Don't Walk into Walls: Creating and Visualizing Consensus Realities for Next Generation Videoconferencing

Nicolas H. Lehment, Philipp Tiefenbacher, and Gerhard Rigoll

Institute for Human-Machine Communication
Technische Universität München
{Lehment,Philipp.Tiefenbach,Rigoll}@tum.de

Abstract. This contribution examines the problem of linking two remote rooms into one shared teleconference space using augmented reality (AR). Previous work in remote collaboration focusses either on the display of data and participants or on the interactions required to complete a given task. The surroundings are usually either disregarded entirely or one room is chosen as the "hosting" room which serves as the reference space. In this paper, we aim to integrate the two surrounding physical spaces of the users into the virtual conference space. We approach this problem using techniques borrowed from computational geometric analysis, from computer graphics and from 2D image processing. Our goal is to provide a thorough discussion of the problem and to describe an approach to creating consensus realities for use in AR videoconferencing.

1 Introduction

To date, videoconferencing is a mode of communication constrained to displays or specialized projective equipment. With current innovations like Google Glass and advanced Head Mounted Displays (HMDs), we can however envision a future where our remote conversation partner is no longer banished to flat displays. Instead our Avatars will appear to walk in far-away offices and labs, rendered as three-dimensional personalities by displays integrated into our glasses. This leads to the question of how we are to define the consensus reality in which we engage. Since we cannot guarantee that all participants are located in infinite, uncluttered spaces, there are bound to be discrepancies between our surroundings. If our opposite were to inhabit a large corner office, while we were to reside in a smaller cubicle, such conflicts would inevitably arise. As our conversation partner strolls over to his desk, his Avatar on our side of the connection might happen to walk right through our cubicle wall. In order to deal with such discrepancies in our environments, we aim to define a consensus reality which uses 3D scans of both rooms in order to identify common layout features. A simplified illustration of our target system is shown in Fig. 1. Note that in this paper we focus on the details of computing the consensus reality, leaving the details of user streaming and rendering to other parties (e.g. [1,2]).

In the following, we start by clarifying the context and present an overview over previous related research. We then present our method for computing the consensus reality,

R. Shumaker and S. Lackey (Eds.): VAMR 2014, Part I, LNCS 8525, pp. 170–180, 2014.

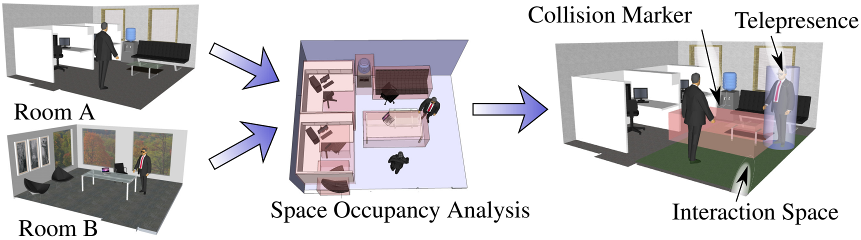


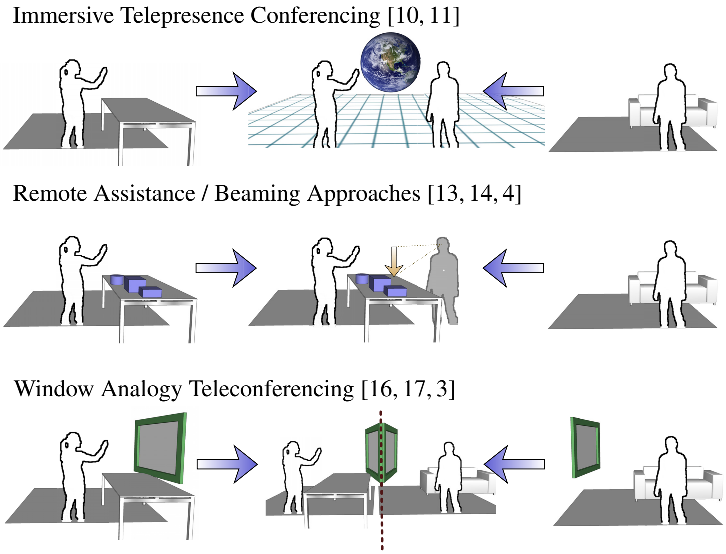
Fig. 1. Example of an augmented reality videoconference with two participants and heterogeneous environments. In a first processing step, the rooms are aligned and obstacles are identified. A map of common uncluttered floorspace is generated and used as the basis for the consensus reality. In the right image, the scene as perceived by participant A is shown: The conversation partner is added to the scene as an Avatar, the uncluttered floorspace is shown in green and a single obstacle, the desk, is represented as a red box. Participant A can thus avoid stepping into the desk and knows the basic layout of the consensus reality.

followed by a description of a simple visualization scheme. Subsequently we summarize our experiences in the conclusion and give a brief outlook on open questions.

