Social Interactions in Multiscale CVEs

Xiaolong Zhang
School of Information
University of Michigan
Ann Arbor, MI 48109-1092, USA
+1 734 647-8032
xiaolong@umich.edu

George W. Furnas
School of Information
University of Michigan
Ann Arbor, MI 48109-1092, USA
+1 734 763-0076
furnas@si.umich.edu

ABSTRACT
A *multiscale Collaborative Virtual Environment* (mCVE) is a virtual world in which multiple users can independently resize themselves to work together on different sized aspects of very large and complicated structures. Interactions among users in an mCVE differ in many ways from those in traditional collaborative virtual environments. In this paper we explore collaboration-related issues affected by multiscale, such as social presence, perception of proximity, and cross-scale information sharing. We also report results of an experiment with our mCVE prototype system, which show the impact of multiscale capabilities on social interactions.

Categories and Subject Descriptors
I.3.6 [**Computing Methodologies**]: Methodology and Techniques -- *interaction techniques*; H.5.3 [**Information Interfaces and Presentation]**: Group and Organization Interfaces – *Computer-supported cooperative work, Synchronous interaction.*

General Terms
Design, Human Factors.

Keywords
Multiscale, CVE, Awareness, Presence, Proximity.

1. INTRODUCTION
Collaborative virtual environments (CVEs) have become an emerging tool in supporting research[28], training[23], education[9][21], and community activities[22]. Many CVE systems are designed for the purpose of using VR technologies to enhance our real world experiences. While many virtual environments (VEs) are designed to simulate reality, it is often valuable to consider how VEs can go beyond reality [18].

Many constraints of everyday physics do not exist in VEs. Physical parameters such as speed and space can be transcended. For example, in the real world, navigation requires traversing physical space between two locations with a certain speed. In a virtual world, however, navigators can "teleport" themselves directly to a destination without traversing space. The absence of these physical constraints provides many opportunities for innovation in the design of VEs and CVEs.

A *multiscale* Collaborative Virtual Environment (mCVE) exploits such an opportunity, supporting collaborative work on huge structures by allowing people to manipulate size scales explicitly, in a way not possible in the real world. A person in a single-user multiscale Virtual Environment (mVE) can manipulate the scale of the whole virtual space. Working on a virtual planet, for example, the user can magnify the virtual world to see the atomic structures of objects on that planet, or shrink the world to see how this planet is related to others. The user does not need microscopes and telescopes, but can simply magnify or shrink the whole world to examine objects at various length-scales. For multiple users working together, an mCVE allows them to collaborate using such re-scaling capabilities, enhancing their ability to control and manage large and complex structures. Imagine two collaborators standing in a VE around a shared planetary model, magnifying and studying it together. If different users want to work at different scales, it is useful to flip the metaphor, and have users resize themselves relative to the world. Each user can do so independently, and the result is an mCVE, a world in which ant-sized and giant-sized actors can work together on different aspects of a shared structure.

An mCVE could be a prominent tool, for example, in the support of cross-scale collaboration in scientific research, where increasing complexity requires collaborations among scientists from a variety of fields. The traditional research focuses of individual disciplines are often on different length scales, so cross-scale collaboration may be needed in cross-disciplinary research. For example, the analysis of metal cracks may need collaboration among people from engineering, materials science, and chemistry. Their expertise with different length scales can help investigate problems ranging from the mechanical properties of materials at macroscopic scales (e.g., stress), to those of material structures at a scale of thousands of atomic diameters, to chemical bonds at atomic scales. An mCVE can bring people in different areas together and allows them to work together in a common environmental context, thus making cross-scale collaborations easier. One could envisage that such an mCVE approach will increase efficiency in collaborative research and enable researchers to think in new ways.

The practical effectiveness of mCVEs will be published elsewhere. In this paper we focus on some interesting social dimensions of multi-scale collaboration. First, we briefly introduce what an mCVE is and what it can offer. Next, we examine several emerging issues in multiscale social interactions. Results from a proxemics experiment then illustrate one of the subtle social consequences of multiscale collaboration. A final

CVE'02, September 30–October 2, 2002, Bonn, Germany.

discussion outlines potential application areas of mCVEs and some implications for further research.

2. mCVEs AND THEIR APPLICATIONS

The virtual space in an mCVE is one enhanced by multiscale technologies inspired by 2D multiscale ("zoomable") user interfaces (ZUIs)[24][4]. In 2D multiscale virtual space, the typical metaphor has users scale up and down the environment, zooming in and out as desired. As described above, in mCVEs, we use the dual metaphor, with users scaling themselves up and down and by controlling how big they are, determine at what size scale the virtual world is observed and manipulated. In a collaborative setting, the users appear to each other in various sizes. When scientists are investigating a new material together in an mCVE, for example, some may shrink themselves to the atomic levels, becoming "ants" (or even "nano-ants" or "nanants"), while others remain large (relative "giants" or even "gigants".) Figure 1 shows two users at different scale levels.

