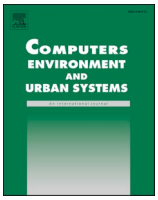
Computers, Environment and Urban Systems 100 (2023) 101923

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus



Inverse augmentation: Transposing real people into pedestrian models

Paul M. Torrens [a,*], Simin Gu [b]

[a] Department of Computer Science and Engineering, Center for Urban Science + Progress, Tandon School of Engineering, New York University, Floor 13, 370 Jay St, Brooklyn, NY 11201, USA
[b] Integrated Design & Media, Tandon School of Engineering, New York University, Floor 13, 370 Jay St, Brooklyn, NY 11201, USA

ARTICLE INFO

Keywords:
Virtual geographic environment
Virtual reality
Geosimulation
Geographic automata system
Pedestrians
Behavioral geography

ABSTRACT

We introduce a scheme for immersing real human users in urban simulations, and for enabling them to transpose their embodied behavior into models. We achieve this by inverse augmentation, flipping traditional philosophies of augmented reality. Rather than beginning with real-world scenes and embellishing them with graphics, we proceed from a base of synthetic, modeled, streetscapes filled with agent characters, which we augment with real human users. Participants are then allowed to use their natural abilities to explore the simulation scenarios. We achieve this by employing mobile virtual reality to allow users to build dynamic presence in a fused geosimulation and virtual geographic environment that they can physically view and walk around in. Our central argument is that inversion of this kind allows for the detail and nuances of human behavior to be brought directly into simulation, where they would traditionally be difficult to capture and represent. We show that close matches between real physical activity on the ground and actions in the model world can be achieved, as measured by spatial analysis and encephalography of user brain activity. We demonstrate the usefulness of the approach with an application to studying pedestrian road-crossing behavior.

"I am here / Come to me before it's too late" (Radiohead, 2016).

1. Introduction

In this paper, we address the problem of how to build realistic human behavior into pedestrian simulation. We propose that inverting the approach of augmented reality (AR) could be useful in bringing observation, theory, and experimentation into alignment in simulation. We consider the case of pedestrian simulation of busy urban streetscapes, for which faithful models of human behavior can be difficult to produce, and where improved connections to reality can be useful.

Our approach flips the classical notion of augmentation to allow real human users to participate directly in pedestrian simulations through immersion and embodiment. In contrast to AR, which usually begins with the real world and uses aspects of virtual reality (VR) to embellish tangible reality, our approach instead takes a detailed and simulation-driven virtual model of a streetscape scene and augments it with the behavior of a *real human user*, who is allowed to use their innate and intuitive abilities of vision and motion to explore experimental scenarios. We will demonstrate how inverse augmentation for pedestrian simulation can be accomplished through combinations of Virtual Geographic Environments (VGEs), geographic automata systems (GAS), and Virtual Reality Environments (VREs).

Placing human users in direct contact with a simulation greatly increases the expectations for the sense of realism that a model ought to convey. In this paper, we focus on three dimensions of that realism: presence, plausibility, and congruence. We demonstrate the usefulness of inverse augmentation with applied experiments to study pedestrian road-crossing behavior.

2. Review

Two philosophies perhaps dominate in the consideration of how to weave authentic representations of human behavior into pedestrian models.

The first philosophy is to employ modeling methods that can reproduce behaviors in simulation. For example, pedestrians can be modeled using machine-learning (Torrens, Li, & Griffin, 2011), kinetic theory (Henderson, 1971), micro-simulation (Gipps & Marksjö, 1985), cellular automata (Blue & Adler, 2001), agent-based modeling (Batty, 1997), individual-based modeling (Zou, Torrens, Ghanem, & Kevrekidis,

* Corresponding author.
E-mail address: torrens@nyu.edu (P.M. Torrens).

https://doi.org/10.1016/j.compenvurbsys.2022.101923
Received 1 July 2022; Received in revised form 13 November 2022; Accepted 29 November 2022
Available online 20 December 2022

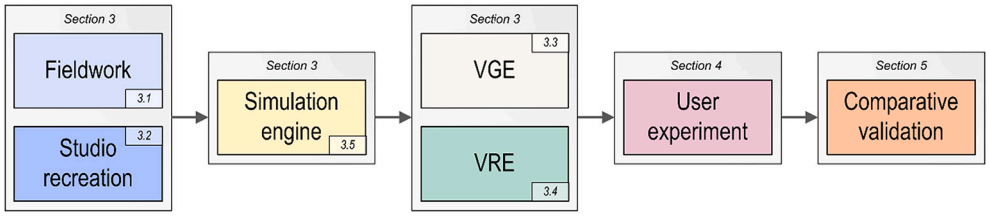


Fig. 1. The methodology used to build the study, with links to discussion in the paper text.

2012), particle modeling (Marschler, Starke, Liu, & Kevrekidis, 2014), character modeling (Pelechano, Allbeck, & Badler, 2008), and geo-simulation (Benenson & Torrens, 2004a). Each addresses different *components* of pedestrian behavior, but none can cover them all. Moreover, parity between methodology and reality perhaps always remains elusive, as our theoretical understanding of human behavior continually advances at a pace that outstrips models' abilities to catch up. (We sidestep the tenet that models should be as simple as possible and that eschew verisimilitude in favor of parsimony (Batty & Torrens, 2001; Moussaïd, Helbing, & Theraulaz, 2011; Torrens, Kevrekidis, Ghanem, & Zou, 2013). In our application, we are interested in facilitating experiments at the level of individual pedestrians: this requires high-fidelity detail.)

A second philosophy puts human users in contact with pedestrian simulations as a way to elicit user behavior. This follows traditions from computer-human interaction (CHI), although it is probably fair to argue that urban simulation has been slow to adopt developments from that field (Crookall, Martin, Saunders, & Coote, 1986). The CHI philosophy is popular in VGE modeling (Lin et al., 2015; Lin & Batty, 2011). VGEs generally combine geographic information systems (GIS) and computer-animated design and drafting (CAD) so that CAD objects are indexed to real world coordinates.

Two CHI schemes are evident in the prior art for VGEs. The first relies on sitting users in front of computer screens that display VGEs ("first-person" viewing). This is common in VGEs for architecture studies. For example, Shushan, Portugali, and Blumenfeld-Lieberthal (2016) used first-person VGEs to examine users' interpretations of homogenous and non-homogenous building designs in urban context, and Shen and Kawakami (2010) used a Web-based first-person 3D environment to explore town planning. Architectural VGEs are usually static (little if anything in the underlying model moves), as the goal is to invite users to view different renderings of built settings. Users may have the ability to pan the camera in the display software to advance the scene in different directions of the VGE, e.g., by engaging a mouse or keyboard. Omer and Goldblatt (2007) and Dong et al. (2022) showed that panning in VGEs can actually reveal wayfinding behaviors. Schwebel, Gaines, and Severson (2008) demonstrated an early implementation of using first-person animations of graphical environments as part of studying children's perceptions of crossing safety. An extension of the first-person VGE approach places human users in front of large two-dimensional displays that surround them. Orenstein, Zimroni, and Eizenberg (2015) and Meir, Oron-Gilad, and Parmet (2015) showed this approach using wrap-around projection. Natapov and Fisher-Gewirtzman (2016) introduced an interesting variation, tracking onlookers' head motion to adapt the camera view in the VGE as they watched on a screen. Roupé, Bosch-Sijtsema, and Johansson (2014) used the *Microsoft Kinect* depth-mapping sensor to translate user body movement into a controller of first-person navigation for a 3D urban environment.