2 Related Work and Context

Videoconferencing is a wide field of research. For our scenario, we focus on concepts which place the conversation partner directly into our environment, e.g. using augmented reality (AR) or at least mobile displays. Our previous work explored the inclusion of AR-elements into classic videoconferencing [3]. Sodhi *et al.* [4] realize a similar concept for hand-held devices in limited tabletop workspaces. Other groups have attempted to integrate conversation partners into remote locations using cylindrical displays [5], mobile social proxies [6,7] or stationary social proxies [8,9]. When we categorize previous work on virtual telepresence by the handling of the physical surroundings, we arrive at three different approaches which are illustrated in Fig. 2.

Firstly, there is the total immersion of both participants in a virtual space. This VR approach disregards the actual environment of the participants and instead provides a virtual meeting space in which interactions and discussions take place. Typical examples would be [10,11]. Another approach uses window metaphors in order to connect two real, physical spaces. In its simplest implementation, this leads to video-conferencing as familiar to users of Skype or Google Hangouts. In recent years, more elaborate versions have evolved such as the perspective sensitive display by Maimone *et al.* [12] or our AR-enhanced videochat [3]. Using the window approach, the participant's spaces are clearly separated as "things on my side of the window" and "things on the other side of the window", effectively avoiding conflicts by suspending immersion. Finally, there is the wide field of remote assistance systems. For these approaches, one user space is selected as the "hosting" space into which the remote supporter is immersed. The space on the side of the remote participant is usually disregarded, since the focus lies on solving a problem in the primary space. Typical examples are found in [13,14,4,15].



Immersive Telepresence Conferencing [10, 11]

Remote Assistance / Beaming Approaches [13, 14, 4]

Window Analogy Teleconferencing [16, 17, 3]

Fig. 2. Current approaches to treating surrounding physical space in remote collaboration research

None of these approaches considers the scenario of mutually integrating conversation partners into each other's physical environment. Therefore we ask: How can we treat scenarios where both participants want to meet on equal footing, talking between themselves in their offices? We propose a distributed augmented-reality approach: We create a consensus reality by combining both participants' offices into one shared virtual space which encompasses both locations.

Thus, we touch upon issues of scanning the room [18,19], computational geometric operations on the resulting scans [20], tracking the conversation partners [21,22] and visualizing the consensus reality [4,23,24].

3 Defining a Consensus Reality

Our approach relies on existing scans of the participating spaces. Such scans can be obtained easily and cheaply using commodity depth sensing cameras and subsequent analysis programs (e.g. using Kinect Fusion [25]). The resulting 3D meshes are then analysed by applying computational geometry techniques such as boolean operations [20]. We thus distinguish between 3 types of spaces: Those unobstructed on both sides of the conversation, those occupied by objects at one conversation partner and finally those occupied on both sides. Unobstructed spaces are then defined as free floor space, while both other types are marked as non-enterable spaces. Optionally, we can identify similar surfaces and mark these as consensus surfaces. Such surfaces can be used to place virtual content objects or 3D models. Since we know for these consensus planes that there is a similar, physical surface in both participants' environment, we can avoid having virtual objects floating mid-air. In a further refinement, we could even create a

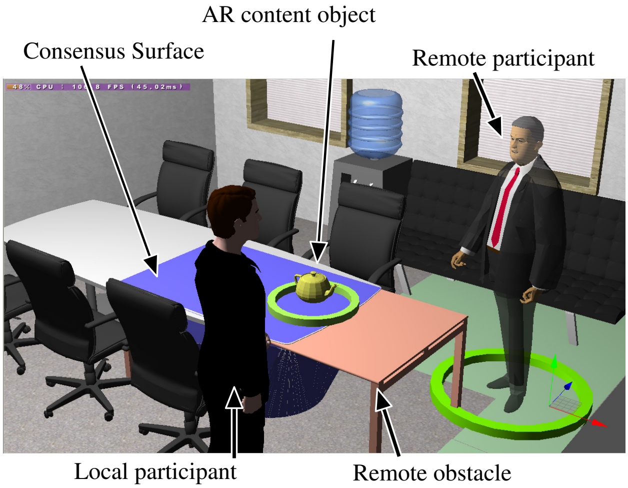


Fig. 3. Visualization of a conference in a consensus reality with static mock-up avatars. The table marked in red is not present in the local environment, but marks a table standing in the conversation partner’s office. For the sake of illustration, HMDs and their fields of view are ignored.

shared physical space, i.e. using the consensus surfaces as entities in a shared simulation of physical properties of virtual objects.