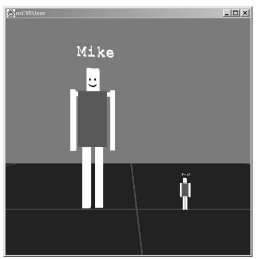


Figure 1: Giant and Ant Users in mCVEs

Working as an ant or a giant, users will have different perception and action domains. Multiscale techniques give users the freedom to control dynamically a whole set of size-related interaction parameters, including viewpoint position (notably viewing distance and eye-height), stereo-eye separation, locomotion speed, and reaching distance. By choosing different working scales, users will see objects rendered with different sizes and various degrees of details, get different overview ranges, and have various size-tuned navigation and objection selection capabilities.

The combination of multiscale and collaboration brings together two important approaches to working with large and complex structures. First, collaboration allows dividing large tasks into sub-tasks and conquering them individually, in parallel. Thus, mCVEs, like CVEs, should be very helpful in supporting the management of structures that require different experts to work together (e.g., in complex engineering design) or where real-time dynamic changes require the simultaneous work of multiple people, just to keep up (e.g., air traffic control.). Second, multiscale techniques provide explicit support for working on increasingly large and complex worlds that demonstrate important structure at many different scales. An mCVE, therefore, should be of particular value when virtual worlds and the tasks within them are too large and complex for a single user, and for working at a single scale. In such situations, multiple actors must work together across different length scales, coordinating small, remotely separated details; or managing the real-time interaction of many details with large scale features.

While an mCVE provides more opportunities for users to work on complicated structures, it also posts new challenges. Interactions among ant and giant users will be different from interactions among users who are at the same or comparable size scale in traditional CVEs.

3. SOCIAL INTERACTIONS IN mCVEs

Dix[10] argues that there are two types of collaboration: communication-centered and artifact-centered. While the former focuses on contents and implications of exchanged messages, the latter emphasizes the mutual understandings of artifacts and users' activities related to artifacts. Although CVEs have been seen as a tool to support communication-centered collaboration[8][22], other technologies (e.g., chatting and video) tend to be much simpler to deploy and easier to use. The capability of CVEs to present objects and data in complicated ways in 3D space indicates the potentials of CVEs in supporting artifact-centered collaboration.

New issues emerge when multiscale is introduced, in mCVEs, to support work on artifacts. Artifacts are usually shared in a workspace, and the presence of participants in the workspace is a critical awareness cue for collaboration[15]. A multiscale 3D space may increase the difficulty in providing appropriate awareness information. In traditional CVEs *place* is an important variable in interaction. Users at different places see different things, and that can interfere with their ability to have common ground and communicate. In mCVEs, *scale* is an additional factor. Users, even at approximately the same place, will see different things when working at different scales (e.g., atoms vs. macro surfaces). The users working at different scales have not only different perceptions, but also different locomotion and manipulation capabilities. Furthermore, artifact-centered collaboration demands the recognition of artifacts referred by other users in a process called *deixis* [10]. Multiscale tools, again, could hinder this process by presenting totally different artifacts at different scales. Discussions following will focus on interaction issues related to social presence, one kind of spatial perception – proximity, and deixis.

3.1 Social Presence

CVEs, as social systems[3][22], need social presence to shape social conventions[2]. An avatar is a very common social presence cue revealing the existence of a user, her location, and her identification[7]. In the mCVE we developed, avatars are still the primary cue for social presence. In addition, avatar size also shows users' interaction scale: how far they can see, how distant they can reach, how fast they can move, etc. – important information for users to interpret and coordinate with each other.

3.1.1 Visibility of Avatars

Scalable avatars introduce new problems for social presence. They could become too small to be seen by others, or be too big to be entirely visible. Without a good view of avatars, social presence and further social interactions could be hurt.

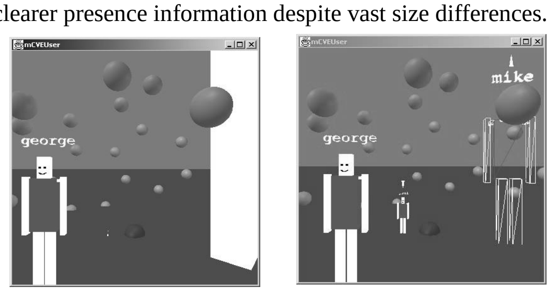
The issue here is the conflict between presenting interaction scale information and other information related to the user in the same object – an avatar. In regular CVEs, the rendered size of an avatar is just related to its distance from the viewer. When the avatar is very small due to a great viewing distance, the viewer is usually not interested in the avatar and embodied information (what direction they are facing, whether they are "smiling") is less important. In an mCVE, however, when the interaction scale of a user is conveyed directly by the avatar size, the embodiments of other information would be hurt due to the poor visibility of features of a small avatar.

One possible solution we considered was to detach the information about interaction scale from the avatar. The avatar would remain a fixed size, to broadcast other relevant social information, and we would use a secondary graphical object, rather than the avatar body, to portray scale information. Using different objects to show different attributes of a user is very common in CVEs. For example, the identity of a user is often represented by a separate graphical object, a nametag, associated with an avatar. Embodiment of identity is separated from that of user's location, view orientation, activities and so on. Often the nametag and the avatar are grouped together as if they are one entity, making the distinction between different embodiments not obvious.

Using a separate graphical object to indicate a user's interaction scale requires a mapping scheme between scale and a certain attribute of the secondary object. Mapping scale onto such attributes as color and shape could make it difficult for users to make size comparisons, because color and shape, being qualitative attributes, can hardly provide quantitative information about scales. Mapping scale onto quantitative attributes, such as object size, is a possible solution, but it also suffers the visibility problem when the secondary object is too small or too big.