A second CHI scheme for VGEs employs VR via head-mounted displays (HMDs) in lieu of computer screens, with the advantage that a three-dimensional VGE may be rendered in a user's vision and properties of distance and size become indexable to real-world counterparts, and the VGE may be experienced with depth and parallax. In these approaches, we suggest that the VGE actually becomes what we would call a "Virtual Reality Environment" (VRE). Promising results from psychology suggest that VR experiences have relevance to real-world behavior (Blascovich et al., 2002; Loomis, Blascovich, & Beall, 1999; Thompson et al., 2004). Work on VREs is relatively nascent. A very early development of the approach was introduced by Simpson, Johnston, and Richardson (2003), well before HMD devices were commercialized. After that, work in this area was relatively dormant until recently. Birenboim et al. (2019) and Nazemi et al. (2021) demonstrated that VR can be used with VGE to assess bicyclers' perception of urban environments (with riders sitting on stationary bikes). Deb, Carruth, Sween, Strawderman, and Garrison (2017) were among the first to use VR to assess pedestrian crossing behavior. Luu et al. (2022) used an unspecified VR system to examine pedestrian road-crossing. Kwon, Kim, Kim, and Cho (2022) deployed VR to examine VGE crossing environments in Korea, using *Unity* and *HTC Vive*. Jo and Jeon (2022) used spatial audio in VR to assess users' perception of VGE-based urban soundscapes in Seoul.

In this paper, we pick up this thread of research. We see four challenges in CHI-based VGE development that we seek to address.

- In the examples that we discussed above, the urban settings are stylized and fictional. Our approach uses *fieldwork to build synergies between real-world urban settings and their VRE representation*, so that there is measurable correspondence between the two.
- HMD-based VGEs remain limited relative to how real people experience urban environments. We will address this challenge *by allowing users to move around in VGEs, using their real bodies*, so that participants can infuse the models with behavior sourced from their real skills, abilities, reasoning, and intentions.
- Most VGE/VRE studies present cities as *ghost towns*. In the prior art represented by Deb et al. (2017), Luu et al. (2022), and Kwon et al. (2022), no pedestrians are included and implementation of traffic is not described: it seems the VGE is simply animated. This is limiting because there are no underlying models of dynamic objects and the VGE has no ability to react to or interact with the user. We will address this limitation *by endowing VGEs with simulation-driven models of people and vehicles*. We anticipate that this could provide new forms of social presence and cognition that are largely missing in existing research.
- Most existing VGE studies that engage CHI rely on rather simple questionnaires as a way to assess users' interactions with model conditions, which present problematic issues of bias and recall. We will address this challenge *by data-mining users' natural interactions with the simulation as they unfold*: we chiefly focus on how users move within the simulation using their natural abilities to walk, turn, and steer; what they choose to gaze upon; as well as their attention, as determined by encephalography.

In addressing these challenges, we reason that we can develop more useful parity between pedestrian models and real-world counterpart scenarios, essentially by making users' connections to models *more realistic*.

3. Methods

To support useful experimentation, we need the model to be realistic

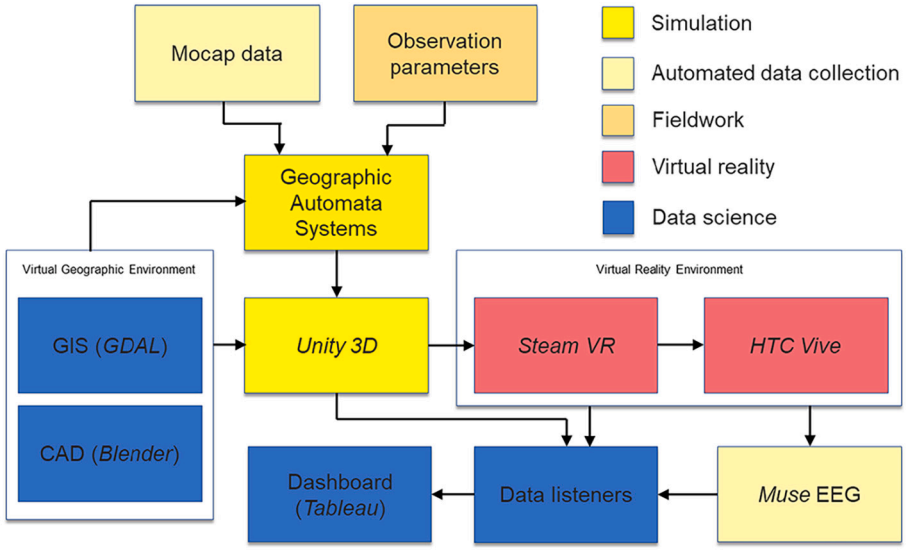


Fig. 2. The information pipeline for the experimentation system.

enough to support authentic user experiences. The prevailing idea is that realistic *presence* is important for building authentic VR experiences (Jung & Lindeman, 2021; Souza, Maciel, Nedel, & Kopper, 2021; Wilkinson, Brantley, & Feng, 2021), although Latoschik and Wienrich (2022) recently argue that *congruence and plausibility* could be key components for realism in VR, whereby congruence evokes how well a user establishes an "objective match between processed and expected information on the sensory, perceptual, and cognitive layers" (p. 4).

Our methodology (Fig. 1) focuses on all three of these dimensions of realism. We use VGEs to establish presence-based realism, relative to counterpart geographic conditions in the real world (Fig. 2). We rely on high-fidelity simulations of synthetic pedestrians and vehicles to establish plausibility. In particular, we built synthetic pedestrians with social interactions that make them relatable to human users, and established timing for model events by matching to real world observations of crossings. Additionally, we calibrated modeled pedestrians' locomotion to motion capture data of real people. We address congruence by delivering the modeled VGEs as VREs. Human users are immersed in the simulations directly, with several channels for establishing sensory realism: through tactile ground contact, vision, optic flow, depth, balance, hearing, and motion parallax.

3.1. Fieldwork

To build matches between our model and reality, we conducted a set of data collection exercises to capture:

- **Geography and geometry of streetscapes**: road width and lanes, sidewalk width, spacing of street obstacles, spacing of building façades, positioning and height of crossing infrastructure.
- **Events**: timing of traffic lights, timing of pedestrian crossing signals, audio for events.
- **Crowd patterns**: crowd density, crowd flow, size of crossing groups.
- **Pedestrian behavior**: movement speed, crossing approach, crossing behavior, crossing speed.
- **Human motion**: locomotion, ambulation.

We observed 1400 sessions across 35 different sites in New York City, which we coded using categories and valences. We used Light Detection and Ranging (LiDAR) to supplement our manual observations (Fig. 3). We placed an eight-beam LiDAR at the roadside to (1) *map* factors at the streetscape as dynamic volumetric representations of crossing context, including built structures, street furniture, civic infrastructure such as crossing signals, and individual vehicle and pedestrian motion (Fig. 3); to (2) *measure* relative distances (crossing widths, distances between buildings and sidewalks, road lane size); and to (3) *time* the speed of crossing pedestrians and durations spent waiting before crossing. These mappings, measurements, and timings were used to recreate matching agent dynamics in the VGE/VRE environment (Fig. 2). We also recorded and localized (with GPS) 360° audio of crossing environments.