For our work, we assume the two participating rooms to have an even and uninterrupted floor plane. We do not consider multi-level rooms, stairs or sloped surface planes. We also require the scans to cover the entire floor of the intended interaction space and all obstacles within. Fig. 4 shows an exemplary scenario where a consensus reality spanning two rooms is constructed. A triangular mesh of counter-clockwise connected vertices is assumed, but does not need to be complete or closed. However, the mesh describing the floor plane of the intended workspace must be complete and free of holes. In the following, we shall consider the meshes $\mathcal{M}_\text{A}$ for room A and $\mathcal{M}_\text{B}$ for room B. The meshes are made up of individual polygons P_k which in turn consist of three vertices $\mathbf{V} = (V_x, V_y, V_z)$:

$$\mathcal{M}_\text{i} = \{P_1, P_2, \ldots, P_n\} \tag{1}$$

$$P_k = \left\{ \begin{pmatrix} V_{1,\text{x}} \\ V_{1,\text{y}} \\ V_{1,\text{z}} \end{pmatrix}, \begin{pmatrix} V_{2,\text{x}} \\ V_{2,\text{y}} \\ V_{2,\text{z}} \end{pmatrix}, \begin{pmatrix} V_{3,\text{x}} \\ V_{3,\text{y}} \\ V_{3,\text{z}} \end{pmatrix} \right\} \tag{2}$$

In a first step, we consider the submeshes $\mathcal{M}_{\text{i, Furniture}}$ containing only Polygons P lying at least partially above the floor plane (i.e. $V_{j,\text{z}} > 10\,\text{cm} \quad \exists j \in \{1, 2, 3\}$) and entirely below the ceiling (i.e. $V_{j,\text{z}} < 2\,\text{m} \quad \forall j \in \{1, 2, 3\}$). We use an orthographic projection to render the Vertices V_j contained in $\mathcal{M}_{\text{i, Furniture}}$ to the 2D floor plane. In order to facilitate subsequent processing, we define the floor plane $\mathbf{I}_{\text{i, Furniture}}$ as an 2D array of fixed size $S_\text{f} = s_x \times s_y = 10\,m \times 10\,m$ and discretize coordinates in that plane with a sampling factor of $d_\text{floor} = 1/50m$. Note that $\tilde{\mathbf{T}}^{\text{world}}_{\text{floor}}$ denotes the absolute world transformation of the floor plane, while $\triangle^{\text{filled}}_{ABC}$ signifies a filled triangle drawn by the half-space function or a similar suitable rendering function.

$$
\begin{aligned}
&\forall P_k \in \mathcal{M}_i \\
&\forall V_{kj} \in P_k \\
&\qquad \hat{V}_{kj} = \tilde{\mathbf{T}}^{\text{world}}_{\text{floor}} V_{kj} && (3)\\
&\qquad \hat{v}_{kj,\text{x}} = \lfloor d_{\text{floor}} \cdot \hat{V}_{kj,x} + 0.5 \cdot s_x \rfloor && (4)\\
&\qquad \hat{v}_{kj,\text{y}} = \lfloor d_{\text{floor}} \cdot \hat{V}_{kj,y} + 0.5 \cdot s_y \rfloor && (5)\\
&\forall P_k \in \mathcal{M}_i \\
&\qquad t_k = \tfrac{\hat{V}_{k1,\text{z}} + \hat{V}_{k2,\text{z}} + \hat{V}_{k3,\text{z}}}{3} && (6)\\
&\triangle^{\text{filled}}_{ABC}(A : \hat{v}_{k1} \rightarrow B : \hat{v}_{k2} \rightarrow C : \hat{v}_{k3}, \text{value} : t_k) \Rightarrow \mathbf{I}_{\text{i, Furniture}} && (7)
\end{aligned}
$$