Another problem of separating the embodiment of interaction scale from avatar body was that with a uniform avatar size, users would not be well informed about others' interaction scales. To understand why avatars with the same size behave differently (e.g., different moving speeds), users need to make additional efforts to find the object that embodies scale information, and interpret it. This demands more cognitive work. In comparison, obtaining such information from scalable avatars is more straightforward and direct.

Our solution to the embodiment conflict is to uses avatar size directly to represent the corresponding user's interaction scale – but within limits. Beyond those limits, the avatar size is designed to stay usefully visible to others when it would otherwise be too small or too large. By this, avatars are rendered with what we call "scale-dependent representations", a technique borrowed from 2D multiscale environments, where it is called "semantic zooming". Specifically, we use a technique called "sticky Z" in 2D ZUIs (where there Z was the magnification or scale parameter): When the size of the avatar is beyond a maximum size or a minimum size, it will be rendered with a size-fixed representation. Figure 2 compares the same view of three avatars before and after they are rendered semantically. In (a), the big block on the right is a huge avatar, and only part of its body is visible. The avatar on the left is seen as normal. The third avatar is too tiny to be seen easily. With scale-bounded representations, both the tiny and giant avatars appear with a visible size in (b). The small white, pointy "caps" above their nametags indicate the visible body size is not their real size. The large avatar is also rendered as a wire-frame model to let the viewer see the world behind the body. In this way, users can get clearer presence information despite vast size differences.



(a) (b)

Figure 2: Semantically Rendered Avatars

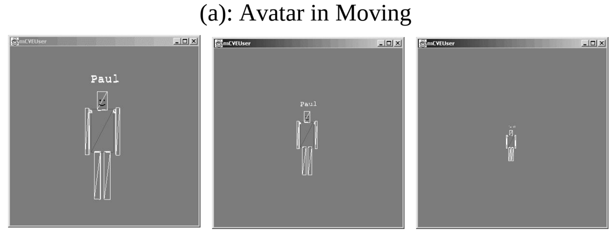
If the viewer does need the information about the real size of an avatar, various strategies are possible. A mouse-over event or a toggle tool can switch the representation of the avatar between the real size and the distorted size. In a more sophisticated dual representation method, the avatar can be shown in two sizes at once. A too-small avatar might be seen as a bright red point at its true location, and a larger visible ghost avatar around that point manifests its other visible features. A too-large avatar would be a ghost presence at some reasonable and informative size, with its red wire-frame indicating its true size and position. However, when a user is facing others whose scales are much larger or smaller than hers, she may not always care about their exact scale values of others. Thus, we implemented the toggle tool version that allows the user to retrieve others' scale values whenever they are needed and switch back to scale-bounded representations to reduce the complexity of avatars whenever they are not wanted.

3.1.2 Avatar Representation in Scaling and Moving

In the mCVE, the size of an avatar as presented on the screen is not only determined by that avatar's interaction scale. It is also determined by its distance from the viewer. This presents a challenge for users to identify correctly what other users are doing when the rendered size of their avatars appears to change: have they shrunk or moved away? This is particularly a problem in those circumstances where independent depth cues are lacking.

We explored a design to distinguish the visual results between scaling and moving by differentiating the appearance of avatars in these different actions. While the user is re-scaling, her avatar body is changed from a solid model to a wire-frame one, making it clear to any on-lookers that the user is changing her interaction scale, not her position. Figure 3 compares these two different appearances of an avatar: (a) is the usual representation, while (b) is what an avatar looks like during re-scaling.

(a): Avatar in Moving



(b): Avatar in Re-scaling

Figure 3: Different Avatar Representations

Of course, other design choices will work well as long as they can distinguish avatars in two different action states. For example, the avatar body can be rendered with a bright color during re-scaling to alert other users. Or the avatar can be rendered with other kind of graphical objects. We used the wire-frame body for two reasons. First, switching between a solid and wire-frame body is very easy for users to understand, and this can help not to increase users' cognitive load significantly. Secondly, a wire-frame body can reduce the area of blocked views, especially when an avatar is scaled up and tends to occupy a large amount of screen space.

3.1.3 Social Dominance

Informal experience with the mCVE, as well as existing literature, point to a possible interesting social dominance complication in multiscale collaboration: Avatar size may affect users' perception of social power, and thereby influence their social interactions. In real life, the physical appearance of people, including height, has been found to be a predictor of social dominance[17][26]. A perceived artificial height of users caused by camera placement has been found to affect people's behaviors in video-mediated communication[19]. In traditional CVEs, avatars are usually set to have similar size, and so height itself embeds few social status signals. In mCVEs, however, different avatars can be of dramatically different sizes, and one might expect some social dominance effects as a result. Indeed, in informal use Giant avatars do seem somewhat intimidating to Ants. It is further interesting to wonder what the effect will be of the fluid change of avatar sizes -- different avatar heights could make the same user be perceived with different social powers at different times, or alternatively to mitigate the size/power effect altogether. Further investigation is needed to find out whether and how the avatar size would affect collaboration, and what design strategies might be used to ameliorate unwanted effects. If the impact of height was found to be a significant issue during certain social interactions, avatars may need to be distorted to reduce the negative consequences. In the mCVE we developed, two users can choose to adjust their avatars to be comparable temporarily in a meeting. After the meeting, their avatars are restored to their original sizes.