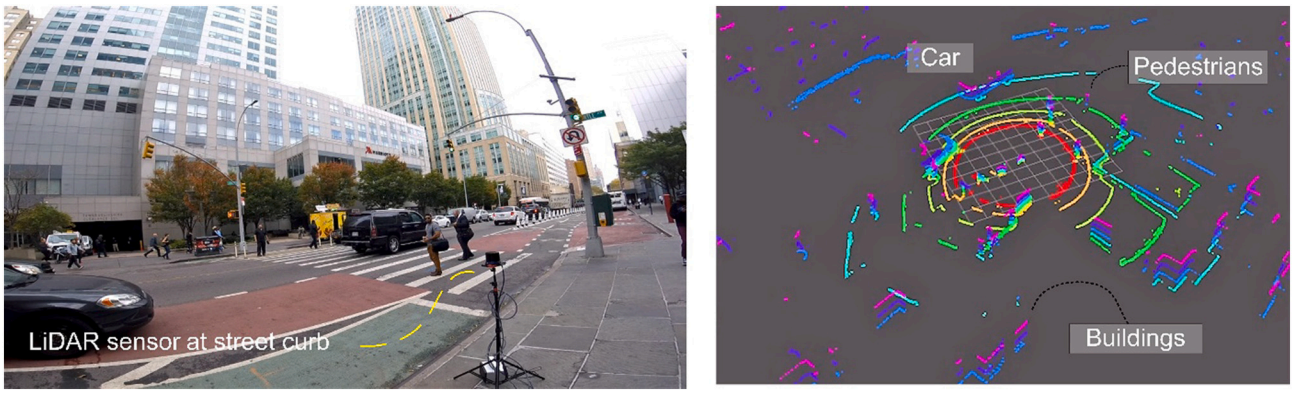


Fig. 3. Site observation in Brooklyn, NY. (Left) video scene; (Right) LiDAR data from the scene.

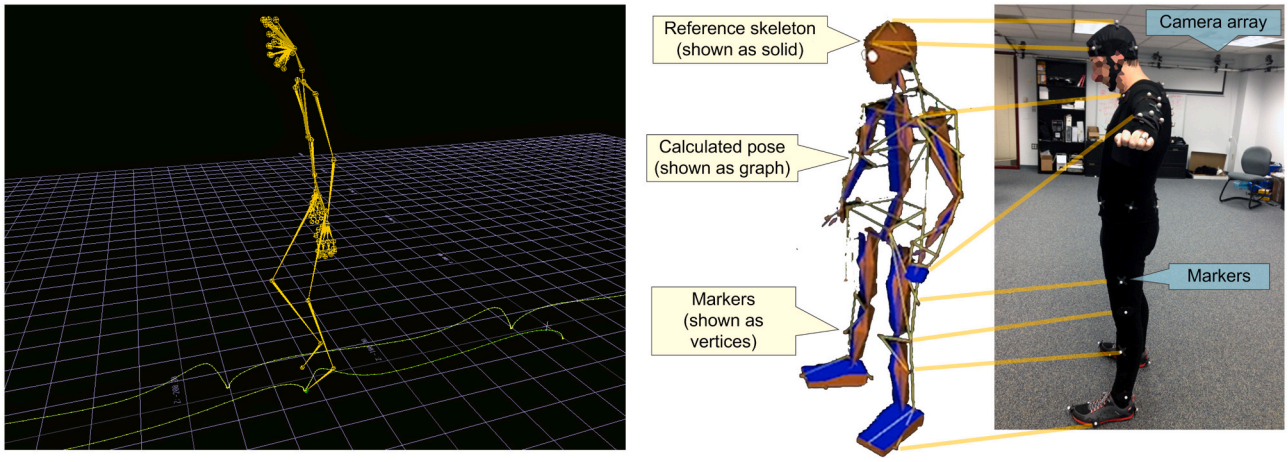


Fig. 4. Locomotion tracking using motion capture. Left: dynamic pose and locomotion path for walking capture. Right: data collection using marker-camera localization.

3.2. Studio recreation

We collected *dynamic* motion capture data of real people in a studio. Our aim was to (1) provide high-resolution (sub-centimeter, sub-decisecond) animation cycles for agent locomotion by velocity and steering decisions in the model (see Section 3.5.3); to (2) create small time geographies of motion (sub-decimeter, sub-second); and to (3) create body language that would convey realistic-looking and realistic-acting counterparts for users when they engage the simulation. Human users would therefore be able to interact with *agents that have been populated from data of real people at quite high resolutions*.

We recorded twenty motion capture sessions, split evenly between male and female subjects (aged between 20 and 48), using an array of 28 camera and infrared (IR) pulses/CMOS sensors in an indoor studio (Fig. 4). Participants in the motion capture sessions were presented with video scenes of New York City streetscapes recorded from different angles, which were displayed on large projection screens around the studio. While wearing motion capture markers (52 markers per participant), they were asked to move as if they were crossing the road in the displayed scenes. The motion capture system had a sensing resolution of 16 Megapixels and a temporal resolution of 260 Hz, producing location accuracy per marker of <1 cm over the 28-camera array. (We will provide details of how we then made this cache of motion capture data available to agents in our model in Section 3.5.3.)

3.3. Virtual geographic environment

We created a VGE replica of a real area of downtown Brooklyn, NY as the simulation environment for our model. LiDAR readings for distances were used to recreate curb, sidewalk, road, and road lane geography and to situate streetscape objects, including traffic lights, crossing signals, and lamp posts. Sidewalks in our field observations were on average 3.65 m wide and road lanes were 3 m. We used a VGE with three lanes and two sidewalks, totaling 16.3 m. Building façades were modeled as simple geometric objects, but textured with imagery matching similar land uses in Brooklyn, such as walls and store fronts.

It is traditionally challenging to implement highly dynamic objects in VGEs, because most underlying GIS do not handle moving objects well. We relied on geosimulation (Benenson & Torrens, 2004b; Torrens & Gu, 2021) to facilitate direct coupling between the simulation, GIS, and VGE/VRE. All objects and agents in the VGE are populated *both* as GIS entities and as automata entities by deploying geographic automata systems (GAS) (Torrens & Benenson, 2005).

Users of the simulation are also represented in the VGE (where we refer to them as "ego-agents"). Each user is mapped to the VGE as an inert geographic automaton: as far as simulated pedestrians know, the ego-agent is an automaton, casting state data that is available for processing by simulated pedestrians' transition rules. However, the ego-agent has no transition function and is instead directly controlled by the human user with their own movement in our lab space. Users were allowed to select their character, which we adjusted manually to represent their real height (Section 3.5.5). Users can drive the speed, velocity, linear acceleration, orientation, angular velocity, and angular acceleration of their ego-agent by simply moving their bodies. (We discuss how this is tracked and updated in the VGE in the next section.) Users may also change their view of the VGE by naturally looking around, up, and down while wearing the HMD.

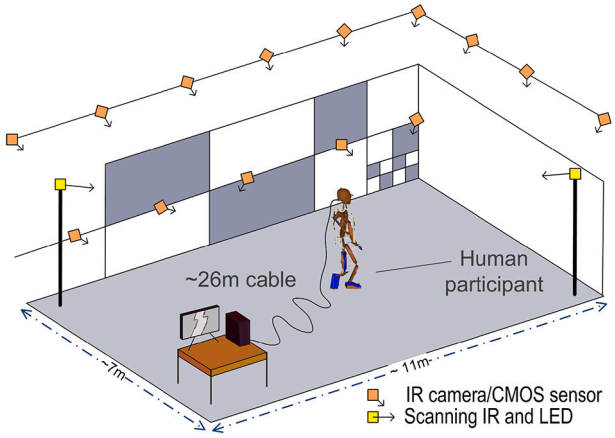


Fig. 5. Geographic passthrough is achieved by tracking a user in-studio as they use the VGE/VRE.