In order to account for walls, we repeat this procedure for all polygons P contained in a single rooms mesh and arrive at a second 2D array $\mathbf{I}_{\text{i, InvWalls}}$ where only the areas without data are set to zero. We perform a bitwise thresholding operation on each pixel $\mathbf{p}(x, y)$ of this array in order to arrive at a map where the volumes behind walls or columns are set to 2.0 m:

$$
\forall \mathbf{p}(x, y) \in \mathbf{I}_{\text{i, Walls}} : \quad \mathbf{I}_{\text{i, Walls}}(x, y) = \begin{cases} 2.0\, m & \text{if } \mathbf{I}_{\text{i, invWalls}}(x, y) = 0 \\ 0.0\, m & \text{else} \end{cases} \tag{8}
$$

We then construct the final map of a single room by combining both maps into the room map $\mathbf{I}_{\text{i}}$:

$$
\forall \mathbf{p}(x, y) \in \mathbf{I}_{\text{i}} : \quad \mathbf{I}_{\text{i}}(x, y) = \begin{cases} \mathbf{I}_{\text{i,Walls}}(x, y) & \text{if } \mathbf{I}_{\text{i,Walls}}(x, y) > \mathbf{I}_{\text{i,Furniture}}(x, y) \\ \mathbf{I}_{\text{i,Furniture}}(x, y) & \text{else} \end{cases} \tag{9}
$$

As we compute the equally sized maps $\mathbf{I}_{\text{A}}$ and $\mathbf{I}_{\text{B}}$ for both rooms A and B, we can then use these to find a map $\mathbf{I}_{\text{Floor}}$ of open floor space common to both rooms:

$$
\begin{aligned}
&\forall \mathbf{p}(x, y) \in \mathbf{I}_{\text{A}} \wedge \mathbf{I}_{\text{B}} : \\
&\qquad \mathbf{I}_{\text{OR}}(x, y) = \begin{cases} \mathbf{I}_{\text{A}}(x, y) & \text{if } \mathbf{I}_{\text{A}}(x, y) > \mathbf{I}_{\text{B}}(x, y) \\ \mathbf{I}_{\text{B}}(x, y) & \text{else} \end{cases} && (10)\\
&\qquad \mathbf{I}_{\text{Floor}}(x, y) = \begin{cases} 1 & \text{if } \mathbf{I}_{\text{OR}}(x, y) > 0.1\, m \\ 0 & \text{else} \end{cases} && (11)
\end{aligned}
$$

Thus we have a first 2D map of shared free space for both participants. Similary, we can use our room maps in order to identify obstacles unique to one room. We can compute $\mathbf{I}_{\text{Obst. A}}$ and $\mathbf{I}_{\text{Obst. B}}$ with boolean operators as follows:

$$\begin{aligned} &\forall \mathbf{p}(x,y) \in \mathbf{I}_{\text{A}} : \\ &\quad \mathbf{I}_{\text{AND}}(x,y) = \mathbf{I}_{\text{A}}(x,y) \wedge \mathbf{I}_{\text{B}}(x,y) && (12) \\ &\quad \mathbf{I}_{\text{Obst. A}}(x,y) = \neg\left(\mathbf{I}_{\text{AND}}(x,y)\right) \wedge \mathbf{I}_{\text{A}}(x,y) && (13) \\ &\quad \mathbf{I}_{\text{Obst. B}}(x,y) = \neg\left(\mathbf{I}_{\text{AND}}(x,y)\right) \wedge \mathbf{I}_{\text{B}}(x,y) && (14) \end{aligned}$$

Finally and maybe most importantly, we can identify not only obstacles common to both rooms in a map $\mathbf{I}_{\text{Common}}$, but also find consensus surfaces $\mathbf{I}_{\text{Surfaces}}$:

$$\begin{aligned} &\forall \mathbf{p}(x,y) \in \mathbf{I}_{\text{A}} : \\ &\mathbf{I}_{\text{HeightDiff}}(x,y) = \|\mathbf{I}_{\text{A}}(x,y) - \mathbf{I}_{\text{B}}(x,y)\| && (15) \\ &\mathbf{I}_{\text{Common}}(x,y) = \begin{cases} \mathbf{I}_{\text{AND}}(x,y) & \text{if } \mathbf{I}_{\text{AND}}(x,y) \leq 1.6\,m \\ 0 & \text{else} \end{cases} && (16) \\ &\mathbf{I}_{\text{Surfaces}}(x,y) = \begin{cases} \mathbf{I}_{\text{AND}}(x,y) & \text{if } \mathbf{I}_{\text{HeightDiff}}(x,y) \leq 0.05\,m \\ 0 & \text{else} \end{cases} && (17) \\ & && (18) \end{aligned}$$

We thus arrive at 2D maps of the 5 different types of spaces. These are also illustrated in Fig. 4.