3.2 Proxemics

The study of proxemics concerns the perception and negotiation of interpersonal distance in social interactions[1][16]. In real life, proximity, the inter-person distance, is important to social interactions. Hall[16] distinguishes four proximity ranges at normal human scale: *intimate* distance (less than 0.45m), *personal* distance (0.45 to 1.2 m), *social* distance (1.2 to 3.6m) and *public* distance (larger than 3.6m). People choose an appropriate distance range based on their social needs, and behave accordingly.

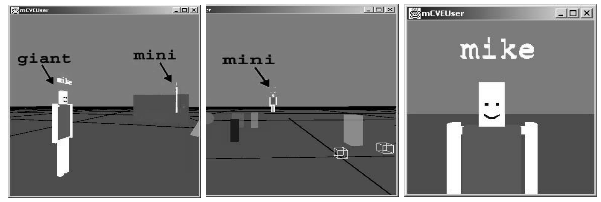
3.2.1 Asymmetrical Proximity Perception

In VEs, proximity, the distance between avatars, has been used to mediate interpersonal communications. The aura, focus, and nimbus mechanism[6] explicitly uses interpersonal distance to enable or disable communications. Becker[3] finds users are quite aware of and sensitive to proximity in graphical environments like CVEs.

Visual information about another person at different distances varies greatly with the proximity range. At *intimate* distances, only one third of the face is easily seen without significant movement of eye and head. At *personal* distance, the whole head and the shoulder can be easily seen, but the other part of the body is out of the range of clear vision. At *social* distance, people will be able to see the whole body of the other. At *public* distance, the whole body and lots of space around it will be visible[16].

Body sizes in real life and avatars size in traditional CVEs do not typically differ much from one person to the next. As a result, what participants can see about each other, and do to each other are fairly comparable, and their sense of proximity is therefore reasonably symmetric.

This symmetry often does not hold in mCVEs, where the size of avatars is no longer uniform. As seen in Figure 4, two avatars at two different scales, a giant and a "mini", are standing face to face while their eye-levels are set equal. With only visual information as proximity cues, scaled avatar size could be misleading. The giant can see the whole body of the mini, and would feel the distance between them as public distance. The mini can only see the big head and the shoulder of the giant, and tend to see the distance more as personal.



(a) (b) (c)

Figure 4: Asymmetrical Proximity Perception. (a) is a third-person view of two users, a giant and a mini. (b) is the view of the mini from the giant. (c) is the view of the giant from the mini.

In general, any asymmetric perception of proximity between users could affect collaborative activities. Actions a user takes based on her own perception of proximity may not be seen as appropriate and acceptable by another with a different perception of proximity. Such asymmetries have been discussed in the literature arising, most notably, from different cultural conventions[16], and resulting in awkward social dances where one person tries to move closer to get a good social distance, and the other backs off feeling an invasion of personal distance.

Such asymmetries arise mightily in mCVEs. If you are an ant, a giant can loom as large as if he were at intimate distance, yet be many of your own steps away (normally associated with public distance). Conversely, to the giant, you-as-ant will be as scarcely visible as someone quite far away (usually associated with remote public distance) yet be within the giant's close arm's reach – the giant's intimate distance. Note that there is not only a strong asymmetry between the two actors, but a strange splitting of the normally linked perception and action definitions of their social distances. For each actor, the visuals indicate one thing (closeness from the ant's view, remoteness from the giant's) yet their action consequences suggest the opposite. If the giant, for example, tried to move closer to get a better view of the ant, the move might be seen by the ant as a incredible invasion of the "private" space, and the ant may respond by retreating more. This misunderstanding of others' actions may affect collaboration performance. A similar case has been observed[6] when different user interfaces (text-based vs. 3D graphic) giver users different perception of proximity.

Note that this problem is independent of the actors' abilities to correctly judge the absolute distance between them. It is related more to how they appear and what they can do to each other at these distances. In real life, our choices of proximity are based on what we want from others and what we want to do to others. With similar or comparable body size and action capabilities, people can affect each other through the same physical distance in an approximately symmetric way, and their understandings of the implications of the distance for their actions on each other tend to be similar. A distance allowing a person to punch (or pat) another also means the latter can punch (or pat) back, and they both know that whatever they do to the other can be done by the other to themselves. This symmetry can also be held in traditional CVEs, where users' perception and action capabilities are similar.

In mCVEs, however, users may choose different interaction scales, and their perception and action capabilities could vary significantly. The same physical distance could have totally different implications for users at different scales. While the giant can quickly approach the mini or easily move objects around the mini, the mini may find it harder to affect the giant in the same way. Therefore, what is important to a user is not the physical distance to others; rather, she needs to know the social implications of the distance: what she can do to others through the distance and how it could affect social interactions.

To understand the social implications of proximity better, the user may need to see the relationship between herself, other users, and the distance, and understand how others may see and feel about the same distance. Providing access both to the other's view, and to a third party view as seen in Figure 4(a) could be helpful.

