AvatAR: An Immersive Analysis Environment for Human Motion Data Combining Interactive 3D Avatars and Trajectories

Patrick Reipschläger
p.reipschlaeger@googlemail.com
Autodesk Research
Toronto, Ontario, Canada
Technische Universität Dresden
Dresden, Germany

Frederik Brudy
frederik.brudy@autodesk.com
Autodesk Research
Toronto, Ontario, Canada

Raimund Dachselt
dachselt@acm.org
Technische Universität Dresden,
Centre for Tactile Internet (CeTi) and
Cluster of Excellence "Physics of Life"
Dresden, Germany

Justin Matejka
justin.matejka@autodesk.com
Autodesk Research
Toronto, Ontario, Canada

George Fitzmaurice
george.fitzmaurice@autodesk.com
Autodesk Research
Toronto, Ontario, Canada

Fraser Anderson
fraser.anderson@autodesk.com
Autodesk Research
Toronto, Ontario, Canada

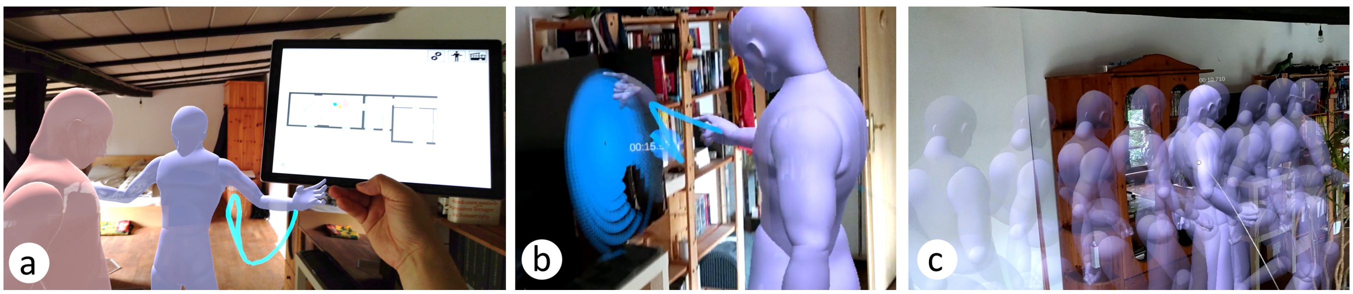

Figure 1: An overview of some of the aspects of *AvatAR*: (a) The virtual avatars of two persons, one with an active 3D trajectory for the left hand, and the tablet with the floor plan visible, which shows the trajectories and position of the two avatars as well as the position and viewing direction of the analyst. (b) A virtual avatar interacts with an interactive displays. Two ghost preview instances are pinned to mark the points when the display was touched and the gaze visualization is active to show what the avatar is looking at. (c) An avatar with the activated *Specter Visualization* technique for the whole body, which shows semi-transparent representations of the avatar for past and future time frames.

ABSTRACT

Analysis of human motion data can reveal valuable insights about the utilization of space and interaction of humans with their environment. To support this, we present *AvatAR*, an immersive analysis environment for the in-situ visualization of human motion data, that combines 3D trajectories with virtual avatars showing people's detailed movement and posture. Additionally, we describe how visualizations can be embedded directly into the environment, showing what a person looked at or what surfaces they touched, and how the avatar's body parts can be used to access and manipulate those visualizations. *AvatAR* combines an AR HMD with a tablet to provide both mid-air and touch interaction for system control, as well as an additional overview device to help users navigate the environment. We implemented a prototype and present several scenarios to show that *AvatAR* can enhance the analysis of human motion data by making data not only explorable, but experienceable.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **Visualization techniques**; *Ubiquitous and mobile computing systems and tools*; **Mixed / augmented reality**.

KEYWORDS

In-situ visualisation, augmented/mixed reality, human motion data, analysing space utilization, motion analysis, Immersive Analytics

ACM Reference Format:
Patrick Reipschläger, Frederik Brudy, Raimund Dachselt, Justin Matejka, George Fitzmaurice, and Fraser Anderson. 2022. AvatAR: An Immersive Analysis Environment for Human Motion Data Combining Interactive 3D Avatars and Trajectories. In *CHI Conference on Human Factors in Computing Systems (CHI '22), April 29-May 5, 2022, New Orleans, LA, USA*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3491102.3517676

CHI '22, April 29-May 5, 2022, New Orleans, LA, USA

ACM ISBN 978-1-4503-9157-3/22/04…$15.00
https://doi.org/10.1145/3491102.3517676

1 INTRODUCTION

The analysis of human motion data can reveal many useful insights into the utilization of space, existing behavior patterns, and the execution of complex work sequences. For example, human motion can be analyzed to better understand how people collaborate in front of large displays[61], what exhibitions visitors attend to and where they spend their time when visiting museums [38], to understand how buildings are utilized to make them more energy-efficient [7], and how people perform complex dance sequences [45]. Efficient analysis of movement data, however, requires tools that can directly relate human movement to the surrounding environment.

Traditionally, analysis of spatial data is performed primarily in desktop environments utilizing 2D [30, 50, 63] or 3D [2, 14, 36] trajectory plots, heatmaps [24, 30, 47], line charts [61, 63], and parallel coordinates [29]. In recent years *Immersive Analytics* has become a strong contender to more traditional approaches for the analysis of spatial data, including motion data, leveraging immersive technologies like Augmented Reality (AR) or Virtual Reality (VR) displays to visualize data directly within the environment it was captured in. These in-situ visualizations enable the direct, spatially aware analysis of motion data with a high degree of immersion [11, 34, 40]. Most of these systems still make use of predominantly 3D trajectories, sometimes accompanied by auxiliary visualizations such as heatmaps [11]. While trajectories alone allow users to gain an overview of the overall position of one or multiple people, they provide little details about their actual behavior and their interactions with the surrounding environment. We believe there is unaddressed potential in using immersive technologies to overcome barriers between people, their data, and the tools they use for analysing and understanding human activities and behavior.

To address limitations of previous approaches, we propose *AvatAR*, an immersive environment for the analysis of human motion data. *AvatAR* uses an Augmented Reality (AR) head-mounted display (HMD) to directly blend virtual content with the real environment and enable in-situ visualizations of human motion data. *AvatAR* combines traditional trajectory-based visualizations with interactive humanoid 3D avatars, providing a detailed representation of a person's posture as well as their movement through and interaction with the environment. In *AvatAR*, the 3D avatars act as a control interface allowing users to manipulate the visualization, control what data is shown, and to scrub through time. Additionally, a tablet interface allows for precise input as well as a concise overview of the current scene to aid the analyst's explorations. Additional visualizations, embedded directly into the environment, show what people looked at, their footprints on the ground, and where they interacted with the environment by touching it. These visualisations aim to create a strong link between data and environment, and to make data not only explorable, but also experienceable from the analyst's point of view. Such a tool could then be used for scenarios such as the analysis of movement patterns in an office context, learning of dance movements, and for analyzing customer behavior in retail and advertising (see section 5 for more details),

In summary, our contributions are:

(a) The design of an immersive analysis environment for the in-situ analysis of human motion data, utilizing interactive 3D avatars, motion trajectories, as well as visualizations embedded into the surrounding environment.
(b) Utilizing the body parts of an avatar as metaphors for accessing and manipulating visualizations directly using mid-air interaction, complemented by a more abstract tablet interface controlled by touch, which acts as an overview device as well.
(c) A prototype implementation of our concept as well as three scenarios highlighting how our techniques can be used for the analysis of human motion data.

2 RELATED WORK

AvatAR builds upon work on the visualization of motion data in general, on research using immersive technologies for visualizations, and on work combining mobile devices with immersive technologies.

2.1 Visualization of Movement Data

As shown by N. and G. Andrienko's analysis of the visual analytics of movement [4], as well as an earlier analysis of the visualization of spatio-temporial data in general [5], the majority of work for visualizing movement data utilizes some form of trajectory, with additional parameters like speed or density encoded in attributes such as color or thickness. The preferred perspective for those visualizations is the bird's-eye view [30, 50, 58], especially for map-based data [24, 44, 47, 56, 59, 63]. Visualising 3D motion data as trajectories in a 3D enviroment allows to also represent elevation as a third spatial dimension and to freely change the perspective when analyzing this data, either in desktop environments [2, 14, 33, 36], immersive environments [11, 17, 34] or on spatially aware mobile devices [13]. These trajectory-based visualisations can be accompanied by other types of visualizations, such as heatmaps [11, 24, 30, 47], line charts [24, 61, 63], or parallel coordinates [29] to show additional aspects of the data. Another example by Kepplinger et al. [33] use gaze information to highlight viewed objects to gain further insights about the behavior of players in a video game.

Space-time cubes [8] are another popular visualization technique for spatio-temporal data, where space is mapped onto two dimensions, often in the form of a map [3, 56, 59], and time is visualized on a third axis. Trajectories can be stacked to visualize additional attributes [3, 59, 66] and space-time cubes have been transferred to virtual environments [25] as well as immersive environments [51, 66]. Although space-time cubes are 3D visualizations in principle, only two spatial dimensions can be visualized, which makes this approach only suitable for data which is either inherently 2D, projected to 2D, or where the third dimension can be neglected. Thus, similar to the 2D techniques above, this approach is not viable for immersive in-situ visualization environments such as *AvatAR*.

Furthermore, there are examples of analysis environments which are less focused on movement itself, but more on how users interact with their environment, especially with interactive displays [10, 42, 61]. Börner and Penumarthy [17] used virtual avatars in a 3D environment to visualize the position of users, although only as approximations and with no detailed reconstruction of a user's posture. Lee et al. [39] visualize skeleton data extracted from surveillance videos for the analysis of human motion. However, the skeletons are superimposed over 2D images without the ability to freely explore the 3D environment and the representation of the humans is limited to stick figures, i.e., connecting the various joints with strokes. Ocupado [47] visualizes building occupancy on a floor plan view aggregated over time, but offers mainly heatmap based visualizations that do not provide any details on a person's posture or how they interacted with the environment.

While there is a vast body of research concerning the analysis of movement data, a lot of it is dedicated to visualizations that can only represent two spatial dimensions. While 3D environments are also common and have shown to effectively visualize human motion data in 3D space, these visualizations mostly use only a single point in time to represent a person's position. Thus, they do not allow the detailed reconstruction of human body postures. With *AvatAR* we want to address this by applying reconstructed skeleton data to humanoid avatars to facilitate more detailed analysis of human behavior. One approach that provides this kind of detailed humanoid avatars is the work by Yu et al. [65]. However, their research is focused on motion guides to teach users, e.g., dance routines and not on the analysis of motion data. Thus, their techniques do not allow to gain insights into another persons behavior within a certain environment, but rather provide insights on the users own movements.