- $\mathbf{I}_{\text{Floor}}$ shows the consensus free space available to both conversation partners.
- $\mathbf{I}_{\text{Obst. A}}$ shows obstacles present only in room A, but not in room B.
- $\mathbf{I}_{\text{Obst. B}}$ shows obstacles present only in room B, but not in room A.
- $\mathbf{I}_{\text{Common}}$ shows obstacles present in both rooms, but not of equal height.
- $\mathbf{I}_{\text{Surfaces}}$ shows obstacles present in both rooms and of equal height, e.g. table tops present in both rooms.

For initializing the consensus reality, we need to determine a reference point prior to map computation. In simple implementations, this might be done using a simple planar marker which serves as a reference point. More interesting however is the dynamic optimization of the map alignment transform $\mathbf{T}_{\text{align}}$. This can be formulated as an optimization problem with the goal of maximizing consensus free space and consensus surfaces. Using a suitable energy function E_{align}, we can solve the following equation for an optimal alignment:

$$\mathbf{T}_{\text{align}} = \text{argmin}_{\mathbf{T}}\, E_{\text{align}}(\mathbf{I}_{\text{A}}, \mathbf{I}_{\text{B}}, \mathbf{T}) \tag{19}$$

In the next step, we can use the resulting maps to communicate the similarities and discrepancies of their surroundings to the conversation partners.

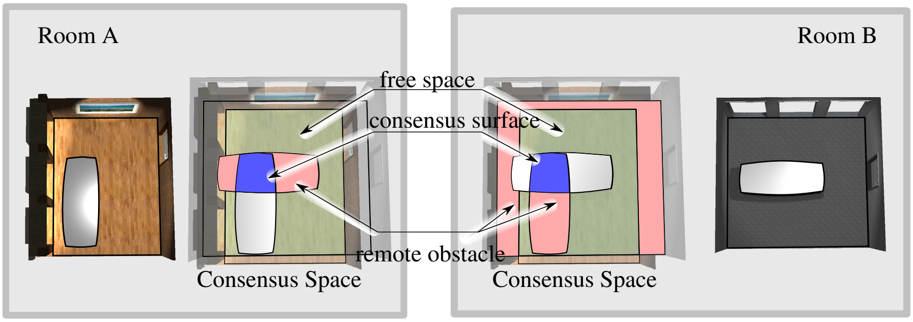


Fig. 4. Consensus Reality as seen from two different rooms. Note especially the different layout of remote obstacles (in red) for the two rooms.

4 Visualizing the Consensus Reality

The procedures described so far are based on geometrical analysis of meshed 3D models. The next challenge lies in communicating the extend and limits of the consensus reality to the conversation participants. In order to study this problem, we use a test-bench approach. The entire AR communication scenario is replicated in the virtual environment of a 4-wall CAVE VR system, an approach suggested in [26,27].

Hence we are free to quickly define and evaluate different modes of display, minutely control scenario parameters and reduce possible experimental noise. Furthermore, we are able to evaluate a wide range of different display technologies simply by simulation. We can evaluate not only existing devices, but also technologies still under development. This enables an effortless and cheap review of different approaches for future systems, ranging from simple Google Glass-style visors up to HMDs covering nearly the entire field of view.

After the mapping procedure described in the previous section, we are presented with the problem of visualizing these 2D maps in a 3D space on devices with a limited field of view (FOV). Here we have to consider different demands and constraints on the visualization. The goal is to provide a non-distracting, intuitive and clear rendering. We consider the different maps in turn, assuming the point of view of an participant in room A:

- $\mathbf{I}_{\text{Floor}}$: Rendering the free floor space can help confidently navigating the consensus space. However, care must be taken to make this visualization as non-intrusive and uncluttered as possible.
- $\mathbf{I}_{\text{Obst. A}}$: As we are already standing in room A and seeing the obstacles directly, no rendering of this map is necessary.
- $\mathbf{I}_{\text{Obst. B}}$: Rendering this map is important. This map contains information crucial to remote obstacle avoidance, such as walls or furniture in the remote location. This visualization helps in avoiding walking through the conversations partners wall, desks etc.
- $\mathbf{I}_{\text{Common}}$: As there is also a local obstacle present, rendering these common obstacles would only make sense if there is a marked difference in shape or height.