3.2.2 Different Distances for Different Actions

The style of interactions also determines the choices of distances. In conversations, besides verbal reactions, each person needs to see the non-verbal response from the other, including facial expression and body language[1]. Hall's "social distance" range is the appropriate choice for casual conversation, because it can clearly present the non-verbal responses as well as support eye contact.

In real life, when users are working together, due to the distribution of the objects, it might not be possible to maintain social distance for conversation. To allow them to stay where they are supposed to be in collaboration, they rely on other tools such as telephone or two-way radios to keep in touch verbally. In a CVE, the need for coordinating with others working on remote objects can also arise, and users may not be able to see each other's avatars in collaboration.

While we can just follow what we do in the real world by providing audio tools to help users keep in touch, we can also think of other ways to address this issue. The challenge here is actually how to have two different kinds of proximity, one for action (distant) and one for communication (close), simultaneously. In the real world, our capability of speaking and doing is unified under the same entity, our body, and we cannot simultaneously place our body at one distance for action and at another for communication. In the virtual environment, however, we can be at two places at once, with multiple embodiments[20].

In situations that require different distances for conversation and action, a secondary avatar, or a *dæmon* avatar[1] can be created to engage in remote conversation while the primary avatar stays for action. In mCVEs, the size of the *dæmon* avatar can be independent of the interaction scale so that two users' *dæmon* avatars can see each other to maintain social distance and "eye-contact" for conversation while their primary avatars stay put, far apart.

One challenge for multiple embodiments is how the *dæmon* avatar should be manifestly related to its primary avatar. Multiple embodiments could confuse other users. When two users are in conversation with their *dæmon* avatars face to face, it could be a problem for a third user to understand what is happening. Are there four users or just two? Which avatar represents the real position and view orientation of the user? It is important, therefore to render the *dæmon* avatar distinctly, so that other users can see that it is not the primary delegate of the user in the virtual environment. For example, while the primary avatar appears as a full body model with solid color, the *dæmon* is rendered as a semi-transparent head. *Dæmon* avatars can also have distinct identity labels or appear as other distinctive different shapes.

Besides giving a *dæmon* avatar an appearance distinguishable from the primary avatar, the correspondence between the primary and *dæmon* avatars should also be clearly indicated. A *dæmon* avatar can be far away from its primary avatar to maintain "social distance" to another user, and when there is more than one *dæmon* avatar, it could be a problem to know which *dæmon* avatar is affiliated with which primary avatar. A visual indication of connection between a *dæmon* with its primary avatar is needed. One choice is to limit the separation between the two. For two users who are close but cannot see each other due to being at different scales, this approach works well. However, for those avatars that are very distant, but still hope to maintain eye contact, limiting the action range of the *dæmon* avatar will not be helpful. A better choice could be to connect two avatars by such attributes as color and shape, to use identity labels to link two avatars as a pair, to highlight the two avatars together when the cursor is over either of them, or other ways.

Having a *dæmon* avatar means users need to see what the *dæmon* avatar sees. If the views of the primary and *dæmon* avatars are not required simultaneously, a toggle tool would be sufficient to let the user switch between two views. If the two views are needed together, a secondary view can be provided. This secondary view is just like the portal tool seen in 2D ZUIs[4], which gives the user an extra view of a distant place and lets the user manipulate the scale of the virtual space presented in the portal window.

Figure 5 is the view of a *dæmon* avatar with its primary avatar in our implementation. The *dæmon* is rendered just as a head with a nametag, the prefix of which tells this object is a *dæmon* and in which the identity of the user is also included. While the primary avatar, which is located at the bottom of the window, is almost out of the viewer's sight, the *dæmon* still maintains eye contact with the viewer. The user can toggle between the views of the *dæmon* and the primary avatar.

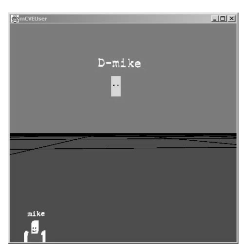


Figure 5: Dæmon Avatar

3.3 Sharing Context Across Scales in Deixis

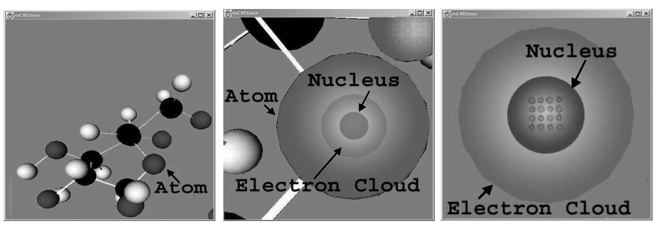
To understand what objects others may refer to, a user may need to see what others are seeing. This requires a tool allowing users to share others' views and to know the working context of others. This can be supported by having multiple views[14] or by seeing others' views[11]. Such techniques may work well in traditional CVEs, where users share the same world with same objects but from different viewpoints. An mCVE, however, could make this context sharing more difficult.

---

[1] The use of "*dæmon*" is inspired by *The Golden Compass*[25].