2.2 Immersive Analysis and Environments

Immersive technologies like AR or VR HMDs are increasingly being used for supporting visual data analysis. The majority of *Immersive Analytics* [26] research is focused on placing classic visualizations, such as scatter plots or bar charts, in immersive environments [15, 28, 31, 46, 49, 51, 62, 66], providing interaction concepts to construct such visualizations [23, 46], fostering collaboration between users [16, 41], or analyzing how users utilize their surrounding space while using these systems [9]. For example, *ImAxis* [23] enables users to combine interactive axes to a variety of visualizations within an immersive environment. Rosenbaum et al. [49] proposed changing the color and size of data items in immersive visualizations such as scatter plots based on the proximity of the user. Zhang et al. [66] and Filho et al. [28] both utilize a VR HMD for visualizing movement data with a space-time cube. However, these solutions are limited in that only two spatial dimension can be represented. *GeoGate* [56] combines a space-time cube in Augmented Reality with trajectories on a map displayed on an interactive tabletop, and provides a ring-shaped tangible object that manipulates the data shown in the space-time cube. Nebeling et al. [46] present a generic Mixed Reality analysis toolkit which, however, lacks dedicated techniques for supporting spatio-temporal data and capabilities for in-situ visualizations.

In contrast to these examples, where the visualizations have little relation to their immediate surroundings, in-situ visualizations, such as those provided by *AvatAR* are placed and analyzed directly in the environment in which the data was recorded. Kloiber et al. [34] visualized human motion in virtual reality using 3D trajectories of the hands and head of multiple people. A triangle fitted to those position serves as a guide for simple pose estimation. The trajectories are visualized in the same virtual environment they were captured to analyze the user's movements. Proxy objects of users' hands and headset offered further insights on what users were looking at or what they were interacting with. Similarly, Lilija et al. [40] and *GhostAR* [18] show proxy objects for hand and head movement, combined with 3D trajectories in a virtual environment. The *MIRIA* toolkit [11] enables the in-situ analysis of movement and interaction data by rendering 3D trajectories using an AR HMD. Additional visualizations such as heatmaps can be placed on the walls or floor to provide further insights. It supports proxy objects that visualize the tracked devices and their orientation overlaid on the trajectory, but provides no detailed visualization of a person's posture. All these works target similar problems as *AvatAR*, but they do not offer the detailed analysis of a person's posture, being limited to either a single position per user [11] or the positions of head and both hands [18, 34, 40] for each time frame. In contrast, *AvatAR* aims to improve upon these approaches by providing a much more detailed reconstruction of a users posture to facilitate the in-depth analysis of their behavior. In contrast to this, the work of Kraus et al. [35] provides detailed 3D models of people and a 3D environment reconstructed from camera images, which can be explored in VR. However, being a VR environment, it lacks the in-situ aspect *AvatAR* is aiming for and relies instead on teleportation for exploration of the environment.

2.3 Combing AR HMDs and Mobile Devices

Mobile devices have been combined with AR displays to support interaction with the immersive content. Early examples from Szalavári et al. [57] and Schmalstieg et al. [52] both utilize handheld sheets that can be used to manipulate AR objects in various ways. More recent examples include *Symbiosis Sketch* [6] which combines a tablet with an AR HMD for improved accuracy during 3D sketching. Users can define a surface in 3D space and then draw on this surface using the tablet. Sereno et al. [53] utilize a tablet for manipulating an AR 3D visualization in a collaborative setting through multi-touch interaction. *MARVIS* [37] shows visualizations in AR that support the data visualization displayed on a tablet's screen. Büschel et al. [12] investigate how a smartphone can be used for panning and zooming an immersive 3D visualization. Hubenschmid et al. [31] utilize spatially aware tablets for multi modal interaction with immersive 3D visualizations. In addition to mobile devices, prior research highlights the benefits of providing touch input for immersive technologies by combining them with large displays [48], interactive tabletops [16], and multi display environments [21, 22]. These examples illustrate the benefits and potential of combining mobile devices with immersive displays, especially to facilitate exploring and interacting with 3D visualisations, and how touch interaction can be used as an alternative to mid-air interaction for manipulating visualizations through precise and tactile input. *AvatAR* leverages the insights from these works by integrating a dedicated tablet interface into the immersive environment.

3 *AvatAR*: OVERVIEW

AvatAR is a tool to analyze human motion data to facilitate understanding of people's movements through and interactions with their surrounding environment. *AvatAR* provides an immersive

Figure 2: The core components of *AvatAR*: The virtual humanoid avatar visualizing a person's detailed posture, a 3D trajectory providing an overview of movement, and the accompanying handheld tablet offering precise control of all visualizations.

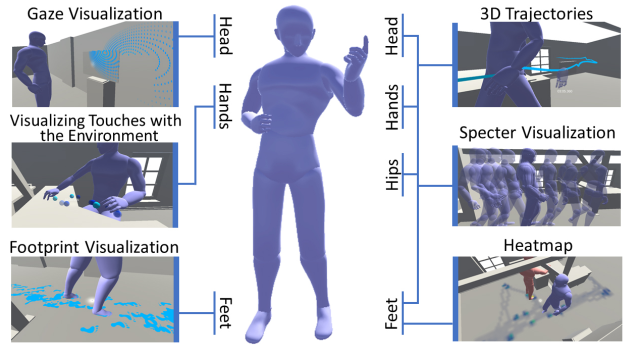

Figure 3: Overview of *AvatAR's* various visualization techniques and the body parts of the virtual avatar from which they can be accessed.

analysis environment in which time-varying motion data is visualized directly in the environment in which it was recorded or to which it relates. Specifically, an analyst wears an Augmented Reality (AR) head-mounted display (HMD) to move freely through the environment and the visualizations of spatially registered data, exploring the data and interacting with the visualizations using mid-air gestures. The data visualization is centered around humanoid avatars, which give a detailed impression of the current posture of the recorded person (Figure 2). The avatars are supported by additional visualizations, such as 3D trajectories that provide an overview of past and future positions of a person, heatmaps and footprint visualizations on the ground that give an overview of the density as well as exact movement of one or multiple people, a visualization of a person's current visual focus, and a visualization of where people touched or interacted with the environment (Figure 3). All of these visualizations form a system of multiple coordinated views in the sense that they show different representations of the same data. They are designed to work in conjunction with each other to provide analysts with an understanding of the current scene, with the avatar as the central element not only for showing the data, but also for interacting with and controlling the visualization through mid-air gestures.

In addition to the HMD, the analyst carries a tablet that provides a 2D floor plan showing the current position and trajectory of all the avatars as well as the analyst's location within the space (Figure 1a). This complements the immersive view of the AR HMD with a summary and can help the analyst with navigation and orientation. The tablet also provides an alternative form of interaction, allowing for direct interaction with the scene and visualizations using precise touch input through a control interface (Figure 2).

In the remainder of this section, we will describe the individual techniques in detail, and how an analyst interacts with them. Interaction with the avatars and visualization techniques is a key component of *AvatAR* for exploring the dataset and gaining new insights. Therefore, we relate our different techniques to the seven categories of interaction identified by Yi et al. [64] and to the user intent to which they translate: **Select** (mark something as interesting), **Explore** (show me something else), **Reconfigure** (show me a different arrangement), **Encode** (show me a different representation), **Abstract/Elaborate** (show me more or less detail), **Filter** (show me something conditionally), and **Connect** (show me related items). While these categories have been developed for traditional information visualization environments, we transfer them to our immersive *AvatAR* environment.

3.1 Avatars

Avatars are virtual humanoid representations of a person at a single point in time (Figure 4a), showing in detail how a person was positioned within a space at a given time. They are the central visualization and interaction component within *AvatAR*, and are constructed from pre-recorded 3D joint data.

The HMD affords freedom of movement so that analysts can freely explore the posture of an avatar and their relation to the environment in AR. By moving to the same position as an avatar in a scene, analysts can better understand the avatar's current situation, for example to learn about comfortable reach from that positions or which objects and features in the environment might be visible to the recorded person. This form of embodiment allows an analyst to use their own senses to experience the data in relation to the environment, which is not possible using more traditional analysis tools.

3.1.1 Visual Representation of Avatars. Avatars are generated by applying pre-recorded skeleton data to a humanoid 3D model in AR, thus reconstructing a person's posture at a specific point in time. This skeleton data can originate from a multitude of sources, such as OpenPose [20], simple RGB cameras [43, 60], or even generated from virtual environments [19, 27]. Avatars are scaled to match the real-world height and physical size of the person they represent, as this might affect how they interact with their physical environment or what they can reach and see. Otherwise, avatars provide few defining features and have the same optical appearance for every

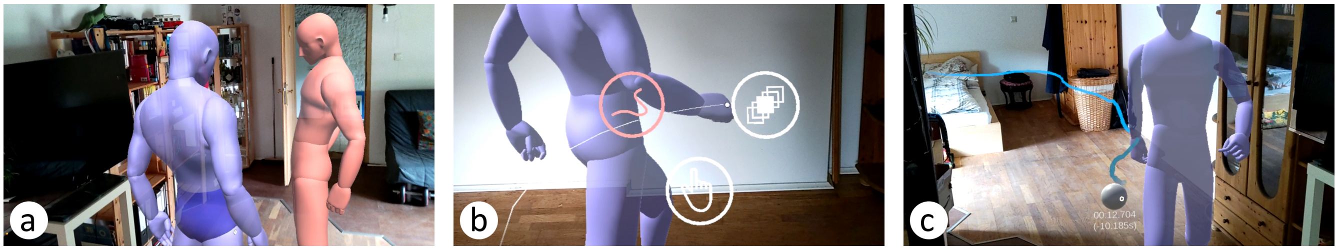

Figure 4: (a) One virtual 3D avatar interacting with the environment with another one walking in the background. (b) Marking menu for the avatar's hand with the item for trajectories being selected. (c) Trajectory visualizing past and future time frames with the avatar at the center. A handle for adjusting the trajectory's length is shown as well.

person, except for being color-coded to help distinguish the avatars of different people from one another (Figure 1a).