The geography and timing of those dynamics are then translated into the VGE. We refer to this as "geographic pass-through" (Fig. 5). We note that this is a very different approach than the typical implementations of VGE: instead of panning a camera with a mouse or keyboard, users naturally walk around our simulation using their innate locomotion abilities; they view the system by turning their heads and looking at things that interest them; they may slow-down and speed up; and they traverse real physical distances that are mapped into the VGE. Because the human participants for the study were all residing in New York City at the time of the experiments, it was hoped that geographic pass-through would establishing a realistic mental map and sense of presence in the VGE/VRE. (We note that users cannot control the ambulation of their ego-agent character; we approximated locomotion and body

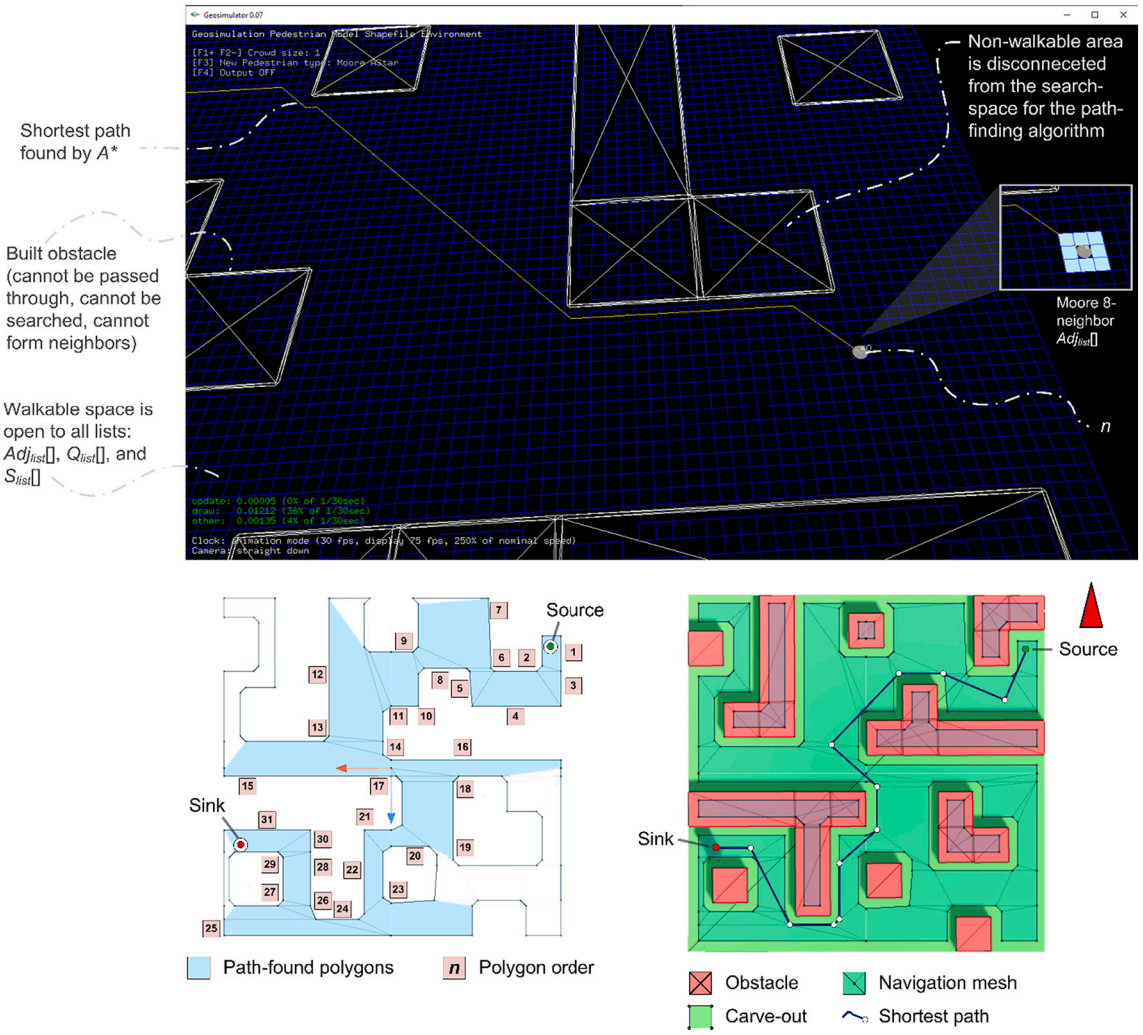


Fig. 6. Top: Synoptic pedestrian movement by path-finding. Grid squares are 1 m^2. Bottom left: path-finding by connected corridor mapping (lavender area). Bottom right: *A** shortest path from the source to the sink (graph with navy edges and white vertices). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

language by matching motion capture data to the speed of the ego-agent, as we will describe in Section 3.5.3.)

3.4. Virtual reality environment

Localization of the ego-agent in the VGE/VRE to the real-world movement of the human-participant was cross-registered in run-time using the *HTC Vive* system. *Vive* employs a pair of scanning infrared and pulsing LED diodes to broadcast light signals from fixed positions like a lighthouse (alternating between horizontal and vertical in 120° swaths). These are picked-up on a user's VR headset [using built-in Imaging Photon Detectors (IPDs)] to infer positioning and orientation by time of flight in real-time (see Niehorster, Li, and Lappe (2017) for an overview) (Fig. 5). As a user moves in a real setting (a studio in our experiments), their position in the VGE/VRE is updated to correspond to their real-world movement. As they move their head from side to side and up and down, the system will redraw the VGE/VRE from that perspective, and as they speed-up/slow-down, the system adjusted the parallax to the HMD to match. This is assisted with built-in G-sensors (to detect changes in acceleration through G-force) and gyroscopes (to detect changes in orientation and angular velocity).

These vector data can also be passed in real-time to the model, to update the positioning and steering data of the automata representation of the user as an ego-agent in the simulation. *In this way, simulated characters can react to real-time geographic data of the user directly from sensors on the HMD. In other words, there is run-time interaction directly between the HMD hardware and the simulated entities.*

The HMD uses stereo vision, which allows for a three-dimensional immersive experience with depth effects. The spatial audio that we discussed also supports this feeling of presence and depth. The field of view is 120° in horizontal (90° for one eye) and the display is rendered at 2448 × 2448 pixels per eye. The animation was refreshed at a rate of 120 Hz, with the result that all motion in the immersive VGE appears smoothly to the user. To gauge the positional accuracy of the *Vive* HMD, we ran experiments alongside motion capture by placing markers on the HMD.

We note that the *Vive* system is tethered to a desktop computer. The cable permitted up to ~26 m of movement, which is sufficient for approaching and traversing a crossing infrastructure (~16.3 m in Brooklyn). (In Section 6 we show an implementation of our system on a wireless *Oculus* HMD, which provides untethered movement.) Our observational experiments revealed that real-world pedestrians often find it hard to localize audio signals of traffic when crossing the road because of ambient city noise. We used spatial audio with the VGE/VRE HMD to play-back recorded streetscape noise, simulated cars honking, and simulated pedestrians talking to ego-agents with realistic spatial audio presence.

