challenges and opportunities for input in HII MEs, we consider a taxonomy of three dimensional user interface components.[21]

Speech-based natural user interfaces provide another interesting set of relatively unbounded affordances. Similar to free-space gestural interfaces, relatively few established interface paradigms exist. Ideas can be expressed in many ways in free human language, so some bounding must occur.

5.1 Locomotion and Travel

Locomotion refers to the process of physical movement within a 3D environment. Requirements for locomotion in MR HII systems are coupled with the spatial mapping employed when rendering data in three dimensions. Since the data presented is not always inherently spatial, or non-spatial aspects may be more salient to some analysis tasks, layout choices can be made to optimize analyst performance. Tradeoffs between overview and detail presentation may require different locomotion affordances, and locomotion may involve scaling information in addition to translation.

HII systems in MR will require micro and macro movements. Depending on the specific requirements, HII systems will involve moving the head-coupled point of view locally, in response to user motion in the physical world. Some systems may also require large-scale locomotion through the information-space. A variety of techniques have been developed to facilitate large scale locomotion in virtual and mixed environments, including teleportation,[22] "Seven-League Boots,[23]" and redirected walking.[24] These approaches rely on intentionally distorting relationships between the physical and virtual space. Redirected walking, for example, relies on the physical world being fully occluded by the VE. The technique attempts to modify the relationship between physical and virtual motion such that the user does not percieve any disconnect. In mixed environments that include elements of the real world, this becomes more challenging. Likewise, the "Seven League Boots" technique amplifies the translation of the user's head-coupled perspective relative to the physical world. This amplification is transparent to the user, so there is less concern about MEs including reference to physical objects breaking the illusion. At the same time, incongruous movements between virtual and physical objects in MEs could prove disorienting to users. These phenomena require further inquiry.

In mixed environments where real world elements are fixed, it may be profitable to consider a question of perspective: is the user moving through a fixed world, or is the user moving the world around their fixed location?

For more abstract data-sets, it may be sufficient to move the data around the user. In this type of abstract visualization, there are no existing preconceptions about "how things work in the physical world" that would be violated by moving the data around the user. Care must be taken, however, to include enough physical objects, or virtual objects tied to the user's position, to guard against simulator sickness brought on by high optical flow.

5.2 Wayfinding

In contrast to locomotion, wayfinding is the cognitive process of determining a path in a 3D environment. Discussion of wayfinding in VEs has primarily focused on spaces that emulate physical environments. During wayfinding in the physical world, people build models based on landmark knowledge, procedural knowledge, and survey knowledge.[25] These types pertain to salient physical features in the space, knowledge of established paths or sequences of moves in the space, and global map knowledge, respectively. While these types of knowledge are also relevant to spatial presentations of information, the variety of possible mappings into 3D space in MEs raises new challenges and opportunities.

Arbitrary and flexible mappings may allow analysts to discover relationships between data points by altering how data is projected into the ME. However, making mappings too flexible or evanescent runs the risk of frustrating human spatial cognition. To aid in positioning oneself within a larger environment, visual maps are a commonly employed user interface strategy. In the context of spatial mappings of arbitrary data, it may be worth considering providing maps of the data-to-spatial mapping space in addition to the local-to-global position space. Effective techniques and best practices for visually representing the relationship between abstract data and 3D space are an open area for future research.

Even in abstract and symbolic MEs, it may be useful to consider best practices from physical world wayfinding and navigation through virtual spaces that emulate the physical world. Properties such as districts, static landmarks, paths, nodes, and borders or edges can be introduced in the visual language of a ME to aid in wayfinding.[26] Design patterns for these arbitrary wayfinding aids have been subject to research, but the work primarily addresses VEs that simulate physical worlds.[27]

Analysts may gain value from examining multiple mappings of arbitrary data to spatial presentations. The importance of maintaining context while altering these mappings bears consideration. Animated changes to mappings may prove useful in maintaining user awareness of context while shifting between alternate spatial mappings. This matches the desire to avoid teleportation in VEs, as users must rebuild their local context after each teleport.[22] The degree of fluidity that can be allowed while retaining context presents an interesting question for further study.

The relative scale of the spatial representation in an HII ME may impact the best practices for wayfinding. Scale and wayfinding aids may be considered in terms of egocentric and exocentric perspectives.[25] In an egocentric context, the virtual environment under consideration is scaled and manipulated in the context of the user's body and proprioception. A VE scaled to represent a sand table might invite exocentric perception from a "god-level" point of view. However, viewing the same environment from the perspective of a person walking the terrain at 1:1 scale would support an egocentric relationship between the user and the world. VEs and MEs give the user the ability to change the scale at which they view the information space, so adapting between egocentric and exocentric wayfinding techniques for the same information space may warrant investigation. Furthermore, MEs may exist within the spatial constraints of a user's physical environment, making it potentially advantageous to manipulate the scale of the virtual components to best fit the workspace.