3.1.2 Manipulating Time. The torso of an avatar acts as a timeline scrubber to directly manipulate the current time of the visualization. This corresponds to the **Explore** interaction category of Yi et al. [64] as it allows analysts to navigate through different points in time to find interesting aspects within the data. To interact with the torso, an analyst points at it and performs a mid-air tap-and-hold gesture, then moves their hand left or right to scrub through time, updating the avatar's visualizations accordingly. Performing a single air-tap gesture quickly without holding starts and stops the playback mode, which advances the dataset in real time. This technique is one of the tools to navigate through the data, with using the time control interface on the tablet being the other one.

3.1.3 Accessing Additional Visualizations. Most visualizations in *AvatAR* are associated with a particular body part of the virtual avatars, from which they can also be activated and controlled. To interact with a specific body part, an analyst points at it with their finger to highlight it. Through a mid-air tap-and-hold gestured they then select the body part and a menu around it shows icons for the available visualizations. Moving the hand towards an icon selects it (Figure 4b).

We use the hip, torso, head, hands, and feet as interactive elements of the avatar (see Figure 3 for an overview). The hips of the avatar represent the overall position of a person and provide access to its corresponding global 3D trajectory. As the head determines where and what a person is currently looking at, we use it to access a visualization of a person's gaze direction. The hands are the primary means of interaction and manipulation in an environment, so they are mapped to a visualization showing where a person interacted with the environment by touching it. The feet are used to access a heatmap of a person's location, as well as a visualization of the footprints a person leaves while moving. 3D trajectories are also available for individual body parts (head, hands, feet) and will only show a visualization of that particular body part to provide more fine grained insights.

These interaction techniques allow an analyst to access the wide range of different visualizations (described in detail in the following sections) through easy to recall menus at the body part of interest. It also enables analysts to quickly toggle visualizations they want to use together in combination for their current task. By tying the visualisations to the different body parts of the avatars, *AvatAR* allows the analysts to focus on the area that is currently of interest to them, such as what interaction with the environment a person has engaged in, or where their visual attention was directed towards. This corresponds to the **Encode** interaction category of Yi et al. [64], because it allows analysts to access different representations of the dataset and choose the one suitable for their current goal. Furthermore, this is also an example of the **Connect** interaction category, as all the techniques accessible from the avatar's body are directly related to the current time frame the avatar shows, and they all change in unison whenever the time frame is altered by the analyst.

3.2 Trajectories

3D trajectories add additional information to the immersive analysis environment. Each trajectory directly corresponds to a particular avatar. Trajectories provide an abstract representation of the location of a person over a period of time. They can be used to gain an overview of the overall movement within an environment. Trajectories are represented as continuous tubes (Figure 4c) painted in the same color as its corresponding avatar. When an analyst looks at either of the endpoints of the trajectory, a handle appears. Analysts can control the length of time visualized by dragging handles through mid-air gestures, thereby reducing or increasing the number of frames shown before and after the avatar's current playback time. This is an example of the **Filter** interaction category of Yi et al. [64], allowing analysts to select the amount of data they want to see. By performing a double-air-tap at the endpoints, the length is set to infinity, thereby visualizing the full length of the dataset. In combination with the avatars, trajectories correspond to the **Abstract/Elaborate** interaction category of Yi et al. [64], because trajectories show an abstract representation of multiple data points, while an avatar shows an elaborated visualization of a single data point. *AvatAR* also provides access to individual trajectories for the hands, feet, and head of a person, allowing for a more focused reference when analyzing data. This in turn is associated with the **Encode** interaction category, because it allows analysts to choose between different representations, as well as the **Abstract/Elaborate** category, since analysts can either focus on the overall position of an avatar or the location of a specific body part. All trajectories are associated with their corresponding avatar and change together with it when the current time is altered.

Figure 5: (a) *Ghost Preview* technique with two pinned hand ghosts touching the display. (b) *Specter Visualizations* technique of just the avatar's hands, showing past and future time frames as semi-transparent instances of the hands. (c) *Gaze Visualization* for two persons, one looking at the display and one looking at the other person.

3.3 Further Integrating Avatars and Trajectories

The immersive environment of *AvatAR* can show both 3D trajectories and humanoid avatars in augmented reality at the same time. We further combine the high degree of details provided by the avatar with the summarizing overview of a trajectory through gaze-controlled ghost previews and specter visualizations.

3.3.1 Gaze-Controlled Ghost Previews. The *Ghost Preview* technique uses the gaze direction of an analyst to show a semi-transparent preview of the avatar at that position along the trajectory. The resulting *Ghost Preview* (Figure 5a) is shown more transparent than the actual avatar. By looking along the trajectory, an analyst can get an understanding of a person's behavior over time without the need to explicitly interact with the avatar or watch the playback in real time. Similar to trajectories that can be used to visualize the motion of the entire body, as well as of individual body parts, the *Ghost Preview* technique is available for each activated trajectory. For example, trajectories of the hands provide an understanding of where a person's hand has been and will go. An analyst can follow along the detailed movements of the hands as *Ghost Previews.* This focused view allows the analyst to understand how a person interacted with the environment without obstructing their view by showing the whole body of the avatar. This is a details-on-demand technique, that corresponds to the **Abstract/Elaborate** interaction category of Yi et al. [64], since it provides details and elaborates on a particular point on the abstract trajectory view. A *Ghost Preview* can be pinned by performing a double air-tap (Figure 5a). This fixes them in place to be used for comparison with other points in time. This corresponds to the **Select** interaction category, allowing analysts to mark interesting points in the data for later comparison. Another double tap on a pinned item removes it again.

3.3.2 Specter Visualization. While the *Ghost Preview* techniques enhances the summative trajectory with the details of the avatar, the *Specter Visualization* follows the opposite strategy of enhancing the avatar visualizations by spatially multiplexing them along the temporal trajectory. To do this, we show the avatar not only for a single point in time, but for past and future time frames as semi-transparent representations in set intervals (Figure 1c). When a person remains stationary, those representations are superimposed on each other, resulting in a more opaque visualization. However, if a person moves, only a semi-opaque specter trail remains visible. This technique provides an accurate representation of a person's posture over time as well as a better sense of the motion of a person than just a single avatar would have. This technique corresponds to the **Abstract/Elaborate** interaction category of Yi et al. [64], since the ordinary trajectories provide an abstract representation while the *Specter Visualization* can provide more insights by offering a detailed view of the same data.

Specter Visualization can be difficult to interpret if too many of the avatar representations overlap. Therefore, the density of the visualization can be adjusted by dragging a handle on the *Specter Visualization* up or down using a mid-air gesture. The duration of frames shown in the past or future can be controlled by dragging the handle inward or outward. This corresponds to the **Filter** interaction category of Yi et al. [64], allowing analysts to increase or reduce the visual complexity of the visualisation by filtering out certain data. These adjustments can be used to prevent the visualization from becoming cluttered, confusing or difficult to read. It can also be further enhanced by using different colors for past and future frames to make it easier to identify the direction of the movement. Similarly to the *Ghost Preview* technique, the *Specter Visualization* technique can be applied to either the whole body of an avatar or just individual limbs like the hands or head, by selecting the corresponding option from the menu described in section 3.1. For example, Figure 5b shows a *Specter Visualization* for only the hands of an avatar. This enables an analyst to focus their analysis on a specific body part while further reducing the complexity of the visualization.

3.4 Environment Visualizations

Visualizations can be embedded directly into the environment to strengthen the mapping and connection between real and virtual objects.

3.4.1 Gaze Visualization. When analyzing human motion in an environment, people's visual focus can provide an analyst with insights into the user's behaviour as well as what they could perceive from their position and whether they had an obstructed view. *AvatAR's Gaze Visualizations* highlights where a person's gaze intersected with the environment (sources of gaze data are discussed in section 6.6.1). To achieve this, a ray is cast in a radial pattern from the middle between the eyes of the avatar into the environment, limited to ten degrees of a person's field of view, to represent their visual focus area (Figures 1b and 5c). When a ray hits a surface

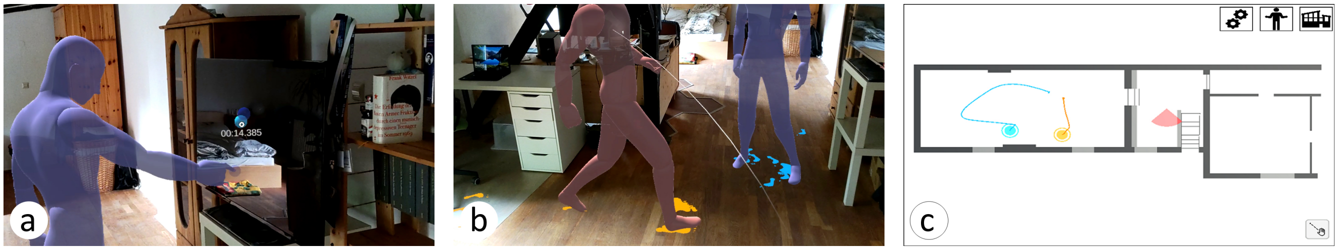

Figure 6: (a) Visualization of interactions with the environment, here showing touches on an interactive display with the center one being highlighted. (b) Visualization of where the feet of two avatars are touching the floor of the environment. (c) Floorplan view on the tablet, showing an overview of the environment, the current position and orientation of the analyst, as well as trajectories and the current position of the avatars.

of the environment, a sphere is created at this position, colored the same as their associated avatars. These spheres can help an analyst to determine whether certain objects in the environment where visisble to a person or if they were blocked from their view, e.g., being obstructed by a piece of furniture. Whenever an avatar's head moves, the visualization is updated to reflect the new position of the avatar's eyes. Overall, this technique is an example of **Abstract/Elaborate**, with the avatars head position and orientation representing the abstract view and the *Gaze Visualization* offering an elaborate view of the data.

3.4.2 Visualizing Touches Within the Environment. *AvatAR* allows an analysts to understand where a person directly interacted with their environment through touch, such as light switches or door handles. While in MIRIA [11] touch points on an interactive surface are visulialized by superimposing a 2D point plot onto an interactive display in AR, we extend this concept with *AvatAR* by showing touch interactions with all surfaces in the environment as well as between different avatars. These touch interactions are embedded directly into the environment at the point of contact, represented as small spheres that are color coded to match the corresponding avatars. Looking at a sphere shows the timestamp the touch occured at. Performing an air-tap moves the current time to this timestamp, allowing analysts to navigate between different events. Spheres prior to the current time frame are shown in a darker shade than future ones. Touch interactions are determined by calculating the distance from the fingers of the avatar to the objects in the environment. Whenever this distance drops below a defined threshold, physical contact with the corresponding surface is assumed, which continues until the threshold is exceeded again. In other words, *AvatAR* aggregates all touch points that contribute to a touch interaction into a single sphere, positioned at the location of the first touch point. The touch visualization technique is related to the **Explore** interaction category of Yi et al. [64] because it allows analyst to directly jump to potentially interesting points within the dataset. It is also an example of the **Connect** interaction category because jumping to a particular point sets the avatar and all active visualizations to the same point in time as well.