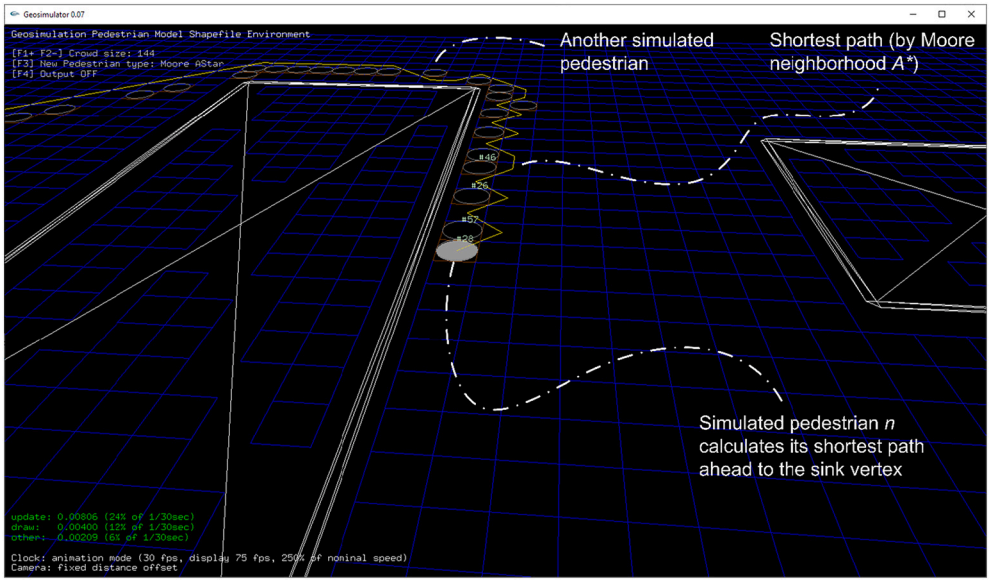


Fig. 7. Local pedestrian movement (filled gray circle) by path-finding to avoid static obstacles (white polygons), while also engaging *dynamic re-finding* to route movement around other (mobile) pedestrians (hollow gray circles). Note the zig-zag pattern to the shortest path (yellow line) to account for near-term collisions with other pedestrians. Grid squares are 400 cm^2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. Pedestrian simulation

Users control their own ego-agents in the simulation. We used GAS to control the physical movement, actions, reactions, and interactions of *simulated* pedestrians (and vehicles). We used slipstreaming (Torrens, 2015) to support exchange of geographical state information (HMD coordinates, VGE positions, navigation graphs and meshes, corridor maps, shortest paths) and non-geographical states (lists, obstacle meshes, scene graphs, textures).

3.5.1. Path-planning and path-finding

Simulated pedestrians use a path-finding model to discover paths through the streetscape and to set intermediate waypoints for interim navigation along those paths (Fig. 6). Paths are calculated and stored as graphs, with edges representing initial movement trails from a source to a sink, and vertices denoting waypoints. The graph search space accepts data on walkable areas from the VGE (by slipstreaming), e.g., so that buildings, lamps, and sign poles are excluded. We used *dynamic* lists to hold and update paths in run-time (Fig. 6). An adjacency list ($Adj_{list}[]$) denotes graph neighbors as vertices open for visitation by path-finding. A priority list ($Q_{list}[]$) keeps track of vertices that remain to be traversed/searched. A shortest-path list ($S_{list}[]$) holds the developing path as it is learned.

We deployed path-finding in two ways. The first, a synoptic approach, uses the A^* heuristic (Hart, Nilsson, & Raphael, 1968) to score the shortest path around *fixed obstacles* and operates by greedy search of the VGE graph to shift vertices between lists. It is offline (simulated pedestrians run it before they move from prescribed origins to destinations) to address what Kuipers and Levitt (1988) referred to as "large-scale space" (i.e., the entire VGE graph at macro-scale, at a cell resolution of 1 m^2). The results from synoptic navigation are introduced to the VGE using *connected corridor mapping* (Nieuwenhuisen, Kamphuis, & Overmars, 2007).

The second approach handled *local* path-finding *as a simulated pedestrian is moving* in particular sub-areas of the VGE (usually over a few square meters). This operates to modify path-finding in the presence of counter-pedestrians. We provide the local routine the locations of obstacles and pedestrians in a fixed first-order Moore neighborhood at a cell resolution of 400 cm^2 (Fig. 7) to adjust paths to dynamic conditions over relatively short distances. (We note that the distance thresholds for synoptic and local path-finding are flexibly configurable in the simulation.) Simulated pedestrians then use steering to return to their synoptic path (Section 3.5.2).

3.5.2. Dyad and crowd steering

In reality, pedestrians usually steer (over distances that are smaller

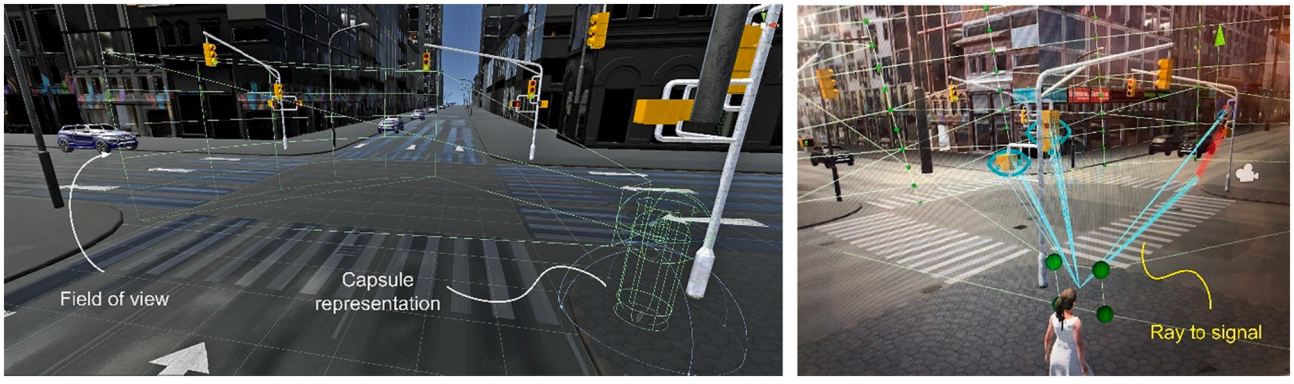


Fig. 8. Vision for simulated pedestrians. Left: A pedestrian capsule (lower right) casts a field of view in its forward vector of movement (bottom right to top left). Pedestrian signals and other simulated pedestrians or ego-agents that fall within this field of view are (dynamically) accessible as state information: shape, instantaneous position for characters, and "cross" and "do not cross" state for signals. Right: three signals are identified in the field of vision; the signal in the forward direction is identified as the relevant crossing information.