5.3 Selection and Manipulation

WIMP ("windows, icons, menus, pointer") interfaces allow 2D software to share a relatively common interface paradigm for selection and manipulation. While these common user interface tasks have been studied in 3D environments, they are far from standardized. In general, 3D selection and manipulation is orders of magnitude more complex than their 2D counterparts, and is also more dependent on the specific technology, task, and context.

3D interaction techniques can be coarsely categorized as either "naturalistic" or "magical" approaches. Naturalistic techniques are modeled after real world behaviors, whereas magical ones are not so restricted and may violate or even completely reimagine the laws of physics. Both approaches have advantages and disadvantages for different types of MR environments and tasks.

Because they are based on the real world, naturalistic techniques are often intuitive and easy to learn. For example, selecting a 3D object may be as simple as reaching out and touching it. To then reposition the object, one might move their hand to place the object in a new position, mimicking the motions one would take in the real world. However, using naturalistic techniques for an extended period time can result in fatigue, and may not be the most efficient way of completing the task. Furthermore, complex naturalistic interactions, such as free-space gestures, may be difficult for a computational system to sense and recognize with sufficient levels of accuracy and robustness. The bounds of potential gesture spaces are very wide, and care must be taken such that gestures are discoverable by the user, and that the ME provides sufficient feedback for the user to determine if attempted gestures are successful.

Magical techniques have many potential advantages. Because they are not limited by the laws of physics, MR systems can support interactions that transcend what is possible in the real world. For example, common magical selection techniques include pointer interaction,[28,29] where a ray cast from a wand, the user's hand, or the user's gaze direction, and image-plane selection,[30] where a ray is cast from the user's head-coupled position, through a wand or the user's hand, and out into the world. Both approaches allow the selection and manipulation of distant objects. The downside of magical techniques is that they often unfamiliar and require a learning curve. However, with sufficient training, they can enable skilled users to accomplish tasks with greater efficiency and less fatigue than naturalistic techniques.

In general, selection and manipulation techniques must strike a balance between discoverability for new or infrequent users, and ease and speed of employment for skilled users. As MR environments continue to grow in popularity, both naturalistic and magical interaction techniques will continue to be major research areas of interest.

5.4 Symbolic Input

Most HII tasks involve entering information into the system, whether to formulate queries, annotate documents or views, or create finished knowledge products for dissemination to others. Screen-based user interfaces for HII benefit from the keyboard, which provides clear affordances for symbolic input, and has the benefit of familiarity to most users. MEs supporting HII tasks will need to provide affordances that support these tasks.

The specific nature of the affordance will vary based on the task and context. Keyboard-based input will remain salient for seated or standing-in-place interactions, but may prove ungainly when the user is locomoting in physical space. Virtual keyboards controlled with gestures or wands address some need for input that moves with the user in physical space, but suffer from limited affordances for haptic feedback, precision, and speed relative to physical keyboards. Symbolic input must be designed with this context of use in mind; context of use may also be determined in part by the required symbolic input affordances.

It is likely that MEs for HII will require the use of a keyboard or keyboard-like device for symbolic input. A variety of keyboards for use in VEs have been investigated. Some work has evaluated miniature keyboards[31] and chorded keyboards,[32] but issues of fatigue and learning curve led to reduced performance in evaluation. Any number of more free-space gestural interfaces for text and symbolic input have been have been considered, but also likely share some of the same problems with user adoption and input speed as chorded keyboards or miniature keyboards.

Most contemporary commercial VR hardware provides input via wands, game controllers, or free-space gesture via cameras. These inputs can be used for symbolic input, with on-screen keyboards or similar metaphors. Laser pointer interaction with a wand can be used to enter text, albeit more slowly than on a physical keyboard, for most experienced users. Various software developed to exploit the affordances of the HTC Vive's wand controllers use the touchpad to type on halves of an on-screen keyboard. This is an innovative way to exploit the affordances of widely-available hardware to allow some form of symbolic input, but remains more cumbersome than use of a keyboard.