3.4.3 Heatmap and Footprint Visualization. To determine in which areas of the environment a lot of movement occurred and which areas people avoided, heatmap visualization are commonly used [11, 24, 30]. *AvatAR* embeds heatmaps directly on the floor of the environment. They can be toggled on or off for each avatar individually by selecting the corresponding icon in the feet's radial menu. If multiple avatars' heatmaps are active, they are combined into a single heatmap visualization.

While heatmaps are suitable for visualizing aggregated movement data, they don't differentiate between the movement of individual people or show the direction of the movement. Therefore, *AvatAR* provides an additional visualization of the footsteps of individual people to show the exact position where a person's feet touched the ground as well as the direction of movement. Whenever a foot of a person touches the ground, a bare footprint marks that position (Figure 6b). This *Footprint Visualization* not only provides an overview of where a person moved, but how fast they moved through the environment. An example would be a person climbing a set of stairs by taking only every second step, which could indicate the person was in a hurry. Another example would be when people walked sideways through a tight space. Both of these would be difficult to determine using a conventional heatmap. In combination, heatmap and *Footprint Visualization* relate to the **Abstract/Elaborate** interaction category of Yi et al. [64], with the first providing an abstract overview of the general movement density and the second providing detailed information of the actual steps a person took.

3.5 Mobile Device Views

While the mid-air input provided by the AR HMD allows analysts to directly interact with AR objects, it often lacks precision for, e.g., the detailed selection of time, and it is difficult to interact with objects from afar. The immersive AR view provided by the HMD is therefore complemented by a tablet (Figures 1a, 6c, 2), which serves as an additional input and output device. The haptic feedback and direct touch input of the tablet serves to mitigate the aforementioned issues. The tablet provides controls for all of the avatars' visualisations, that can be used as an alternative to the radial menus attached to an avatar's different body parts and thus allows manipulating avatars which are either far away or completely or partially obstructed by the environment. The tablet also provides a floor plan, which shows a top-down view of the environment and the avatars in it. Thus, it helps analysts to orientate themselves, navigate within the environment, and to locate avatars even if they are not visible from their current position. Both, the avatar control view and the floor plan, can be switched by pressing the corresponding button on the top of the tablet's screen.

3.5.1 Floor Plan View. The Floor Plan View provides an overview of the surrounding environment to the analyst to help with navigation and orientation. It shows a 2D floor plan of the current building that can be interacted with using the common drag-to-pan and pinch-to-zoom touch interaction techniques (see Figure 1a and 6c). The current position and orientation of the analyst within the environment is shown by a view cone that is updated dynamically as the analyst moves or faces different directions. 2D trajectories, superimposed on the floor plan provide a top down visualization of the time-varying position of each avatar within the data set. In contrast to the 3D trajectories of the AR view, these 2D trajectories show the entirety of the motion instead of a specific time window, to allow the analyst getting and overview over the complete dataset. This in turn relates to **Abstract/Elaborate** [64], with the tablet providing an abstract 2D overview of environment and trajectories and the AR view providing a detail representation of the avatars and trajecotries directly within the environment. All trajectories are color-coded the same way as their corresponding AR counterparts. A circle on the 2D trajectories visualizes the current position of each avatar on the floor plan as a general overview and to make it easier for analysts for find them in AR. Analysts can enter a selection mode by pressing a corresponding button in the corner of the floor plan, to perform a lasso-selection technique using a touch drag gesture. This selects all 2D trajectories crossing the area as well as their corresponding avatars. All other avatars and their associated visualizations are hidden in AR as well as on the floor plan. This is a **Filter** interaction [64]. Performing a double tap anywhere on the floor plan discards the selection and shows all avatars and trajectories again.

3.5.2 Avatar Control View. The Avatar Control View offers a detailed overview of all the avatars in the scene (see Figure 2). It consists of a global time control and dedicated time controls for each avatar in the dataset, containing the avatar's ID, a time slider to manipulate time continuously, buttons for advancing or reverting time in single frames, and a button to toggle play mode where time advances continuously. By default, the global time control slider manipulates time for all avatars in unison, and the time sliders for each avatar are locked. A check box can be toggled to deactivate time synchronization for the avatars, which locks the global time slider and enables to manipulate time independently for each avatar using their dedicated sliders. The time sliders of avatars are colored to match their virtual 3D avatar counterparts to help differentiate them. Different shades on the slider indicate when a particular person was present in the scene, i.e., when there is actual data on a person that can be visualized. This allows an analyst to quickly get an overview over where in the data each avatar is present. Furthermore, by taping the ID field, an avatar can be disabled to filter the dataset. Underneath the time slider a list of buttons can be used to control each avatar's visualizations (see section 3.1), providing an analyst with an alternative to using the radial menus in AR. These buttons can also be found at the global time control for toggling visualizations for all avatars together. This further supports the tablet's role as an overview device by allowing an analyst, e.g., to quickly enable the heatmap view for all avatars to gain an understanding of the overall movement in a particular area.

4 PROTOTYPE

To explore the technical feasibility of our visualization techniques in a real environment, we developed a prototype implementation of *AvatAR*. It was implemented for Microsoft HoloLens 2 and a Microsoft Surface Pro 7 as accompanying tablet. The applications were developed in the Unity 3D engine[1] with the Mixed Reality Toolkit[2]. Both devices communicate with a central server using TCP and a custom protocol to synchronize state. Our *AvatAR* prototype provides a visualization environment for human motion data. It should be noted that it does not offer a complete processing pipeline that would also include capturing, processing, and refinement of these data.

4.1 Data Sets

As sample data sources for our visualizations we used the CMU Panoptic Studio datset [32, 55], JTA dataset [27], as well as a custom data set for one of the co-author's home-office environment consisting of three people interacting in the space (see the accompanying video figure for the last example). Due to the ongoing COVID–19 pandemic we were not able to newly record in-situ data in an actual office or shop.

4.2 Spatial Occlusion and Environment Model

Although *AvatAR* uses HoloLens' spatial mapping for tracking and spatial orientation, a custom-built 3D model is used to determine the spatial occlusion of virtual objects by real objects. The real environment was measured in detail for this purpose and recreated in the 3D modeling program 3ds Max. Integrated into our Unity prototype, the resulting 3D model is manually aligned to match the real space geometry at the first start of the program. Occlusion of virtual objects was achieved by using a custom shader for the environment model that only writes to the depth buffer but not into the frame buffer. The result is a much more accurate spatial occlusion model than the one created by the HoloLens' built-in capabilities. This enhances immersion and allows to perform raycasts necessary for our environment visualizations using Unity's physics engine and the 3D model. Consequently, the gaze visualization's visual imprints closely match real world geometry, increasing its representational accuracy. The floor plan used in the tablet view is also based on the custom 3D model of the environment, which was reduced to two dimensions in 3ds Max before being imported into Unity.

4.3 Preparing *AvatAR* for a Specific Use Case

In principle, *AvatAR* can be configured to support any environment. The necessary steps are the selection of the data source, a suitable environmental model, and the corresponding floor plan. Selecting the data set is the most important step. *AvatAR* uses an internal format for representing the collective joint data, but offers converters for a variety of formats, such as the standard OpenPose joint format [20]. Once a data set is selected and parsed, the avatars are

[1]https://unity.com/
[2]https://github.com/microsoft/MixedRealityToolkit-Unity

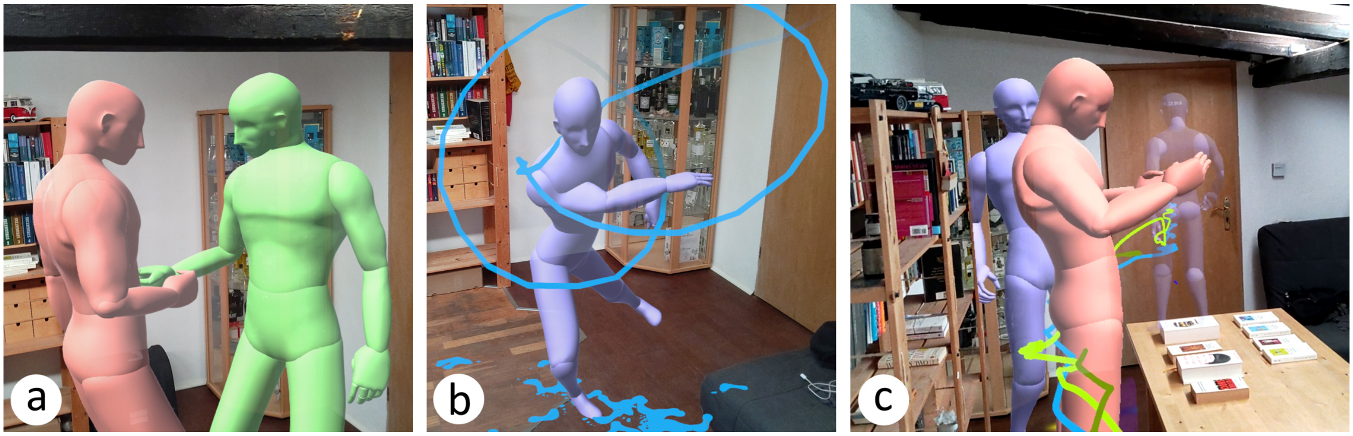

Figure 7: (a) The hands of avatars connecting as people shake hands to greet each other. (b) A person performing a dance choreography with their hand movements visible as a trajectory and the positions where their feet touch the floor as foot steps. (c) People having to move through a narrow passage to get past each other in a retail store scenario.

placed into the environment and the scene can be explored using our previously described techniques. An important consideration is the origin of the coordinate system in which the joint positions are represented and how it relates to the real environment. *AvatAR* in its current implementation does not have automatic alignment as this would require extensive knowledge of the target environment. Instead, a gesture-based interface allows a user to manually align the tracking data with the real environment at the first start of the program.

The next step is to provide a 3D model of the environment which is used for the spatial occlusion of virtual objects by the real environment. This can either be created manually, as in the current prototype, or for example be extracted from a BIM system or be created by a LIDAR scan. Alternatively, the built-in functionality of the HoloLens can also be used for this purpose, although it offers lower precision. The last step is to provide a floor plan to show on the tablet for orientation. In our current prototype, this was created manually based on the 3D model, but can be extracted from an existing BIM system or provided by other sources. *AvatAR* assumes that the 3D model and the floor plan have the same coordinate origin as the data set to ensure the data is aligned and displayed correctly.