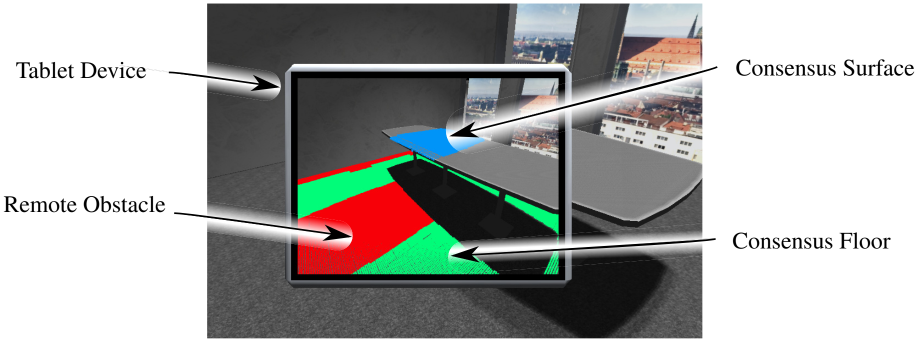


Fig. 5. Consensus Reality rendered as pointcloud on a simulated tablet device. Shown is the view from room A, as seen in Fig. 4. Different colors denote different maps, in this case $\mathbf{I}_{\text{Surfaces}} \equiv$ blue, $\mathbf{I}_{\text{RemoteObstacle}} \equiv$ red and $\mathbf{I}_{\text{Floor}} \equiv$ green.

- $\mathbf{I}_{\text{Surfaces}}$: Rendering this map enables the conversation partners to find flat surfaces common to both rooms. Thus, the local user in room A would be able to place a virtual object on the remote's user table in room B instead of having it hang mid-air.

We use a pointcloud rendering approach in order to visualize the different maps. Since the maps also contain height information, we can draw the point clouds such that they appear to hover over the real objects themselves.

The pointclouds $\mathcal{P}_{\text{i}} = \{\mathbf{A}_{i1}, \mathbf{A}_{i2}, \ldots, \mathbf{A}_{iN}\}$ are computed using the reverse mapping from the map array to the 3D space of the room:

$$\forall \mathbf{p}_j(x,y) \neq 0 \in \mathbf{I}_{\text{i}}:$$

$$\mathbf{A}_{j,\text{x}} = \frac{x-0.5\cdot s_x}{d_{\text{floor}}} \tag{20}$$

$$\mathbf{A}_{j,\text{y}} = \frac{y-0.5\cdot s_y}{d_{\text{floor}}} \tag{21}$$

$$\mathbf{A}_{j,\text{z}} = \mathbf{p}_j(x,y) \tag{22}$$

These pointclouds can then be rendered into the scene, as shown in Fig. 5, or used as the basis for more elaborate visualizations, e.g. using surface meshes or wireframe models.

5 Discussion and Conclusion

In this paper, we show a method for computing and visualizing consensus realities from existing 3D scans of rooms. These consensus realities are intended for use in 3D AR videoconferencing scenarios with non-homogeneous surroundings. The 2D mapping approach is a rather simplistic solution and easily implemented. The reduction to a 2D plane also reduces computational complexity especially for full room models: Whelan *et al.* [28] arrive at 1.2×10^6 vertices for a single room (LAB dataset). The simplification to the floor plane therefore significantly accelerates the consensus space computation,

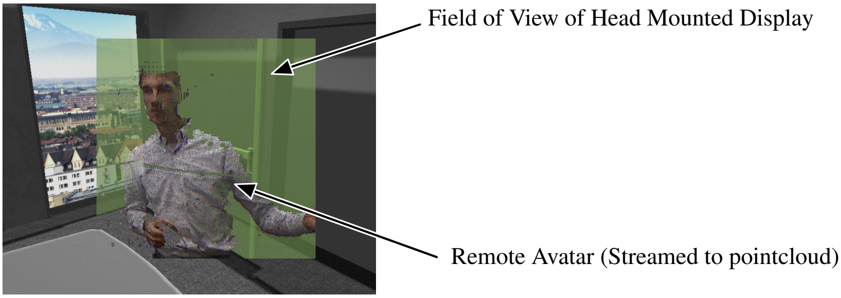

Field of View of Head Mounted Display
Remote Avatar (Streamed to pointcloud)

Fig. 6. Visualization of a dynamic remote avatar streamed to a pointcloud and seen through a simulated HMD device in our CAVE

however at the cost of losing height information during the projection. Our approach performs well for rooms of limited size with an even, continuous floor plane. In the implementation presented here, we assume a rigid mapping of participating conference rooms, i.e. no *redirected walking* [29,30].