3.3.1 Scale-Based Semantic Representations

Earlier we used the technique of "scale-dependent representation" to keep others' avatars from becoming too small or too large when they resize. This scale-based representation technique has numerous other uses for helping even individual users work in multiscale worlds. Like "semantic zooming" in ZUIs[4], the mCVE we developed can present any objects with successive models that do not just reveal geometric refinements as they get larger. Instead, objects as they enlarge can be rendered with different semantically meaningful visual representations, showing alternate structures and characteristics of objects to enhance user understanding at different scales. This is what allowed avatars to stay meaningful instead of shrinking out of sight, for example. Images in Figure 6 show another, non-social example -- the views of the structure of a substance at three different scales. Its molecular structure is seen in (a). When scaling herself down, the user sees the increase of the structure's size, and at the same time the atom is fading-out and the atomic structure inside atoms, the electronic cloud and the nucleus, is appearing in (b). Continuing scaling down, in (c), the atoms disappear, and the user can clearly see the nucleus and the electron cloud when the structure inside the nucleus begins to emerge. Each representation shows different characteristics of the substance, and the user is semantically informed of the multiscale characteristics by these different views. To investigate new materials, for example, scientists need this tool to get objects of interest at different scales.



(a) (b) (c)

Figure 6: Scale-Based Semantic Representations

Scale-based semantic representations, however, present a problem for collaboration, however: context sharing is more difficult. Users at different scales would see quite visually different renderings of even the "same" virtual objects. How could a user seeing the virtual world as in Figure 6(a) share working context with others who see the world as in Figure 6(c)? Simply sharing each other's view or knowing the orientation of views would not help much, because the two views are so diverged that nothing common can be found to relate them.

The divergent views caused by the semantic rendering of different interaction scales can be considered a kind of *subjective views*[27][29], with which users tailor what they see based on their own interests. One challenge in subjective views is the mutual understanding of each other's contexts. A common view relating both diverged views might provide some help[29]. Subjective views seen in traditional CVEs are usually created by rendering the same objects with different representations (e.g., a solid model vs. a wire-frame model), or adding/hiding some objects to match users' different interests[27]. However, most of these subjective views are about the same world and the same or similar scale. In such situations, an objective view of the world may indeed help users to understand others' context.

While a static common view that interests both users might be effective in traditional CVEs, it is not sufficient for users with subjective views from scale-based semantic representations in mCVEs. When two users are seeing Figure 6(a) and 6(c) respectively, what should the common view be? Is 6(b) a good candidate? Of course, it is possible that users can figure out how their views are related by comparing three images. However, in more complicated scenes, separated by many orders of magnitude, finding a useful static view that includes objects appearing in both views could be a challenge.

The objects of interest in Figure 6(a) and 6(c) are hierarchically related. Conceptually, their relationship is one between an ancestor and its descendent, similar to the relationship between two nodes of 0 and 1 in a tree in Figure 7. Because they are very close hierarchically, a static view could be structured so that the displayed contents in the view include objects seen in both views, as in 6(b). Users can see the relationship between two nodes, and understand the connection between objects in the two views through the common view. However, if the relationship between the objects two users are interested in is like that between nodes 1 and 2, bringing both nodes together in a static common view could be difficult. These two nodes are related to each other through node A, their least common ancestor, so the static common view that clearly demonstrates their relationship should include both of them as well as node A. However, the scale difference between these two nodes and A could be significant, making it impossible to create a view of A without nodes 1 and 2 disappearing. To inform users of the relationships between what they see, a static common view is not adequate.

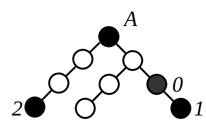


Figure 7: Relationship of Objects in Views

3.3.2 Dynamic View Transition

One way to address this issue is to use a dynamic view, instead of just a static one, to bring two divergent views together. We designed this dynamic view as an animation to show the transition between two views. To connect the views of Figure 6(a) and 6(c), for example, an animation can be created by showing more intermediate views, like 6(b), between them, and through animations, users will know how their views can be transformed from one to another.

Generally speaking, the view of a user can be written as $V(P, O, S)$, where P, O, and S are the view position, view orientation, and scale of the user respectively, and the view animation between two views can be written as

$$V_0 \overset{f}{\Leftrightarrow} V_1$$

where $V_0(P_0, O_0, S_0)$ and $V_1(P_1, O_1, S_1)$ are the views of two users respectively, and f is the view transition function, determined by the path between V_0 and V_1. Inspired by the Space-Scale Diagram[13], V_0 and V_1 can be seen as two points in a seven-dimensional view space that is defined by P(three variables), O(three variables), and S. The f is a path connecting these two points. The animation can be created by assembling views along the path, the trajectory of which can appear in any form. In our implementation, the path f is a simple piece-wise linear function.

When the relationship between the interesting objects in two views is more complicated, like that between nodes 2 and 1 in Figure 7, directly linking two views may not suffice to help users see the big picture. One solution to this problem is to find the structure that is the least upper bound of the objects in the two views, and then to generate paths between these two views and passing through the view of the bounding structure. For the case of nodes 2 and 1, the node A is used to create the view transition as two segments, written as:

$$V_2 \overset{f_2}{\Leftrightarrow} V_A \overset{f_1}{\Leftrightarrow} V_1$$

where V_A is the view of the node *A*.

There is one challenge for the design of this kind of two-segment view animation. It is required to identify what objects users are seeing in a given view so that the least upper bound of the contents in two views can be calculated. Therefore, a function that maps an arbitrary view Vi to its view contents has to be pre-defined. For a very complicated and very large structure, it could be a daunting task to create such a function.