For systems where the user will be seated or standing at workstation, use of the familiar full-size keyboard may be desirable. Current work in VR uses chromakey techniques to composite the view from a HMD-mounted camera into the VE displayed on a fully-occluded HMD. This technique is effective but requires the physical elements that are to be removed to be painted in a specific color, and lighted appropriately. These constraints are likely not appropriate to the working environment of analysts performing HII tasks. Depth sensors provide a more feasible approach, assuming that appropriate techniques can be deployed to isolate the desired elements of the physical world, such as the user's body and input devices. Latency and avoiding simulator sickness remains an issue with this approach, when a fully-occluded HMD is employed.

6. SEARCHING AND BROWSING FOR INSIGHT

When confronted by a problem for analysis, users must choose whether to search for specific information or browse through available information in search of relevant insight. Browsing and searching strategies are employed by users trying to identify relevant information to satisfy an information need.

The Library Information Sciences discipline has studied browsing behavior in the context of library patrons exploring collections.[33] While this approach is built around the document as an artifact, similar strategies occur when browsing a collection of information, and could be considered in terms of navigation through information visualizations, whether in MR or in screen-based interactions.

The Information Retrieval (IR) community engages with algorithmic strategies to identify and obtain relevant information from collections based on queries. The particular features and query techniques used vary depending on the collection and context. Fine details of IR approaches are beyond the scope of HII with MR, but effective MR HII systems will need to provide understandable user interfaces to IR systems in order for users to identify and retrieve the information they need. The most obvious example of a query is the free-text search box of contemporary search engines. Other queries include query-by-example,[34] where the user provides a source piece of content and features are used to retrieve similar content. This query mode has been explored for images,[35] or user sketches of images.[36] Geographic Information Systems (GIS) have employed query-by-sketch.[37] Query by example in graph structure applied to the social graph could provide an analyst with valuable insight into patterns. Research into query building in graph data structures has explored 6DOF interations.[38]

Whatever the query format, effective HII systems must permit users to search for and retrieve relevant information. Matching IR techniques with query entry techniques and visual analytics approaches to visualize the results in an actionable way will be an important facet of MR HII systems.

7. SHARING MIXED-REALITY EXPERIENCES

No human endeavor is completed in isolation. HII in MR must accomodate this reality. For reasons of cost management, not all research will directly engage with multi-user systems, but any complex fielded system will need to accommodate communications between users who are physically and/or virtually co-present, as they interact with information in MEs to accomplish tasks. Sharing MR experiences is an emergent area of study, and presents a wealth of new and challenging areas of inquiry. For this reason, this section takes a more theoretical look at the classes of research problems presented by HII in ME. This

Depending on the technology stack used to present the ME, different techniques and approaches may need to be leveraged to capture and present telepresent participants, and to integrate physically co-present participants with virtual elements. However the task is achieved technically, it remains important for users to be able to share body language and facial expressions that accurately reflect their state, with relatively high accuracy and low latency. Approaches to this challenge include Holoportation,[39] where a participant is captured by sensors located in an environment and projected into remote ME, or the instrumentation of HMDs.[40]

7.1 Asymmetry in Relative Positions

MR and telepresence necessarily create some level of asymmetry between non co-located participants. This asymmetry is likely most challenging for systems closer to the AR end of the RV continuum. Virtual elements must be placed in context to a real environment that varies between participants, but may include multiple participants copresent in one physical location. VR systems have a somewhat simpler set of challenges, since virtual elements and participants may be unmoored from physical context.

Display solutions must determine how to present the telepresent users to the local users in a way that is consistent with their current environment, and also consistent with the remote users' relationship to the data. Redirecting gaze and gesture may be necessary if relative positions of users and environment differs across environments due to physical space and layout constraints. VR approaches that occlude more of the physical environment of each participant have less overhead in terms of synchronizing with the physical environment of multiple users.

7.2 Asymmetry in Local View

Challenges still remain if, for example, the ME allows users to hold different views of the same information-space simultaneously. Users' views into a MR information space may drift in and out of symmetry depending on their needs. It may be fruitful to consider user-local and team-global lenses on a ME, and define ways to share control of the team-global representation, while still allowing individual users to maintain and manipulate user-local representations.

Different classes of HII tasks may require greater or lesser communication and synchronization between users. A meeting to present results to a superior or pass tasking to subordinates has different requirements, for example, compared to a collaboration between analysts.

7.3 Asymmetry in Information

Participants may differ in preferences, affordances, and need-to-know. Personalized display technologies mean that two colleagues in the same ME can observe overlapping but not identical views, filtered according to their requirements and access.