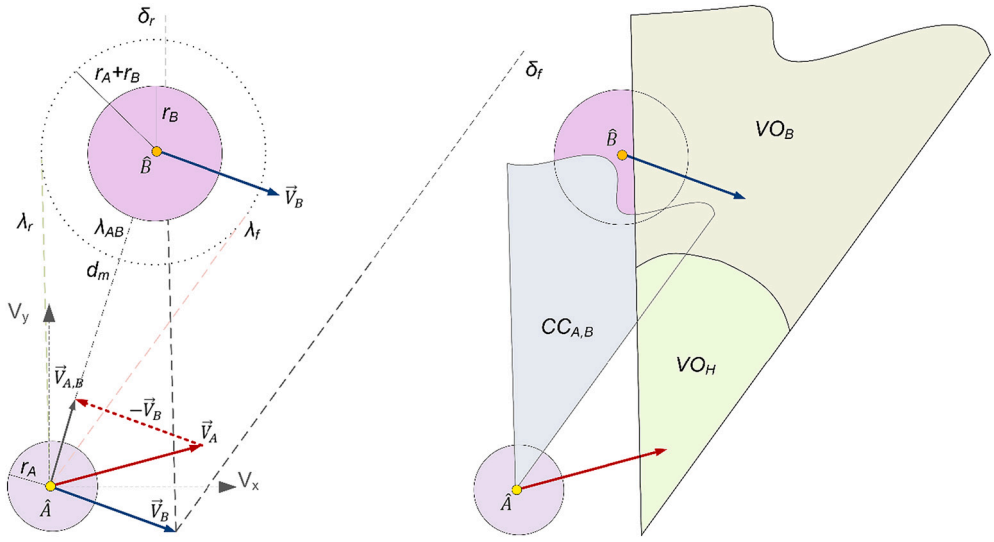


Fig. 9. Velocity obstacle (VO) for two simulated pedestrians A given B given a time horizon H.

than path adjustments) to avoid collisions with things that they *see* in their path. We implemented synthetic vision for simulated pedestrians as a field-of-view collider (Fig. 8). (We discuss vision in more detail in Section 3.5.5, where it is tied to social agency by gaze.) Simulated pedestrians were designed to steer in response to users' ego-agents by Reynolds (1999) steering. Relative to users' ego-agents, simulated entities need to be *reactive*, because information about what human users might do next is unknown to them.

Steering among simulated pedestrians, however, needs to be *interactive*, as responses from both parties need to be reconciled. We accomplished this with Velocity Obstacles (VOs). VO implementations are prone to oscillation problems. To avoid this, we used Reciprocal Velocity Obstacles (RVO) (Snape, Van Den Berg, Guy, & Manocha, 2011; van den Berg, Patil, Sewall, Manocha, & Lin, 2008), in which the steering to avoid a VO is shared. We implemented RVO for both pairwise and multiplicative steering. RVOs proceed from collision cones that we poll from simulated pedestrians' vision.

$$CC_{A,B} = \left\{ \vec{V}_{A,B} \mid \lambda_{A,B} \cap \hat{B} \neq 0 \right\}$$

Above, $CC_{A,B}$ is the collision cone from simulated pedestrian A to simulated pedestrian B (see Fig. 9). $\vec{V}_A$ is the velocity of simulated pedestrian A with instantaneous center-point $\hat{A}$ and radius r_A. $\vec{V}_B$ is the velocity of simulated pedestrian B with instantaneous center-point $\hat{B}$ and radius r_B. $\vec{V}_{A,B}$ is the relative velocity of A with respect to B. $\lambda_{A,B}$ is the line of $\vec{V}_{A,B}$. (λ_r and λ_f are tangent velocities.)

$$RVO_{A,B}\left(\vec{V}_A, \vec{V}_B\right) = \left\{ \vec{V}_A' \mid 2\vec{V}_A' - \vec{V}_A \in VO\left(\vec{V}_B\right) \right\}$$ if a reciprocal collision, where:

$VO = CC_{A,B} \oplus \vec{V}_B$ if a dyadic collision.

$VO = \cup_{i=1}^{m} VO_{B_i}$ if a multiplicative collision.

$$VO_h = \left\{ \vec{V}_A \mid \vec{V}_A \in VO, \left\| \vec{V}_{A,B} \right\| \leq \frac{d_m}{T_h} \right\}$$ if a time horizon collision

VO is the Minkowski sum ($\oplus$) of $CC_{A,B}$ and V_B. $\vec{V}_A \in VO$ denotes that the vector of simulated pedestrian A is within VO. $\left\| \vec{V}_{A,B} \right\|$ is the Euclidean norm of the relative velocity of A with respect to B). d_m is the shortest relative distance between simulated pedestrian A and simulated

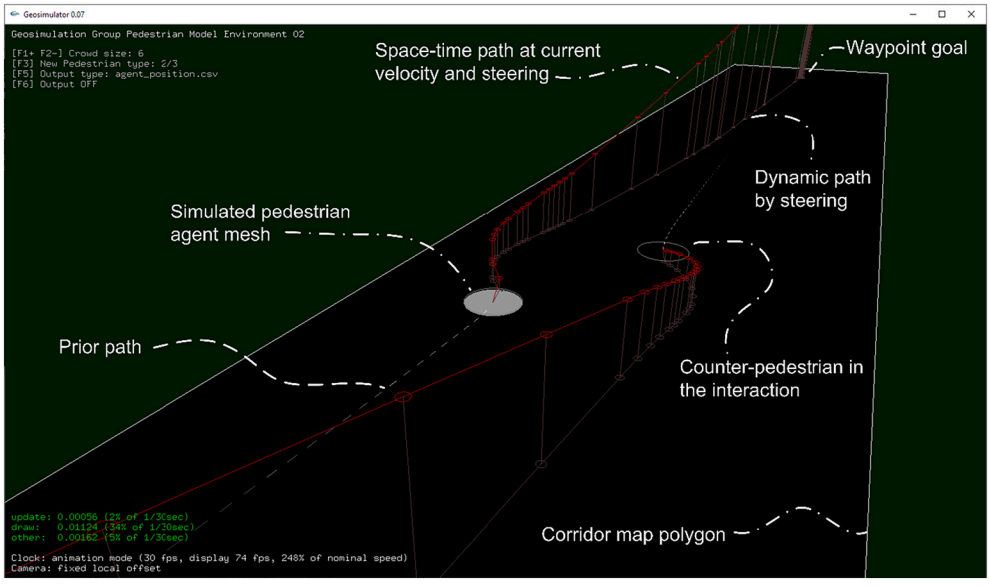


Fig. 10. Time geography in steering. Hyper-local movement of a simulated pedestrian (filled gray circle) by steering outside a counterpart RVO to avoid collision with another simulated pedestrian (hollow gray circle) within a shortest-path traversal. The pedestrian decouples from its planned path to temporarily pursue a path by steering (olive green polyline). This produces a new space-time path (three-dimensional red polyline) for movement. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

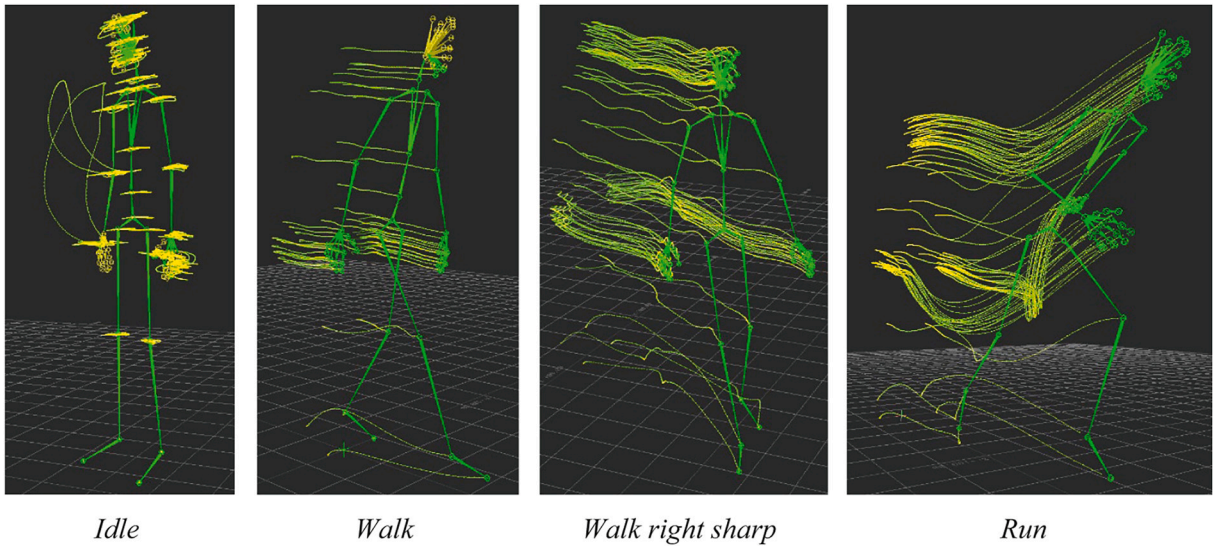


Fig. 11. Examples of motion capture trajectories for base locomotion behaviors.