After these steps have been performed, *AvatAR* is adapted to the new environment and ready for in-situ analysis of the data. Since *AvatAR* is a research prototype and not a finished product, currently those changes must be performed in the Unity 3D engine and no dedicated interface is provided within the prototype.

5 USE CASE SCENARIOS

In this section we want to demonstrate the transferability of our concepts to different application domains by presenting three use-case scenarios: 1) analysing movement patterns in an office context, 2) learning of dance movements, and 3) retail and advertising. Through these scenarios we illustrate how the individual techniques can be used to gain insights into spatial tracking data, as well as how workflows combine the techniques, showcasing *AvatAR's* depth and breadth.

This paper is accompanied by a video figure demonstrating the overall concept of *AvatAR*, as well as three short video clips as supplementary material showcasing these three use case scenarios.

5.1 Scenario 1: Analysing movement patterns in an office context

Leah works as a facility manager in a small company. After working entirely from home for a period of time due to COVID–19, employees are anticipated to come back into the office more frequently. Leah wants to use this opportunity to optimize the work space layout and identify places that are frequented by many people and therefore might pose issues regarding workplace hygiene. A pre-pandemic work session of the employees was captured with a camera-based tracking system set up in the office, and Leah loads this data set into *AvatAR* for analysis.

Leah starts her analysis by looking at the trajectories in the floor plan view of the tablet to get a general overview of where people were moving. She immediately notices a pattern where a person is moving back and forth in their office. After taking a look of the corresponding heatmap visualization an the 3D trajectories, Leah pins several Ghost Previews of the person. She notices, that the person is most likely making a phone call while restlessly wandering around the office. Although Leah views this behavior with curiosity, it does not pose a health risk, and she continues to explore the data set further.

Leah looks again at the floor plan and notices that two people's trajectories are moving towards and then away from each other. Leah sets the current time of the visualization to just before the encounter and starts the playback. She observes how the hands of the avatars connect as the two people shake hands to greet each other before heading to their respective desks (see Fig. 7a). Leah makes a note to remind the employees to keep their distance from each other and avoid physical contact as much as possible.

She notices several people moving and staying for a longer time in one of the larger offices. Leah also goes there and finds a digital whiteboard at the corresponding location. When manipulating time and watching the avatars, Leah notices people are interacting with the whiteboard and are having lively discussions with each other. She realizes that this is a popular place for having ad-hoc meetings and also notices that people pay little attention to the distance they keep from each other. Leah makes a mental note to enact a strict mask requirement at such meetings in the future as they pose a high risk setting.

Leah activates the *Touch Visualization* for several of the people present at the office to investigate where they had direct contact with the facilities. She detects many touches at a table in one of the larger offices and moving over there, notices a French press and other coffee-making equipment. Activating several touch points, Leah identifies at least three different people interacting with the equipment for a longer time. Leah makes a mental note to replace the French press with an automatic coffee maker, which is easier to disinfect and therefore easier to work into her hygiene concept.

5.2 Scenario 2: Remote Analysis of a Dance Choreography

Walter decides to take up his old hobby of dancing again after years of hiatus. However, due to an ongoing pandemic, Walter is unable to personally attend any schools. Looking for a solution online, he finds a school which offers to digitize his dance moves, which will be then analyzed by one of the instructors. Working through a tutorial, Walter sets up several cameras to capture his performance as skeleton data and sends it to Eliza, his assigned instructor, to provide feedback for his performance. Since for this scenario, the environment is not of importance and the recording can be played back anywhere, Eliza uses one of the rooms of the dancing school to watch a playback of the recording on her HMD. She can walk around the avatar to fully understand Walter's body posture and movement. In addition, she can pause the playback at any time to inspect the current posture in great detail (see Fig. 7b). Analyzing the *Footstep Visualization* on the floor, Eliza is able to determine if Walter placed his feet correctly after landing a jump and if his stride is big enough. When unsure, Eliza can also shadow Walter's movements and use her own experience to determine if the moves have been performed correctly. As a result, Eliza is able to provide Walter with detailed feedback on how to improve his dancing. She even provides Walter with a recording of her own performance of the same routine, that Walter can replay using *AvatAR* on his own to learn from Eliza's example.

In a similar way, *AvatAR* could be used for physiotherapy: a therapist can follow the patient's recorded movement to understand how well the patient is progressing without the need for daily in-practice appointments.

5.3 Scenario 3: Understanding and Improving Customer Experience

Sam wants to use *AvatAR* to understand customers' experience for a clothing store and optimize its layout. Using trajectories, Sam analyzes how customers arrive and move through the store and finds several bottlenecks, where a lot of the trajectories converge. Enclosing one such area on the tablet, Sam filters the dataset to only those people that moved through this area. Turning on the replay, it becomes now clear customers have to take turns in passing between the shelves or need to squeeze past each other (see Fig. 7c). Customers further seem to quickly vacate the area to move out of each other's way, which Sam confirms using the heatmaps. By comparing these observations against the store's sales records later, Sam discovers unusually low sales for items on display in that area, confirming that the store layout negatively affects business and needs to be changed.

Sam shadows a single avatar throughout the store to experience the store and its offerings from the perspective of a customer. Using trajectories as a guide, Sam uses the *Ghost Preview* to quickly gain an overview of a persons posture at different points in time. Pinning several of these previews for a couple of different avatars allows Sam to create several personas, illustrating different customer types. Sam plans to show these to marketing colleagues, and to convince investors that a better store layout is needed.

To provide customers with a safe and clean shopping experience, Sam uses *AvatAR's* visualization of touches with the environment to see which surfaces are frequently touched and therefore need to be cleaned more often. In addition, the *Gaze Visualization* allows Sam to find which areas of the store customers look at. Sam notices that customers' gaze often falls onto a shelf next to the exit as they leave the store. Sam notes this down as an ideal placement for advertisement of upcoming sales as well as grab-and-go promotional items.

These three scenarios illustrate how *AvatAR's* techniques can be applied in analysis tasks in three different application domains. In the following section we will discuss how the techniques support a holistic understanding through individual and combined use.

6 DISCUSSION

In this work, we explored how different types of visualizations can be applied in AR to gain a better understanding of people's movement and interactions with the environment. *AvatAR's* aim is not only to allow analysts to passively observe spatial information, but enable full immersion into the experience. *AvatAR* provides different ways of synthesising a scene, either through "live" in-the-moment replay, longitudinal understanding through summary visualisations, as well as highlighting interactions with the environment.

6.1 Power in combination

The purpose of *AvatAR's* individual techniques is to enable insights into how people utilise the space (*3D Trajectories*, *Footprint Visualization*), areas of interest (*Gaze Visualisation*, *Heatmap*), as well as detailed views of people's actions (*Avatars*, *Touches with the environment*). Each of these act as basic building block enabling a detailed analysis of a particular aspects of a dataset and the overall interaction of the people. In our scenarios the individual components also demonstrate the general workflow. However, each technique is designed to be used in combination with other techniques, which has the potential to gain even further insights than using each technique in isolation. In particular, *AvatAR* allows analysts to combine detailed views (*Avatars* and *Ghost Preview*) with overview visualisations (*Trajectories*, *Spectre Visualization*, and *heatmap*). In addition, the time-based techniques (*Trajectories*, *Spectre Visualization*) not only provide a summary visualisation over a period of time, but further support analysts' awareness. By combining these components, insights can be gained, comparisons made, and interactions understood. For example, the *Spectre Visualization* as well as several pinned *Ghost Previews* can be used to create a static representation of a person's movement over a longer period of time for deeper understanding as well as sharing. Also, simply watching an avatar's spectre trails and footsteps in a space could tell an analyst whether to expect large or small movements, where avatars come from and go to, and how many avatars to expect in the scene.

The three scenarios explore a potential future of spatial data analysis using *AvatAR's* capabilities. While the target audience is grounded in existing personas (i.e., shop owners wanting to understand a store design's effectiveness; dancers wanting to learn a new routine; or physiotherapists wanting to understand progress of their clients), the underlying data for such analysis might not yet exist on a large scale: *AvatAR* requires human motion data with a high level of detail, recorded within a potentially large area (from room scale to building scale). While prototypical tracking infrastructure can create such data, as mentioned previously, it is not yet readily available and can only be created with high investment costs regarding time and money. However, the further development of algorithms that can calculate human skeletons from conventional camera images can significantly reduce this cost. While an in-the-wild deployment of *AvatAR* would have yielded deeper insights into the system's usability and use, the aim of the current work was to demonstrate the feasibility of those techniques and their power in combination to gather a holistic understanding of a scene.

6.2 Differentiation of *AvatAR* to existing approaches.

Several prior works have explored how human motion data can be visualised and analysed. After having described *AvatAR* and potential usage scenarios in the previous sections, in the following we want to highlight how *AvatAR* extends and differs from these prior works.

Ocupado [47] allows an analysts to explore temporal building occupancy data through 2D bar graphs, line charts, and similar charts. It further allows to filter the data a floorplan view. In contrast, AvatAR provides in-situ visualisations that are directly embedded in the environment they were recorded in. Furthermore, *AvatAR* visualises detailed motion data through 3D visualisation, rather than presence data. *AvatAR's* connected handheld device is similar to Ocupado's interactive floor plan in that it allows selection of regions of interest. Additionally, AvatAR's floor plan view visualises movement data (position, orientation, trajectories) of the people in the dataset as well as the position of the analyst in the physical space.

Similar to Yu et al.'s work [65], which explored how an extended reality system can be used to teach users a dance or exercise routine, our second scenario (section 5.2) comprises a dancing scenario. In contrast to this prior work, *AvatAR* does not provide motion guidance as it is aimed at analysis tasks, not at learning tasks. Similarly to Yu et al.'s work *AvatAR* allows analysts to step into a first-person view, but additionally *AvatAR* also allows them to observe data from a third-person perspective. Through this embodiment, *AvatAR* makes the data not only explorable by the analysts, but experienceable as they can "step into" the recorded user's position, with an emphasis on observing and analysing, rather than guidance or learning.