7.4 Asynchronous Access and Artifacts

Participants may also interact with the same information and problems asynchronously. This suggests the need to save and share arrangements or presentations of information. A saved view or presentation could have utility as a boundary artifact that facilitates understanding between analysts regarding a particular analysis problem.[41]

8. CONCLUSION

The confluence of Mixed Reality and Human Information Interaction creates opportunities for deep inquiry within disciplines and for the fusion of results across traditional disciplinary boundaries. Research must address when and how MR can best serve HII tasks to improve outcomes. HII contexts must be evaluated to determine their suitability for MR representation, and some tasks will be more suitable for spatial MR interfaces than others. Opportunities exist in the realm of data representation, translation between abstract and MR MR environments, and designing effective user interfaces for MR HII applications.

9. ACKNOWLEDGEMENTS

The work depicted here is sponsored by the U.S. Army Research Laboratory (ARL) under contract number W911NF-14-D-0005. Statements and opinions expressed and content included do not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

[1] Milgram, P., Takemura, H., Utsumi, A., and Kishino, F., "Augmented reality: A class of displays on the reality-virtuality continuum," in [*Photonics for industrial applications*], 282–292, International Society for Optics and Photonics (1995).

[2] Azuma, R. T., "A survey of augmented reality," *Presence: Teleoperators and virtual environments* **6**(4), 355–385 (1997).

[3] Fidel, R., [*Human information interaction: an ecological approach to information behavior*], MIT Press (2012).

[4] Cook, K. A. and Thomas, J. J., "Illuminating the path: The research and development agenda for visual analytics," tech. rep., Pacific Northwest National Laboratory (PNNL), Richland, WA (US) (2005).

[5] Papert, S., [*Mindstorms: Children, computers, and powerful ideas*], Basic Books, Inc. (1980).

[6] Peter, L. and Varian, H. R., "How much information," research project, University of California at Berkeley (2003). Retrieved from http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/.

[7] Turner, V., "The digital universe of opportunities: Rich data and the increasing value of the internet of things." https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm (2016). Accessed 9 Mar 2017.

[8] Chen, C. and Chen, C., [*Mapping scientific frontiers*], Springer (2003).

[9] Clark, H. H., Brennan, S. E., et al., "Grounding in communication," *Perspectives on socially shared cognition* **13**(1991), 127–149 (1991).

[10] Pirolli, P. and Card, S., "Information foraging.," *Psychological review* **106**(4), 643 (1999).

[11] Patterson, E. S., Roth, E. M., and Woods, D. D., "Predicting vulnerabilities in computer-supported inferential analysis under data overload," *Cognition, Technology & Work* **3**(4), 224–237 (2001).

[12] Card, S. K., Mackinlay, J. D., and Shneiderman, B., [*Readings in information visualization: using vision to think*], Morgan Kaufmann (1999).

[13] Phan, T., Krum, D., and Bolas, M., "ShodanVR: Immersive visualization of text records from the Shodan database," in [*IA'16: Proceedings of the 2016 Workshop on Immersive Analytics at VR'16*], (2016).

[14] Allison, R. S., Gillam, B. J., and Vecellio, E., "Binocular depth discrimination and estimation beyond interaction space," *Journal of Vision* **9**(1), 10–10 (2009).

[15] Ware, C. and Franck, G., "Evaluating stereo and motion cues for visualizing information nets in three dimensions," *ACM Trans. Graph.* **15**, 121–140 (Apr. 1996).

[16] Bartram, L. and Ware, C., "Filtering and brushing with motion," *Information Visualization* **1**, 66–79 (Mar. 2002).

[17] Plumlee, M. D. and Ware, C., "Zooming versus multiple window interfaces: Cognitive costs of visual comparisons," *ACM Trans. Comput.-Hum. Interact.* **13**, 179–209 (June 2006).

[18] Irani, P. and Ware, C., "Diagramming information structures using 3d perceptual primitives," *ACM Trans. Comput.-Hum. Interact.* **10**, 1–19 (Mar. 2003).

[19] Ware, C. and Knight, W., "Using visual texture for information display," *ACM Trans. Graph.* **14**, 3–20 (Jan. 1995).

[20] Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., schraefel, m. c., and Wobbrock, J. O., "Reducing legacy bias in gesture elicitation studies," *interactions* **21**, 40–45 (May 2014).

[21] Bowman, D., Kruijff, E., LaViola, J., and Poupyrev, I., [*3D User Interfaces: Theory and Practice*], Addison-Wesley, New York, New York (2004).

[22] Bowman, D. A., Koller, D., and Hodges, L. F., "Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques," in [*Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*], 45–52, 215 (Mar 1997).