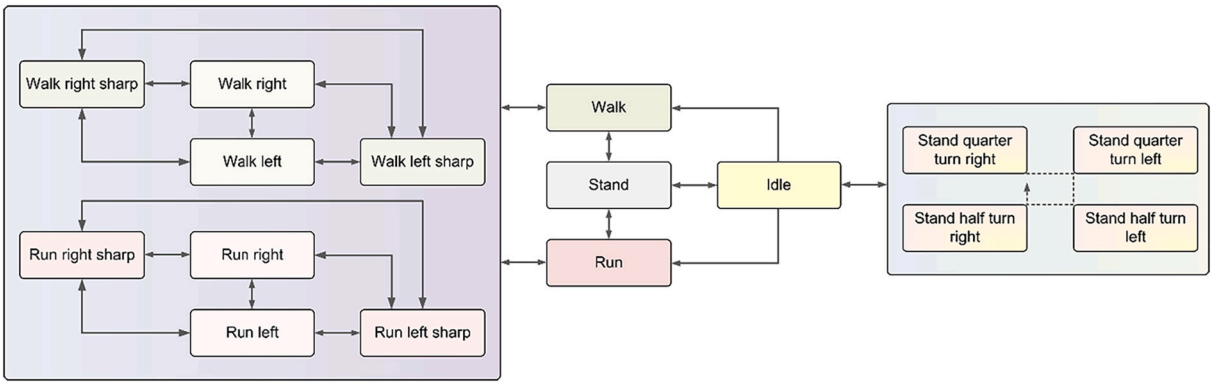


Fig. 12. Motion blending state transition diagram.

pedestrian B (Fiorini & Shiller, 1998). $\overrightarrow{V}_A'$ is a subsequent vector for simulated pedestrian A, which it uses to avoid a collision along $\overrightarrow{V}_A$.

For dyadic collisions, $RVO_{A,\ B}$ is considered as pairwise. We also adapted the RVO for two additional considerations. First, we added a *time component*: A avoids collision with B by steering such that $A(t) \cap B(t) = \varnothing$ if $\overrightarrow{V}_A(t) \notin VO(t)$, where t denotes a future time (Fiorini & Shiller, 1998). To discount collisions beyond a particular *time horizon*, T_h, we introduced VO_h. Second, we addressed cases with *multiple potential collisions*: $VO = \cup_{i=1}^{m} VO_{B_i}$, where m is the number of simulated pedestrians being considered for collision from i to m (Fiorini & Shiller, 1998).

Once free from near contact, the simulated pedestrian will engage a Reynolds (1999) seek-based steering with a lookahead time to get back to its shortest path (as in Fig. 10). If the amount of steering required to negotiate contact with multiple simulated pedestrians takes them too far away from the shortest path, the simulated pedestrian can calculate a new path from its current location to the sink.

Our use of the time horizon in RVO and lookahead times in Reynolds-type steering implements a form of *time geography* in the model: we endow the simulated pedestrians with a sense of the window of space and time available for them to reach goals within a prescribed temporal agenda (Fig. 10). This is tied to the simulated pedestrians' velocity so that they can speed-up to account for lost time in excessive steering maneuvers. This is handled straightforwardly through the nesting of ($t' \rightarrow t' + 1) \in t$.

3.5.3. Locomotion and action blending with motion capture data

For vectors produced by synoptic navigation and hyper-local steering, a matching animation cycle is invoked to articulate simulated pedestrians with matching gait, posture, and stride, i.e., *locomotion*. Animations were applied to the root motion node for each simulated pedestrian. So, movement routines drive the pedestrian's root velocity, while animation matching that velocity is applied by locomotion routines. We employed motion blending (Kovar & Gleicher, 2003) to combine data into sequences of motion (using *Unity*'s *Mecanim* system), using the motion capture data described in Section 5 (Fig. 11). We adopted 15 different animation cycles for locomotion. Three covered high-level locomotion states (idle/stand, walk, and run); 12 were used for turning within each of the three high-level states (Fig. 12). The turning animations produce locomotion for angular velocity output from Reynolds-type steering. Blends of each animation (e.g., from idle, to walk, to run) were created by first aligning animations to common spatial and temporal nodes in a graph so that transition between them would be anchored to common reference points. Blending was then used to create a weighted average across the animations by linear interpolation.

3.5.4. Movement relative to crossing signals

Crossing a road usually requires that pedestrians combine movement, halting, and visual perception of conditions and signals. We used an event model to control pedestrian crossing signal switching relative to vehicle traffic lights, with timing gleaned from our fieldwork (7 s. head start time and a 30 s. cross time). For simulated pedestrians,

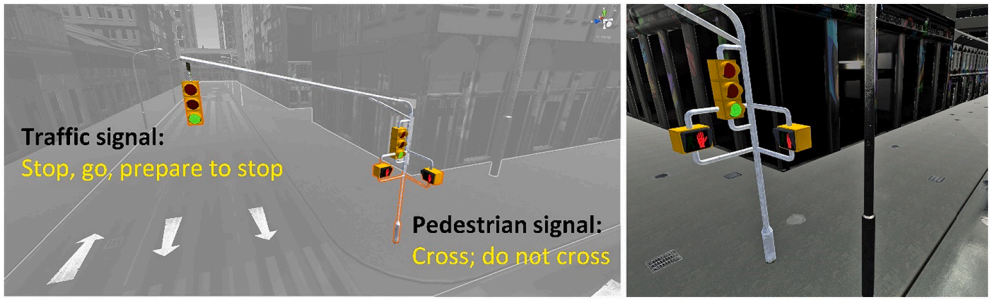


Fig. 13. Left: traffic and pedestrian crossing signals. Right: user-participant view in simulation.

Fig. 14. Avatar representation for simulated pedestrians. Left: the 13 characters. Right: detail.

Fig. 15. (Left) Dyadic gaze. The female pedestrian is looking at the ego-agent (i.e., right at you, the reader), which is in her direction of travel and field of view; the male pedestrian is looking at the female pedestrian as she is in his direction of travel and in his field of view. (Right) Group gaze. The female pedestrian on the left-hand side of the illustration has already crossed, so the pedestrian signal is low-priority in her gaze. The pedestrian on the right-hand side is closer and gets top priority. The child pedestrian is relatively far away in the field of view and receives a priority of 3.

crossing areas were coded along the sidewalks at crossing junctions, as sub-sinks for simulated pedestrians to move toward (initially, as part of path-planning). As simulated pedestrians approach crossing vertices, they switch to vision-based steering and use their synthetic vision to interpret crossing signals in their vision, following the signal that (1) falls within their field of vision, (2) in the forward direction of their path-planned movement. *Signal perception* is therefore based purely on simulated pedestrians' views of the ambient geographic information around them, such that they glance from signal-to-signal as they approach a crossing (Fig. 8). Simulated pedestrians transition their locomotion to an idle animation and their movement to a halt state if signals indicate that it is not safe to cross. If the crossing signal changed from "cross" to "do not cross" while simulated pedestrians were in the roadway, they would increase their movement velocity and transition from a walk animation to a run animation.