Kraus et al. [35] developed a system for 4D scene reconstruction based on image and video footage, which also includes OpenPose skeleton detection and supplementation of humans with 3D avatars. Their system allows an analyst to explore the data on a desktop as well as in a VR scene. In contrast, *AvatAR* displays the recorded data in the environment it was recorded in. It further provides visualisations that go beyond playback of the data and help in understanding of space utilization (e.g., heatmap and trajectories) and interactions with the environment (e.g., gaze, touch, and footprint). Furthermore, Kraus et al's system and *AvatAR* could complement each other by using the output of their full data procession pipeline as input to *AvatAR*, providing the time variant skeleton data loaded into *AvatAR* for in-situ study and analysis.

MIRIA [11] is an in-situ AR analysis environment similar to *AvatAR*, enabling the exploration of movement and interaction data through 3D trajectories and additional visualizations like heatmaps. While MIRIA provides an overview of movement over time, *AvatAR* is concerned with providing an overview using trajectories and heatmaps as well as a high degree of detail by visualizing a person's full body and posture, which could lead to more nuanced insights of an individual person's behavior. Another key difference is that MIRIA allows analysts to place additional visualizations on the floor or walls of an environment, while *AvatAR* embeds visualisations directly into the environment (Gaze, Footprint, and Touch visualization), directly resembling the geometry of a space. This potentially leads not only to a higher degree of immersion, but also enables a more detailed analysis of touches with surfaces of complex, non-planar geometry.

Lilija et al. [40] presented a system for analyzing the movement of objects in a VR environment. Similar to *AvatAR*, 3D trajectories provide a summary of an object's movement, proxy objects showing an object's current position, and time controls enable scrubbing through time. However, their work supports exploration of virtual spaces that recreate a real environment, whereas *AvatAR* supports in-situ exploration of the data. Additionally *AvatAR* offers several different visualizations beyond 3D trajectories, therefore posing a greater potential to gain insights into the data by giving a richer set of tools.

In summary, what differentiates *AvatAR* from other works is our focus on immersive in-situ analysis using detailed 3D humanoid avatars to understand detailed actions, complemented by 3D trajectories to understand space utilization, as well as additional visualizations embedded directly into the environment to highlight areas of interest. These visualizations can be accessed directly through the avatars as well as through a tablet device. *AvatAR* thus represents an important aspect in the overall design space for analyzing human motion data, not through individual techniques themselves, but through the entirety of the system and the way the different techniques integrate with each other.

6.3 Future Evaluation of *AvatAR*

Throughout the course of this work, we have sought to demonstrate the potential utility of *AvatAR*. We showed the practical feasibility of *AvatAR* through our prototype implementation (see section 4), as well as outlined possible use cases with the scenarios in section 5. However, it is unknown how much *AvatAR* actually supports the analysis of human motion data, as we were not able to conduct an evaluation of our system with real users due to the ongoing COVID–19 pandemic. In the future we want to build on this foundation and conduct an in-the-wild study within a concrete use case, such as an office or a construction site, in which real users use *AvatAR* to solve real problems. Such a study can help to gain insight into how *AvatAR's* techniques are used to solve real-world analysis problems, which techniques are preferred by users, and what weaknesses or enhancement potential users identify in the respective techniques.

6.4 Changes within the Environment

One of the potential advantages of embedding in-situ AR visualization into the real environment is that, in addition to their personal perspective, it becomes easier for analysts to take the perspective of a person being analyzed. This requires that the environment that analysts encounter during their analysis corresponds to the one that existed at the time the movement data was captured. However, our living and working environments are subject to constant change and both smaller objects, such as cups, keyboards, or mobile devices, as well as pieces of furniture, such as tables and chairs, can easily be moved into different positions. This perceived mismatch between environment and a person's action can on the one hand be confusing for the analyst. On the other hand, the detailed visualization can help to identify such discrepancies within the dataset. For example, when an avatar sits in mid-air, where there is no chair in the real world, an analysts could easily notice that something must have changed.

Work by Lilija [40] allows users to track objects in virtual reality and use them to jump directly to the time any changes to them might occur. Transferring such techniques into the real world requires tracking of manipulated objects in addition to people occupying it. In controlled conditions this might be easily achieved by, e.g., placing optical markers on them, but outside of such environments, tracking objects poses a significant technical challenge. Additionally, relevant objects would need to be represented as virtual objects in AR, meaning they need to be digitised. However, once both tracking data and digitized model have been created, 3D trajectories could be used to provide an overview of the objects' movement with the model representing their current position at the current point in time.

Besides moving objects, the overall condition of the environment can also change between recording and analysis of data. This includes aspects such as the lighting conditions, temperature, humidity and others, which can also have an influence on the behavior of people. For example, a person might be influenced to change their workplace by direct sunlight falling onto their screen, which might not be clear when the analysis is performed under different lighting conditions. Capturing such environmental conditions and visualizing them can therefore lead to further insights.

6.5 In-Situ vs. Remote

AvatAR's core ability is to visualize data in the same environment it was captured in, as demonstrated through our first and third scenario. In addition, *AvatAR* can be used for *remote* scenarios as demonstrated in our second scenario: Analysing data of people recorded in a different than the playback environment. Another such example is the CMU Panoptic Studio dataset [32, 55], which has been recorded in small dome, resulting in detailed 3D human recordings. Using *AvatAR*, these recordings can be played back in any other environment of at least similar size for detailed analysis. This can be especially useful for analysing complex motions, for example in learning and teaching scenarios.

Scaling of the analyzed environment is another interesting aspect. While in-situ visualizations need to be viewed in the same scale as the environment they where recorded in to keep their relationship intact, this is not the case for remote scenarios. Here, the scaling can be adjusted to fit the analysis being performed, catering to classic overview and detail techniques. For example, for a building, it is possible to look at a single floor as a world-in-miniature view to get an overview of the movement patterns, and then drill down into individual areas of the building to look at the interaction between specific people.

6.6 Relation to the Dataset and Limitations

In the following, we discuss considerations and limitations of *AvatAR* in how well it represents the underlying dataset with regard to the level of detail visualized, its length, and the number of people—and thus avatars—included. It should be mentioned that *AvatAR* is not a system designed for capturing tracking data, but a dedicated visualization framework that can be used to analyze already existing datasets from different sources.

6.6.1 Level of Detail of the Dataset. The various visualizations that make up *AvatAR*, most notably the virtual avatars themselves, are created using skeletal data, or more specifically, a list of absolute positions for the various joints of the human body. *AvatAR* is agnostic to the tracking system used to create this joint data: Any tracking system that can provide a set of joint positions in 3D space can be used with *AvatAR* to reconstruct virtual avatars. Examples include the CMU Panopticum Studio [32, 55], Vicon motion tracking systems, Microsoft's Kinect, as well as systems using simple RGB cameras [43, 60]. Even visualization of data generated from virtual sources [19, 27] is possible. While some methods of data capture may only detect a person's position and orientation, others may be able to capture detailed motion, down to movement of fingers and facial expression.

We have focused on datasets providing a moderate level of fidelity, which can be achieved using relatively fast processing of commercial RGB camera footage. In fact, our prototype fully supports the standard OpenPose 18 or 24 joint positions for the body [20] and 2x21 joint positions for the hands [54]. Therefore, all datasets that provide this level of fidelity or less can be be displayed and *AvatAR* is not limited regarding the complexity of the motion sequence. One aspect of *AvatAR* where our prototype could be expanded to benefit from even more detailed datasets is the *Gaze Visualization* technique. Currently, we use the position and orientation of the head to reconstruct the likely view cone of a person. With the availability of eye tracking data, this visualization can be improved to use the actual movement of the eyes as the basis for the visualization.

Another technique which is strongly influenced by the fidelity of the dataset is the visualization of touches with the environment. If data for each finger is available, the visualization can be created with a high degree of precision by calculating the distance for each finger with the environment individually, using a threshold of roughly one centimeter to register hits. When no finger data is available, the position of the hand is used instead to calculate possible interactions with the environment. However, the threshold needs to be much higher and is prone to generate false positives where a person's hand was close to a surface without actually touching it. *AvatAR*'s prototype implementation supports both versions and gracefully degrades to using the hand joints if no finger data is provided by the dataset.

Even more so, *AvatAR* can even be useful when fewer joint data is present than OpenPose's 18 joints, as the visualisations and concepts gracefully degrade: When only a single position is available for each person, trajectories can be used instead of avatars. When at least a position and orientation is known, an avatar can be placed into a scene facing the correct direction to provide better insights of how a person may have fit into a particular environment. If data for legs and feet were missing, a floating torso can be rendered, which could still allow analysts to gain further insights into position, actions, and interactions. The position and orientation of the head, or the presence of enough joints to reconstruct it (e.g. eyes, nose, or forehead), enables the use of the *Gaze Visualization*, and foot joints data in turn enables the *Footprint Visualization*. Therefore, the avatars and accompanying visualization techniques can adapt to the fidelity of the data set, providing rough visualizations when little joint data is present and providing fine-grained visualizations when used with a rich data set.

6.6.2 Length of the Dataset. When analyzing human motion data, there are different goals that the analyst can pursue in their analysis. For example, the analyst may be interested in the detailed inspection of short, specific motion sequences or in the longer-term understanding of larger time periods to identify motion patterns. *AvatAR* aims to support the former, focusing on datasets that are a few minutes up to an hour in length and *AvatAR's* time controls on the tablet are designed with such a use case in mind. For larger datasets in the range of hours or even days, the current controls and filtering mechanics are most like insufficient and need to be adapted. Furthermore, mid-air interaction alters the current time frame relative to the avatar, which is useful for advancing several seconds into future or past, but not suitable for skipping longer time periods. However, there is already a body of research for exploring large timelines (cf. [1]) which can be integrated into the control view of *AvatAR's* tablet interface to mitigate these issues.

6.6.3 Number of supported Avatars. Some of the proposed visualization techniques are limited regarding the number of people that can be analyzed comfortably, especially when the people are very close to each other. *AvatAR* was designed to and works best when analyzing small groups of up to around ten people. While more people are technically supported, it is advisable to use the tablet to filter the selection to people relevant to the current analysis to avoid a visual information overload. One concept of *AvatAR* is to provide not only a single 3D trajectory for each person in the dataset, but several ones for the different limbs of a person. Due to the large number of trajectories that can be generated by this, it can be difficult to distinguish between different trajectories belonging to different people and body parts. To further reduce complexity, it may therefore be beneficial to show only a single trajectory per avatar. The same applies when activating additional visualization techniques which leave visual imprints in the environment, such as the *Gaze* or *Footstep Visualization*. An extreme example is the *Specter Visualization* technique, which is impractical to use for more than one person at a time, as the large amount of intertwined semi-transparent avatar representations make it very hard to make out any distinguishing features.