The visualization with pointclouds in flat planes is a technically elegant solution, but suffers from the limited field of view of the users' devices. If mapped to the floor plane, the point cloud becomes barely visible at times and requires the user to look down in order to orient himself in the consensus reality. For remote obstacles, the rendering often results in pointclouds and artifacts floating in space without a visible connection to the physical surroundings. We therefore propose the introduction of more powerfull visualization techniques in future incarnations of our system. Nevertheless, pointclouds are a viable starting point for more elaborate displays and will remain attractive especially due to their flexibility in rendering and interaction. As we can see in Fig. 6, the pointcloud approach can also be used for rendering the users themselves. This is an active field of research in itself [1,2].

The use of a CAVE for fast evaluation of different rendering methods has proven to be very promising. We found that considering the same visualization techniques on different devices can lead to surprising insights which might have been missed otherwise. For instance, just varying the field of view for a see-through HMD can lead to a markedly different perception of the consensus space. In future works, we aim to extend the parallel testing and development of AR visualizations using the flexibility provided by the CAVE to a more formal and systematic development approach for AR experiences.

In the coming years, we are likely to see a convergence of three major technological developments: Affordable and lightweight consumer HMDs, as foreshadowed by the OculusRift and W. Steptoes current adaption to AR scenarios, cheap depth sensing cameras similar to the Kinect and the spread of high-bandwidth network connections. In consequence, we can expect AR videoconferencing to become a serious alternative to window-constrained videochats. As we move away from window analogies and virtual meeting spaces, the inclusion of our surroundings is bound to become an important factor in this development. We hope that the methods and insights outlined in this paper will help bringing videoconferencing away from limited flat screens into the rich and complex spaces of our everyday life.

References

1. Kammerl, J., Blodow, N., Rusu, R., Gedikli, S., Beetz, M., Steinbach, E.: Real-time compression of point cloud streams. In: ICRA, pp. 778–785 (2012)
2. Ruhnke, M., Bo, L., Fox, D., Burgard, W.: Compact rgbd surface models based on sparse coding. In: AAAI (2013)
3. Lehment, N.H., Erhardt, K., Rigoll, G.: Interface design for an inexpensive hands-free collaborative videoconferencing system. In: ISMAR, pp. 295–296 (2012)
4. Sodhi, R.S., Jones, B.R., Forsyth, D., Bailey, B.P., Maciocci, G.: Bethere: 3d mobile collaboration with spatial input. In: ACM SIGCHI, pp. 179–188 (2013)
5. Kim, K., Bolton, J., Girouard, A., Cooperstock, J., Vertegaal, R.: TeleHuman: Effects of 3D perspective on gaze and pose estimation with a life-size cylindrical telepresence pod. In: Human Factors in Computing Systems, pp. 2531–2540 (2012)
6. Adalgeirsson, S.O., Breazeal, C.: Mebot: a robotic platform for socially embodied presence. In: HRI, pp. 15–22 (2010)
7. Michaud, F., Boissy, P., Labonte, D., Corriveau, H., Grant, A., Lauria, M., Cloutier, R., Roux, M.A., Iannuzzi, D., Royer, M.P.: Telepresence robot for home care assistance. In: Multidisciplinary Collaboration for Socially Assistive Robotics, pp. 50–55 (2007)
8. Venolia, G., Tang, J., Cervantes, R., Bly, S., Robertson, G., Lee, B., Inkpen, K.: Embodied social proxy: mediating interpersonal connection in hub-and-satellite teams. In: SIGCHI, pp. 1049–1058 (2010)
9. Steptoe, W., Normand, J.M., Oyekoya, O., Pece, F., Giannopoulos, E., Tecchia, F., Steed, A., Weyrich, T., Kautz, J., Slater, M.: Acting rehearsal in collaborative multimodal mixed reality environments. Presence: Teleoperators and Virtual Environments 21(4), 406–422 (2012)
10. Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., Strehlke, K., Moere, A.V., Staadt, O.: Blue-c: A spatially immersive display and 3D video portal for telepresence. In: SIGGRAPH, pp. 819–827. ACM (2003)
11. Kurillo, G., Bajcsy, R.: 3D teleimmersion for collaboration and interaction of geographically distributed users. Virtual Reality 17(1), 29–43 (2013)
12. Maimone, A., Fuchs, H.: Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras. In: ISMAR, pp. 137–146 (2011)
13. Adcock, M., Anderson, S., Thomas, B.: Remotefusion: Real time depth camera fusion for remote collaboration on physical tasks. In: VRCAI, pp. 235–242 (2013)
14. Gurevich, P., Lanir, J., Cohen, B., Stone, R.: Teleadvisor: A versatile augmented reality tool for remote assistance. In: SIGCHI, pp. 619–622 (2012)
15. Oyekoya, O., Stone, R., Steptoe, W., Alkurdi, L., Klare, S., Peer, A., Weyrich, T., Cohen, B., Tecchia, F., Steed, A.: Supporting interoperability and presence awareness in collaborative mixed reality environments. In: VRST, pp. 165–174 (2013)
16. Billinghurst, M., Cheok, A., Prince, S., Kato, H.: Real world teleconferencing. IEEE Computer Graphics and Applications 22, 11–13 (2002)
17. Maimone, A., Yang, X., Dierk, N., State, A., Dou, M., Fuchs, H.: General-purpose telepresence with head-worn optical see-through displays and projector-based lighting. In: IEEE VR, pp. 23–26 (2013)
18. Salas-Moreno, R.F., Newcombe, R.A., Strasdat, H., Kelly, P.H., Davison, A.J.: Slam++: Simultaneous localisation and mapping at the level of objects. In: CVPR, pp. 1352–1359 (2013)
19. Newcombe, R.A., Davison, A.J., Izadi, S., Kohli, P., Hilliges, O., Shotton, J., Molyneaux, D., Hodges, S., Kim, D., Fitzgibbon, A.: Kinectfusion: Real-time dense surface mapping and tracking. In: ISMAR, pp. 127–136 (2011)