This animated view transition technique may also be needed in traditional CVEs when users are distributed at very distant places. When the users are very far apart, it is a challenge for them to understand each other's contexts by just sharing their local views. A common objective view that includes two very distant views could mean that the detail of the contents of each view cannot be clearly exhibited due to the large spatial span of the common view. When the context information is available, the content information is missing. In an important sense, this is really a multi-scale problem - the scale of local views and global separation are quite different – ad as such can use support, even if a full suite of multiscale tools is not provided. In addition to using traditional tools to deal with this focus and context problem, such as Fisheye Views[12], the view transition animation devised here could also be helpful by allowing users to see how two views are related and transformed from one to another across space.

4. EXPERIMENT

A desktop mCVE system was implemented by using Java 3D and Java Shared Data Toolkit (JSDT). Based on this prototype, we conducted a series of tests. Here we report the results of a test related to social interactions: how multiscale affects proxemics.

In the test, each subject was required to move to a comfortable "conversation" distance from the avatar of another, inert user who appeared at one of two different scales. Each subject encountered, in successive trials, a sequence of four different avatars in an almost empty virtual environment. Presentations of avatars formed a 2x2 design (avatar size x avatar eye-level). Two avatars were 2.5 times taller than the viewer; two were 2.5 times shorter. Two avatars were positioned to have the eye level equal to the eye level of the subject's avatar, and two stood on the same ground as the subject's avatar. These two different avatar positions reflect the fact that an mCVE can be used to present two different types of virtual worlds, one with a ground plane (e.g., a virtual city) and one without (e.g., a virtual galaxy). The different sizes of avatars represent different interaction scales of other people that subjects may meet in mCVEs. The dependent measure was final distance between a subject's viewpoint and the inert "other" avatar. Six subjects participated in the experiment.

An ANOVA shows main effects of the eye-level difference ($F_{1,20}$ =12.85, p= 0.0019) and avatar size ($F_{1,20}$=9.72, p=0.0054), and a strong interaction ($F_{1,20}$=13.23, p=0.0016). (Figure 8) In the test, at least five of the size subjects seemed to use the visibility of the whole body of the other avatar as the criterion to judge the distance. They stopped at the point where further movement would lose part of the avatar body. This observation can help to understand what factors may affect the choice of social distance.

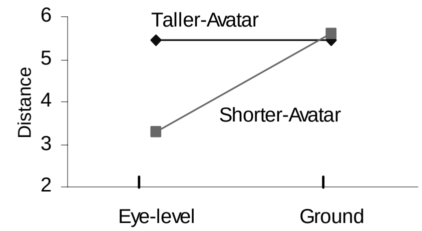


Figure 8: Distance Comparison

Given the fact that each user has a fixed view angle regardless of scale, the size and the vertical position of the avatar have clear geometric effect. As seen in Figure 9, for viewer *A* with a view angle, α, to see the whole body of an avatar B_1, the viewing distance has to be *D*. If the avatar, B_1, flies up with its body size unchanged (B_2), the distance becomes *D'*. Shrinking the size of B_2 while keeping its vertical position consistent, the preferred view distance to B_3 by the viewer becomes *D"*. Clearly, with the same view angle as that of the viewer A, the avatar B_1 can see the entire body of the viewer's avatar A, but B_3 cannot. While the distance *D"* is preferred by *A*, it is not appropriate for B_3.

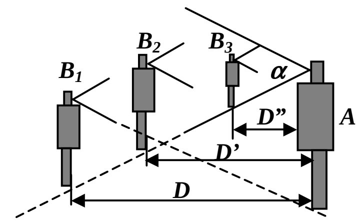


Figure 9: Avatar Size, Position and Social Distance

Social positioning problems seen in traditional CVEs[3] are indeed therefore likely to be even more serious in mCVEs. In traditional CVEs, factors like participants' different culture backgrounds may contribute to the varied understandings about the closeness. In mCVEs, however, the different sizes of avatars will have a dramatic effect on the negotiation of mutually acceptable closeness. Users need tools, like a third-person view and other's view, to inform them the social implications of the distance for each other's actions, and understand that the implications are asymmetrical to users at different scales.

5. DISCUSSION

In this paper, we discussed several social interaction issues in an mCVE. The various users' abilities to grow or shrink relative to the scale of virtual environments gives them different perception and action capabilities at different interaction scales. This in turn raises several scale-related social issues, including difficulties in maintaining social presence, asymmetries in proximity perception, and problems in cross-scale context sharing.

Resolving such issues is important because the application of mCVEs could be very broad. In this paper, we primarily focused on examples of using mCVEs to manage objects and structures across different length-scale levels. Actually, objects and structures can also exhibit various multiscale characteristics along other dimensions, such as temporal (e.g., weather patterns) and granularity (e.g., demographic distribution). When objects and structures are modeled in virtual environments, their temporal or granularity attributes can be mapped onto extrinsic spatial dimension (e.g., x, y, and z coordinates) in a virtual space[5]. For example, timelines are usually built by mapping time to one of the x, y, or z spatial coordinates. Multiscale technology can become a powerful tool to help people to understand multiscale characteristics of these objects and structures. Therefore, an mCVE may prove to be an effective tool in such areas as biology, public health, space physics, management information systems, marketing, and engineering.