To summarize, *AvatAR's* concepts are designed to support analysts in environments that are not too crowded, aiming at the detailed analysis of individuals and small groups. Future work is needed to expand these techniques to longer periods of time, larger environments, or where the goal is to understand collective crowd behavior.

7 CONCLUSION

We presented *AvatAR*, an immersive visualization environment for the in-situ analysis of human motion data. Our contribution consists of the design of a set of visualization techniques that take advantage of the combination of virtual humanoid avatars with 3D trajectories, as well as techniques for embedding visualizations directly into the real environment. All these techniques were implemented in a prototype, and we presented three scenarios demonstrating how *AvatAR's* prototype can help with spatial analysis tasks. Building on this strong foundation, we consider the next step for future work to apply our prototype to a real-world analysis problem and to perform an extensive evaluation of our concepts in this context. We are confident that *AvatAR* has the potential to provide a real benefit in the analysis of human motion data by enabling analysts to occupy the same space as the data and immerse themselves with the virtual avatars being analyzed.

ACKNOWLEDGMENTS

This work was partially funded by DFG grant 389792660 as part of TRR 248 (see https://perspicuous-computing.science) and the DFG as part of Germany's Excellence Strategy Clusters of Excellence EXC-2068 390729961 Physics of Life and EXC 2050/1 390696704 Centre for Tactile Internet with Human-in-the-Loop (CeTI) of TU Dresden.

REFERENCES

[1] Wolfgang Aigner, Silvia Miksch, Wolfgang Müller, Heidrun Schumann, and Christian Tominski. 2008. Visual Methods for Analyzing Time-Oriented Data. *IEEE Transactions on Visualization and Computer Graphics* 14, 1 (2008), 47–60. https://doi.org/10.1109/TVCG.2007.70415

[2] G. Andrienko, N. Andrienko, G. Fuchs, and J. M. C. Garcia. 2018. Clustering Trajectories by Relevant Parts for Air Traffic Analysis. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (Jan 2018), 34–44. https://doi.org/10.1109/TVCG.2017.2744322

[3] Gennady Andrienko, Natalia Andrienko, Heidrun Schumann, and Christian Tominski. 2014. *Visualization of Trajectory Attributes in Space–Time Cube and Trajectory Wall.* Springer Berlin Heidelberg, Berlin, Heidelberg, 157–163. https://doi.org/10.1007/978-3-642-32618-9_11

[4] Natalia Andrienko and Gennady Andrienko. 2013. Visual analytics of movement: An overview of methods, tools and procedures. *Information Visualization* 12, 1 (2013), 3–24. https://doi.org/10.1177/1473871612457601 arXiv:https://doi.org/10.1177/1473871612457601

[5] Natalia Andrienko, Gennady Andrienko, and Peter Gatalsky. 2003. Exploratory spatio-temporal visualization: an analytical review. *Journal of Visual Languages & Computing* 14, 6 (2003), 503 – 541. https://doi.org/10.1016/S1045-926X(03)00046-6 Visual Data Mining.

[6] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *CHI '18*. ACM, 185:1–185:15. https://doi.org/10.1145/3173574.3173759

[7] Elie Azar and Carol Menassa. 2012. Agent-Based Modeling of Occupants and Their Impact on Energy Use in Commercial Buildings. *Journal of Computing in Civil Engineering* 26 (07 2012), 506–518. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000158

[8] B. Bach, P. Dragicevic, D. Archambault, C. Hurter, and S. Carpendale. 2014. A Review of Temporal Data Visualizations Based on Space-Time Cube Operations. In *EuroVis - STARs*, R. Borgo, R. Maciejewski, and I. Viola (Eds.). The Eurographics Association. https://doi.org/10.2312/eurovisstar.20141171

[9] A. Batch, A. Cunningham, M. Cordeil, N. Elmqvist, T. Dwyer, B. H. Thomas, and K. Marriott. 2020. There Is No Spoon: Evaluating Performance, Space Use, and Presence with Expert Domain Users in Immersive Analytics. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (Jan 2020), 536–546. https://doi.org/10.1109/TVCG.2019.2934803

[10] Frederik Brudy, Suppachai Suwanwatcharachat, Wenyu Zhang, Steven Houben, and Nicolai Marquardt. 2018. EagleView: A Video Analysis Tool for Visualising and Querying Spatial Interactions of People and Devices. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces* (Tokyo, Japan) *(ISS '18)*. Association for Computing Machinery, New York, NY, USA, 61–72. https://doi.org/10.1145/3279778.3279795

[11] Wolfgang Büschel, Anke Lehmann, and Raimund Dachselt. 2021. MIRIA: A Mixed Reality Toolkit for the In-Situ Visualization and Analysis of Spatio-Temporal Interaction Data. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3411764.3445651

[12] Wolfgang Büschel, Annett Mitschick, Thomas Meyer, and Raimund Dachselt. 2019. Investigating Smartphone-Based Pan and Zoom in 3D Data Spaces in Augmented Reality. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services* (Taipei, Taiwan) *(MobileHCI '19)*. Association for Computing Machinery, New York, NY, USA, Article 2, 13 pages. https://doi.org/10.1145/3338286.3340113

[13] Wolfgang Büschel, Patrick Reipschläger, Ricardo Langner, and Raimund Dachselt. 2017. Investigating the Use of Spatial Interaction for 3D Data Visualization on Mobile Devices. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (Brighton, United Kingdom) *(ISS '17)*. ACM, New York, NY, USA, 62–71. https://doi.org/10.1145/3132272.3134125

[14] Stefan Buschmann, Matthias Trapp, and Jürgen Döllner. 2016. Animated visualization of spatial–temporal trajectory data for air-traffic analysis. *The Visual Computer* 32, 3 (2016), 371–381. https://doi.org/10.1007/s00371-015-1185-9

[15] P. Butcher, N. John, and P. Ritsos. 2021. VRIA: A Web-Based Framework for Creating Immersive Analytics Experiences. *IEEE Transactions on Visualization and Computer Graphics* 27 (2021), 3213–3225.

[16] Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs, and Harald Reiterer. 2018. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173664

[17] Katy Börner and Shashikant Penumarthy. 2003. Social Diffusion Patterns in Three-Dimensional Virtual Worlds. *Information Visualization* 2, 3 (2003), 182–198. https://doi.org/10.1057/palgrave.ivs.9500050 arXiv:https://doi.org/10.1057/palgrave.ivs.9500050

[18] Yuanzhi Cao, Tianyi Wang, Xun Qian, Pawan S. Rao, Manav Wadhawan, Ke Huo, and Karthik Ramani. 2019. GhostAR: A Time-Space Editor for Embodied Authoring of Human-Robot Collaborative Task with Augmented Reality. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 521–534. https://doi.org/10.1145/3332165.3347902

[19] Zhe Cao, Hang Gao, Karttikeya Mangalam, Qizhi Cai, Minh Vo, and Jitendra Malik. 2020. Long-term human motion prediction with scene context. In *ECCV*.

[20] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[21] M. Cavallo, M. Dholakia, M. Havlena, K. Ocheltree, and M. Podlaseck. 2019. Dataspace: A Reconfigurable Hybrid Reality Environment for Collaborative Information Analysis. In *IEEE VR '19*. 145–153. https://doi.org/10.1109/VR.2019.8797733

[22] Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree, and Mark Podlaseck. 2019. Immersive Insights: A Hybrid Analytics System ForCollaborative Exploratory Data Analysis. In *25th ACM Symposium on Virtual Reality Software and Technology* (Parramatta, NSW, Australia) *(VRST '19)*. Association for Computing Machinery, New York, NY, USA, Article 9, 12 pages. https://doi.org/10.1145/3359996.3364242

[23] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas, and Kim Marriott. 2017. ImAxes: Immersive Axes As Embodied Affordances for Interactive Multivariate Data Visualisation. In *UIST '17*. ACM, 71–83. https://doi.org/10.1145/3126594.3126613

[24] Haoran Dai, Yubo Tao, and Hai Lin. 2020. Visual analytics of urban transportation from a bike-sharing and taxi perspective. *Journal of Visualization* 23, 6 (2020), 1053–1070. https://doi.org/10.1007/s12650-020-00673-8

[25] Philip DeCamp, George Shaw, Rony Kubat, and Deb Roy. 2010. An Immersive System for Browsing and Visualizing Surveillance Video. In *Proceedings of the 18th ACM International Conference on Multimedia* (Firenze, Italy) *(MM '10)*. Association for Computing Machinery, New York, NY, USA, 371–380. https://doi.org/10.1145/1873951.1874002

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

[27] Matteo Fabbri, Fabio Lanzi, Simone Calderara, Andrea Palazzi, Roberto Vezzani, and Rita Cucchiara. 2018. Learning to Detect and Track Visible and Occluded Body Joints in a Virtual World. In *European Conference on Computer Vision (ECCV)*.

[28] J. A. W. Filho, W. Stuerzlinger, and L. Nedel. 2020. Evaluating an Immersive Space-Time Cube Geovisualization for Intuitive Trajectory Data Exploration. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (Jan 2020), 514–524. https://doi.org/10.1109/TVCG.2019.2934415

[29] H. Guo, Z. Wang, B. Yu, H. Zhao, and X. Yuan. 2011. TripVista: Triple Perspective Visual Trajectory Analytics and its application on microscopic traffic data at a road intersection. In *2011 IEEE Pacific Visualization Symposium*. 163–170. https://doi.org/10.1109/PACIFICVIS.2011.5742386

[30] N. Hoobler, G. Humphreys, and M. Agrawala. 2004. Visualizing competitive behaviors in multi-user virtual environments. In *IEEE Visualization 2004*. 163–170. https://doi.org/10.1109/VISUAL.2004.120

[31] Sebastian Hubenschmid, Johannes Zagermann, Simon Butscher, and Harald Reiterer. 2021. *STREAM: Exploring the Combination of Spatially-Aware Tablets with Augmented Reality Head-Mounted Displays for Immersive Analytics.* Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445298

[32] Hanbyul Joo, Tomas Simon, Xulong Li, Hao Liu, Lei Tan, Lin Gui, Sean Banerjee, Timothy Scott Godisart, Bart Nabbe, Iain Matthews, Takeo Kanade, Shohei Nobuhara, and Yaser Sheikh. 2017. Panoptic Studio: A Massively Multiview System for Social Interaction Capture. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017).