[23] Interrante, V., Ries, B., and Anderson, L., "Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments," in [*2007 IEEE Symposium on 3D User Interfaces*], (March 2007).

[24] Razzaque, S., Kohn, Z., and Whitton, M. C., "Redirected walking," in [*Proceedings of EUROGRAPHICS*], **9**, 105–106, Citeseer (2001).

[25] Thorndyke, P. W. and Hayes-Roth, B., "Differences in spatial knowledge acquired from maps and navigation," *Cognitive psychology* **14**(4), 560–589 (1982).

[26] Ingram, R., Benford, S., and Bowers, J., "Building virtual cities: Applying urban planning principles to the . . .," in [*PROCEEDINGS OF THE ACM SYMPOSIUM ON VIRTUAL REALITY SOFTWARE AND TECHNOLOGY (VRST96*], 83–91, ACM Press (1996).

[27] Vinson, N. G., "Design guidelines for landmarks to support navigation in virtual environments," in [*Proceedings of the SIGCHI conference on Human Factors in Computing Systems*], 278–285, ACM (1999).

[28] Bolt, R. A., "“put-that-there”: Voice and gesture at the graphics interface," *SIGGRAPH Comput. Graph.* **14**, 262–270 (July 1980).

[29] Bowman, D. A. and Hodges, L. F., "An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments," in [*Proceedings of the 1997 Symposium on Interactive 3D Graphics*], *I3D '97*, 35–ff., ACM, New York, NY, USA (1997).

[30] Pierce, J. S., Forsberg, A. S., Conway, M. J., Hong, S., Zeleznik, R. C., and Mine, M. R., "Image plane interaction techniques in 3d immersive environments," in [*Proceedings of the 1997 Symposium on Interactive 3D Graphics*], *I3D '97*, 39–ff., ACM, New York, NY, USA (1997).

[31] Thomas, B., Tyerman, S., and Grimmer, K., "Evaluation of text input mechanisms for wearable computers," *Virtual Reality* **3**(3), 187–199 (1998).

[32] Bowman, D. A., Rhoton, C. J., and Pinho, M. S., "Text input techniques for immersive virtual environments: An empirical comparison," in [*Proceedings of the Human Factors and Ergonomics Society Annual Meeting*], **46**(26), 2154–2158, SAGE Publications Sage CA: Los Angeles, CA (2002).

[33] Chang, S.-J. L., "Chang's browsing," in [*Theories of Information Behavior*], Fisher, K. E., Erdelez, S., and McKechnie, L., eds., ch. 7, 69–74, Information Today, Oxford (2005).

[34] Zloof, M. M., "Query by example," in [*Proceedings of the May 19-22, 1975, national computer conference and exposition*], 431–438, ACM (1975).

[35] Smith, J. R. and Chang, S.-F., "Visualseek: A fully automated content-based image query system," in [*Proceedings of the Fourth ACM International Conference on Multimedia*], *MULTIMEDIA '96*, 87–98, ACM, New York, NY, USA (1996).

[36] Di Sciascio, E., Mingolla, G., and Mongiello, M., "Content-based image retrieval over the web using query by sketch and relevance feedback," in [*International Conference on Advances in Visual Information Systems*], 123–130, Springer (1999).

[37] Egenhofer, M. J., "Query processing in spatial-query-by-sketch," *Journal of Visual Languages & Computing* **8**(4), 403–424 (1997).

[38] Ware, C. and Bobrow, R., "Motion to support rapid interactive queries on node–link diagrams," *ACM Trans. Appl. Percept.* **1**, 3–18 (July 2004).

[39] Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., Kim, D., Davidson, P. L., Khamis, S., Dou, M., Tankovich, V., Loop, C., Cai, Q., Chou, P. A., Mennicken, S., Valentin, J., Pradeep, V., Wang, S., Kang, S. B., Kohli, P., Lutchyn, Y., Keskin, C., and Izadi, S., "Holoportation: Virtual 3d teleportation in real-time," in [*Proceedings of the 29th Annual Symposium on User Interface Software and Technology*], *UIST '16*, 741–754, ACM, New York, NY, USA (2016).

[40] Olszewski, K., Lim, J. J., Saito, S., and Li, H., "High-fidelity facial and speech animation for vr hmds," *ACM Trans. Graph.* **35**, 221:1–221:14 (Nov. 2016).

[41] Bergman, M., Lyytinen, K., and Mark, G., "Boundary objects in design: An ecological view of design artifacts," *Journal of the Association for Information Systems* **8**(11), 546 (2007).