Future research efforts can be extended in two directions. First, we hope to investigate other general social interaction issues, such as the impact of scale on users' activities. For example, how can users benefit from multiscale tools in such collaborative activities as navigation? Second, we would like to study potential task-specific social interaction issues. What problems may emerge when users are working on structural materials or when they are managing a nested hierarchical file system? What social issues emerge in users' adoption of this new technology in different task domains? To explore these questions, it is important to find what specific tools are needed to make mCVEs valuable in different research disciplines, and then integrate those tools into our generic mCVEs and then deploy them to real users.

6. ACKNOWLEDGMENTS

This research is funded in part by Microsoft Research.

7. REFERENCES

[1] Argyle, M., Dean H.: Eye-contact, Distance And Affiliation. Sociometry, 1965, 38(3), pp. 289-304.

[2] Becker, B., Mark G.: Constructing Social Systems through Computer-Mediated Communication. Virtual Reality, No. 4, 1999, pp. 60-73.

[3] Becker, B., Mark G.: Social Conventions in Collaborative Virtual Environments. Proceedings of CVE'98, pp. 47-55.

[4] Bederson, B.B., Hollan J.D., Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics. Proceedings of UIST'94, 1994, pp. 17-26.

[5] Benedikt, M.: Cyberspace: Some proposals. In Cyberspace: First Steps, MIT Press, 1991, pp. 273--302.

[6] Benford, S. et al: Managing Mutual Awareness in Collaborative Virtual Environments. Proceedings of VRST'94, 1994, pp. 223-236.

[7] Benford, S. et al: User Embodiment in Collaborative Virtual Environments, Proceedings of CHI'95, 1995, pp. 242-249.

[8] Greenhalgh, C., Benford, S.: MASSIVE: a distributed virtual reality system incorporating spatial trading. In Proceedings of the 15th International Conference on Distributed Computing Systems, 1995, pp. 27–34,

[9] Corbit, M., DeVarco, B.: SciCentr and BioLearn: Two 3-D implementations of CVE science museums. Proceedings of CVE 2000, 2000, pp. 65-71.

[10] Dix, A. J: Computer-supported cooperative work - a framework. In D. Rosenburg and C. Hutchison (Eds.), Design Issues in CSCW, Springer Verlag, 1994, pp. 23-37.

[11] Fraser, M. et al: Supporting Awareness and Interaction through Collaborative Virtual Interfaces Collaborative Spaces. Proceedings of UIST'99, 1999, pp. 27-36

[12] Furnas, G. W.: Generalized Fisheye Views. Proceedings of CHI'86, 1986, pp. 16-23.

[13] Furnas, G.W., Bederson, B.B.: Space-Scale Diagrams: Understanding Multiscale Interfaces Papers: Navigating and Scaling in 2D Space. Proceedings of CHI'95, 1995, pp. 234-241.

[14] Gaver, W., et al: One is not enough: Multiple views in a media space. Proceedings of INTERCHI'93, 1993, pp. 335-341.

[15] Gutwin, C., Greenberg, S.: The Effects of Workspace Awareness Support on the Usability of Real-Time Distributed Groupware. ACM Transactions on Computer-Human Interaction, 6 (3), 1999, pp. 243-281.

[16] Hall, E. T.: The Hidden Dimension. Doubleday. 1966.

[17] Hensley, W.: Height as a Measure of Success in Academe. Psychology, 3(1), 1993, pp. 40-46.

[18] Hollan, J., Stornetta S.: Beyond Being There. Proceedings of ACM CHI'92 Conference on Human Factors in Computing Systems, 1992, pp. 119-125.

[19] Huang, W. et al: Camera Angle Affects Dominance in Video-Mediated Communication. Proceedings of CHI 2002, 2002, pp. 716-717.

[20] Jää-Aro, K: Distorted, disjointed and multiple embodiments—high-tech horror cabinet or useful tools? COTECH working paper, 1996.

[21] Johnson, A. et al: The Round Earth Project: Collaborative VR for Conceptual Learning. In IEEE Computer Graphics and Applications, 19(6), 1999, pp. 60-69.

[22] Lea, R. et al: Virtual Society: Collaboration in 3D Spaces on the Internet. Computer Supported Cooperative Work. Journal of Collaborative Computing, No. 6, 1997, pp. 227-250.

[23] Oliveira, C. et al: A Collaborative Virtual Environment for Industrial Training. Proceedings of VR 2000, 2000, p. 289.

[24] Perlin, K., Fox D.: Pad: An Alternative Approach to the Computer Interface. Proceedings of SIGGRAPH'93, 1993, pp. 57-64.

[25] Pullman, P.: The Golden Compass, Ballantine Books, New York, 1995.

[26] Schwarts, B., et al: Dominance Cues in Nonverbal Behavior. Social Psychology Quarterly. Vol. 45, No. 2, 1982, pp. 114-120.

[27] Smith G., Mariani J.: Using Subjective Views to Enhance 3D Applications. Proceedings of VRST '97, 1997, pp. 139-146.

[28] Sonnenwald, D., et al: Designing to support collaborative scientific research across distances: The nanoManipulator environment. In Collaborative Virtual Environments. London: Springer Verlag, 2001, pp. 202-224.

[29] Snowdon, D., Jää-Aro, K.: A Subjective Virtual Environment for Collaborative Information Visualization. Virtual Reality Universe'97, 1997.