[33] Daniel Kepplinger, Günter Wallner, Simone Kriglstein, and Michael Lankes. 2020. *See, Feel, Move: Player Behaviour Analysis through Combined Visualization of Gaze, Emotions, and Movement.* Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376401

[34] Simon Kloiber, Volker Settgast, Christoph Schinko, Martin Weinzerl, Johannes Fritz, Tobias Schreck, and Reinhold Preiner. 2020. Immersive analysis of user motion in VR applications. *The Visual Computer* 36, 10 (2020), 1937–1949. https://doi.org/10.1007/s00371-020-01942-1

[35] M. Kraus, T. Pollok, M. Miller, T. Kilian, T. Moritz, D. Schweitzer, J. Beyerer, D. Keim, C. Qu, and W. Jentner. 2020. Toward Mass Video Data Analysis: Interactive and Immersive 4D Scene Reconstruction.. In *Sensor*, Vol. 20. https://doi.org/10.3390/s20185426

[36] D. Lange, F. Samsel, I. Karamouzas, S. J. Guy, R. Dockter, T. Kowalewski, and D. F. Keefe. 2017. Trajectory Mapper: Interactive Widgets and Artist-Designed Encodings for Visualizing Multivariate Trajectory Data. In *Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Short Papers* (Barcelona, Spain) *(EuroVis '17)*. Eurographics Association, Goslar, DEU, 103–107. https://doi.org/10.2312/eurovisshort.20171141

[37] Ricardo Langner, Marc Satkowski, Wolfgang Büschel, and Raimund Dachselt. 2021. MARVIS: Combining Mobile Devices and Augmented Reality for Visual Data Analysis. In *Proceedings of the 2021 ACM Conference on Human Factors in Computing Systems* (Yokohama, Japan). ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3411764.3445593

[38] Joel Lanir, Tsvi Kuflik, Julia Sheidin, Nisan Yavin, Kate Leiderman, and Michael Segal. 2017. Visualizing museum visitors' behavior: Where do they go and what do they do there? *Personal and Ubiquitous Computing* 21 (04 2017). https://doi.org/10.1007/s00779-016-0994-9

[39] Bokyung Lee, Michael Lee, Pan Zhang, Alexander Tessier, and Azam Khan. 2019. Semantic Human Activity Annotation Tool Using Skeletonized Surveillance Videos. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers* (London, United Kingdom) *(UbiComp/ISWC '19 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 312–315.

https://doi.org/10.1145/3341162.3343807
[40] Klemen Lilija, Henning Pohl, and Kasper Hornbæk. 2020. *Who Put That There? Temporal Navigation of Spatial Recordings by Direct Manipulation.* Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3313831.3376604
[41] Tahir Mahmood, Willis Fulmer, Neelesh Mungoli, Jian Huang, and Aidong Lu. 2019. Improving Information Sharing and Collaborative Analysis for Remote GeoSpatial Visualization Using Mixed Reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 236–247. https://doi.org/10.1109/ISMAR.2019.00021
[42] Nicolai Marquardt, Frederico Schardong, and Anthony Tang. 2015. EXCITE: EXploring Collaborative Interaction in Tracked Environments. In *Human-Computer Interaction – INTERACT 2015*, Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler (Eds.). Springer International Publishing, Cham, 89–97.
[43] Dushyant Mehta, Oleksandr Sotnychenko, Franziska Mueller, Weipeng Xu, Srinath Sridhar, Gerard Pons-Moll, and Christian Theobalt. 2018. Single-Shot Multi-Person 3D Pose Estimation From Monocular RGB. In *3D Vision (3DV), 2018 Sixth International Conference on.* IEEE. http://gvv.mpi-inf.mpg.de/projects/SingleShotMultiPerson
[44] Dinara Moura, Magy Seif el Nasr, and Christopher D. Shaw. 2011. Visualizing and Understanding Players' Behavior in Video Games: Discovering Patterns and Supporting Aggregation and Comparison. In *Proceedings of the 2011 ACM SIGGRAPH Symposium on Video Games* (Vancouver, British Columbia, Canada) *(Sandbox '11)*. Association for Computing Machinery, New York, NY, USA, 11–15. https://doi.org/10.1145/2018556.2018559
[45] A. Nakazawa, S. Nakaoka, T. Shiratori, and K. Ikeuchi. 2003. Analysis and synthesis of human dance motions. In *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, MFI2003*. 83–88. https://doi.org/10.1109/MFI-2003.2003.1232637
[46] Michael Nebeling, Maximilian Speicher, Xizi Wang, Shwetha Rajaram, Brian D. Hall, Zijian Xie, Alexander R. E. Raistrick, Michelle Aebersold, Edward G. Happ, Jiayin Wang, Yanan Sun, Lotus Zhang, Leah E. Ramsier, and Rhea Kulkarni. 2020. *MRAT: The Mixed Reality Analytics Toolkit.* Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376330
[47] Michael Oppermann and Tamara Munzner. 2020. Ocupado: Visualizing Location-Based Counts Over Time Across Buildings. *Comput. Graph. Forum* 39, 3 (2020), 127–138. https://doi.org/10.1111/cgf.13968
[48] Patrick Reipschläger, Tamara Flemisch, and Raimund Dachselt. 2021. Personal Augmented Reality for Information Visualization on Large Interactive Displays. *IEEE Transactions on Visualization and Computer Graphics* 27 (2 2021), 1182–1192. Issue 2. https://doi.org/10.1109/TVCG.2020.3030460
[49] René Rosenbaum, Jeremy Bottleson, Zhuiguang Liu, and Bernd Hamann. 2011. Involve Me and I Will Understand!–Abstract Data Visualization in Immersive Environments. In *Advances in Visual Computing*, George Bebis, Richard Boyle, Bahram Parvin, Darko Koracin, Song Wang, Kim Kyungnam, Bedrich Benes, Kenneth Moreland, Christoph Borst, Stephen DiVerdi, Chiang Yi-Jen, and Jiang Ming (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 530–540.
[50] D. Sacha, F. Al-Masoudi, M. Stein, T. Schreck, D. A. Keim, G. Andrienko, and H. Janetzko. 2017. Dynamic Visual Abstraction of Soccer Movement. *Computer Graphics Forum* 36, 3 (2017), 305–315. https://doi.org/10.1111/cgf.13189 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13189
[51] Michael Saenz, Ali Baigelenov, Ya-Hsin Hung, and Paul Parsons. 2017. Reexamining the cognitive utility of 3D visualizations using augmented reality holograms. In *IEEE VIS Workshop on Immersive Analytics: Exploring Future Interaction and Visualization Technologies for Data Analytics.*
[52] D. Schmalstieg, A. Fuhrmann, and G. Hesina. 2000. Bridging multiple user interface dimensions with augmented reality. In *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*. 20–29. https://doi.org/10.1109/ISAR.2000.880919
[53] Mickael Sereno, Lonni Besançon, and Tobias Isenberg. 2019. Supporting Volumetric Data Visualization and Analysis by Combining Augmented Reality Visuals with Multi-Touch Input. In *EuroVis '19 - Posters*. https://doi.org/10.2312/eurp.20191136
[54] Tomas Simon, Hanbyul Joo, Iain Matthews, and Yaser Sheikh. 2017. Hand Keypoint Detection in Single Images using Multiview Bootstrapping. In *CVPR*.
[55] Tomas Simon, Hanbyul Joo, and Yaser Sheikh. 2017. Hand Keypoint Detection in Single Images using Multiview Bootstrapping. *CVPR* (2017).
[56] S. Y. Ssin, J. A. Walsh, R. T. Smith, A. Cunningham, and B. H. Thomas. 2019. GeoGate: Correlating Geo-Temporal Datasets Using an Augmented Reality Space-Time Cube and Tangible Interactions. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 210–219. https://doi.org/10.1109/VR.2019.8797812
[57] Zsolt Szalavári and Michael Gervautz. 1997. The Personal Interaction Panel – a Two-Handed Interface for Augmented Reality. *Computer Graphics Forum* 16, 3 (1997), C335–C346. https://doi.org/10.1111/1467-8659.00137 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-8659.00137
[58] Anthony Tang, Michel Pahud, Sheelagh Carpendale, and Bill Buxton. 2010. VisTACO: Visualizing Tabletop Collaboration. In *ACM International Conference on Interactive Tabletops and Surfaces* (Saarbrücken, Germany) *(ITS '10)*. Association for Computing Machinery, New York, NY, USA, 29–38. https://doi.org/10.1145/1936652.1936659
[59] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko. 2012. Stacking-Based Visualization of Trajectory Attribute Data. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (Dec 2012), 2565–2574. https://doi.org/10.1109/TVCG.2012.265
[60] Timo von Marcard, Roberto Henschel, Michael Black, Bodo Rosenhahn, and Gerard Pons-Moll. 2018. Recovering Accurate 3D Human Pose in The Wild Using IMUs and a Moving Camera. In *European Conference on Computer Vision (ECCV)*.
[61] Ulrich von Zadow and Raimund Dachselt. 2017. GIAnT: Visualizing Group Interaction at Large Wall Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2639–2647. https://doi.org/10.1145/3025453.3026006
[62] J. A. Wagner Filho, M. F. Rey, C. M. D. S. Freitas, and L. Nedel. 2018. Immersive Visualization of Abstract Information: An Evaluation on Dimensionally-Reduced Data Scatterplots. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 483–490. https://doi.org/10.1109/VR.2018.8447558
[63] J. A. Walsh, J. Zucco, R. T. Smith, and B. H. Thomas. 2016. Temporal-Geospatial Cooperative Visual Analysis. In *2016 Big Data Visual Analytics (BDVA)*. 1–8. https://doi.org/10.1109/BDVA.2016.7787050
[64] Ji Soo Yi, Youn ah Kang, John Stasko, and J.A. Jacko. 2007. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (Nov 2007), 1224–1231. https://doi.org/10.1109/TVCG.2007.70515
[65] Xingyao Yu, Katrin Angerbauer, Peter Mohr, Denis Kalkofen, and Michael Sedlmair. 2020. Perspective Matters: Design Implications for Motion Guidance in Mixed Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 577–587. https://doi.org/10.1109/ISMAR50242.2020.00085
[66] Meng-Jia Zhang, Jie Li, and Kang Zhang. 2015. Using Virtual Reality Technique to Enhance Experience of Exploring 3D Trajectory Visualizations. In *Proceedings of the 8th International Symposium on Visual Information Communication and Interaction* (Tokyo, AA, Japan) *(VINCI '15)*. Association for Computing Machinery, New York, NY, USA, 168–169. https://doi.org/10.1145/2801040.2801072