20. Granados, M., Hachenberger, P., Hert, S., Kettner, L., Mehlhorn, K., Seel, M.: Boolean operations on 3D selective nef complexes: Data structure, algorithms, and implementation. In: Di Battista, G., Zwick, U. (eds.) ESA 2003. LNCS, vol. 2832, pp. 654–666. Springer, Heidelberg (2003)
21. Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A.: Real-time human pose recognition in parts from a single depth image. In: CVPR (2011)
22. Baak, A., Müller, M., Bharaj, G., Seidel, H.P., Theobalt, C.: A data-driven approach for real-time full body pose reconstruction from a depth camera. In: CVPR, pp. 1092–1099 (2011)
23. Livingston, M., Gabbard, J., Swan, J., Edward, I., Sibley, C., Barrow, J.: Basic perception in head-worn augmented reality displays. In: Human Factors in Augmented Reality Environments, pp. 35–65. Springer, New York (2013)
24. Kantonen, T., Woodward, C., Katz, N.: Mixed reality in virtual world teleconferencing. In: IEEE VR, pp. 179–182 (2010)
25. Izadi, S., Newcombe, R.A., Kim, D., Hilliges, O., Molyneaux, D., Hodges, S., Kohli, P., Shotton, J., Davison, A.J., Fitzgibbon, A.: Kinectfusion: real-time dynamic 3d surface reconstruction and interaction. In: SIGGRAPH, pp. 23:1–23:1 (2011)
26. Ragan, E., Wilkes, C., Bowman, D., Hollerer, T.: Simulation of augmented reality systems in purely virtual environments. In: IEEE VR, pp. 287–288 (2009)
27. Lee, C., Bonebrake, S., Hollerer, T., Bowman, D.: A replication study testing the validity of ar simulation in vr for controlled experiments. In: ISMAR, pp. 203–204 (2009)
28. Whelan, T., McDonald, J., Kaess, M., Fallon, M., Johannsson, H., Leonard, J.J.: Kintinuous: Spatially extended kinectfusion. In: Workshop on RGB-D: Advanced Reasoning with Depth Cameras (2012)
29. Steinicke, F., Bruder, G., Jerald, J., Frenz, H., Lappe, M.: Analyses of human sensitivity to redirected walking. In: VRST, pp. 149–156 (2008)
30. Kulik, A., Kunert, A., Beck, S., Reichel, R., Blach, R., Zink, A., Froehlich, B.: C1x6: A stereoscopic six-user display for co-located collaboration in shared virtual environments. In: SIGGRAPH Asia, pp. 188:1–188:12 (2011)