Crossing signals also appear as a visual signal for *human participants* to view and interpret. We used signal icons that match those that appear in New York City: a Portland orange-colored upraised hand for "do not cross" and a white-colored walking person for "cross".

Traffic signals were used to control the flow of vehicle traffic agents in the model. State data from the traffic lights are passed directly to cars to interpret. We also used visual red–yellow–green lights to indicate the signal activation visually in the model for human participants (Fig. 13).

3.5.5. Personalization

Simulated (and ego-agent) pedestrians were instantiated in the VGE/VRE from among a set of 13 different avatars, with varying age, height, and sex (Fig. 14) (using *Adobe Mixamo* meshes). Each had varying volumetric collision potential as a capsule representation for RVO (see Fig. 8). We used motion retargeting (Gleicher, 1998) to scale motion capture data to different character rigs (using *Unity*'s built-in re-targeting). For simulated pedestrians, we set the base and limit conditions for RVO velocity (maximum linear acceleration and deceleration, maximum angular acceleration and deceleration, preferred speed) from ranges of pedestrian crossing that we observed via LiDAR in fieldwork. (For ego-agents, velocity comes directly from user action.)

3.6. Gaze-based social interaction and personality

Agents were specified with different ***personalities***, governing (1) how they treat collisions with ego-agents (a scalar weight, ranging from always yield to hold your line); and (2) their relative rank-attraction to gazing at other simulated pedestrians (and ego-agents). At the lower-end of the scale, the RVO between simulated pedestrians was established with a relatively narrow time horizon; at the higher-end the time horizon was set to a relatively high value.

Rank-attraction was used to support a ***social gaze system*** for simulated pedestrians, implemented by *proactive vision*. As any object passed within a simulated pedestrian's field of view geometry (between 140° and 190°, following evidence from eye-tracking studies (Kitazawa & Fujiyama, 2010); see Fig. 8), the capsule for that geographic automaton was registered as a gaze object with shape, distance from the simulated pedestrian, and direction from their facing view (we term this as "gaze-return"). Simulated pedestrians were allowed to check the information from gaze-return against priority lists that they maintained. For simulated pedestrians and ego-agents that registered as colliders, their velocity and agent type were also passed as information with gaze-return. Simulated pedestrians that were closest in distance were given priority, and the agent would adjust its gaze to stare at them (Fig. 15). The gaze between simulated pedestrians and the ego-agent of human users established a level personal contact between users and simulated entities, which we used to create *social plausibility*.

If an ego-agent was detected and judged to be collision-imminent, the simulated pedestrian would also invoke a ***conversational interaction***. They would say either, "Get away from me", "Focus!", or "I'm sorry", in a

male or female voice per their sex. These vocalizations were played using spatial audio, such that a human participant would hear the speech ahead, to the right or left, or behind them as they passed simulated pedestrians. These gaze-and-audio combinations were used to reinforce human users' sense of the "cone of direct gaze", i.e., the range of direct gaze from counterparts that an observer regards as being directed at them (Gamer & Hecht, 2007). The cone of direct gaze is regarded as a relatively strong signal of state exchange in social interactions (Lobmaier, Savic, Baumgartner, & Knoch, 2021). It usually is interpreted as a signal of others' willingness to communicate (Balsdon & Clifford, 2018); following it with an audio exchange reinforces this in the simulation.

3.7. Vehicle model

Cars were modeled as GAS, although with relatively straightforward rules. Cars followed a navigation mesh through the city, traveling on a circuit from a prescribed set of shifting origins and destinations. (We introduced a separate road mesh for cars to plan their movement in the model; this ensured that cars only drove on the road in simulation.) Cars' dynamic motion was controlled by (1) velocity and velocity-matching; (2) adhering to traffic lights; (3) avoiding collisions with other cars; and (4) interacting with simulated pedestrians and the ego-agents of human participants. Cars did not change lanes in the simulation. Developing a more sophisticated vehicle model is a goal for future work, which we address in Section 6.

Fig. 16. Vehicle model.

The model to handle car motion, interactions with pedestrians, and adherence to traffic lights was based on a cellular automata following model (see Torrens (2004) for details). Under free movement conditions ahead of them, cars will proceed at a user-specified velocity. They will slow-down if approaching another car (by velocity-matching), and also slow if they encounter a nearby pedestrian in the roadway ahead of them (abruptly slowing to a stop, *if possible*), or a red-light traffic signal (slowing gradually to a stop).

If a simulated pedestrian (or the ego-agent of a human participant) came within a user-specified distance of a car, the vehicle would play an audible honking sound as it encounters a looming collision. This sound is designed to alert human participants in the simulation of impending collisions. Again, the sound clips were played using spatial audio in the VR headsets used by human participants, which helps them to localize the source of the warning from the simulated car.

At relatively high speed, cars may be unable to stop and may therefore collide with pedestrians. In our simulation, this only occurs with the ego-agents of human participants, i.e., simulated pedestrians and cars will always successfully navigate collision-free movement because both parties adhere to safe crossing. If a car is at a standstill (e.g., ahead of accelerating from a stop upon a green traffic light) and it encounters a pedestrian in the roadway, the car will remain at a stop, effectively yielding to the pedestrian's movement. If a traffic light turns from green to yellow to red while a car is in a junction, the car will proceed through that junction (unless it encounters a pedestrian in the way).

3.8. Data listeners and gaze detection by ray-casting

We established a set of data-listeners to collect data from the simulation during run-time. Almost any and every single piece of information that runs through our outputs from the simulation may be stored. This renders the environment tremendously useful for testing, particularly when one considers that real human participants are involved in the experiments. We streamed these to a data dashboard for ease of analysis during and after experiments. Data were collected on positions and orientations of each object—dynamically—in the simulation, in world and local coordinates (between counterpart objects). Gaze objects were handled using a dedicated listener that first registered the presence of objects within user-participants' field of vision collider and noted their position. Second, we used a ***ray tracer*** to draw rays between the eyes of avatars representing user-participants and those objects. Rays with the shortest distance along the path of travel of the ray-cast were then recorded as "gazed upon objects" (see Fig. 8).

4. Experiments

We ran two experiments using the system. Each involved a different set of human participants. In total, 43 user-participants assisted us in the study across the two experiments. We might remind the reader that *we collect huge amounts of data per participant in the system.*

Both experiments were relatively free-form in conditions, as we were interested in *how* users would behave in the VGE/VRE. Participants were told that they would appear on a sidewalk in a replica of a Brooklyn-type streetscape and we asked them to "walk around and cross the road four times at the intersections you will see ahead of you". Participants were not told how to cross the roads, whether to jaywalk or not, or given any instructions about what paths to follow. Users were not informed that simulated pedestrians would gaze at them, that they would avoid collisions with them, that they could "talk", or that cars would stop to avoid a collision with them and honk to warn them of impending danger. Users were only permitted to perform the experiment once.

Fig. 17. The VGE/VRE experiment. Left: human participant. Right: his immersive view of the simulation.

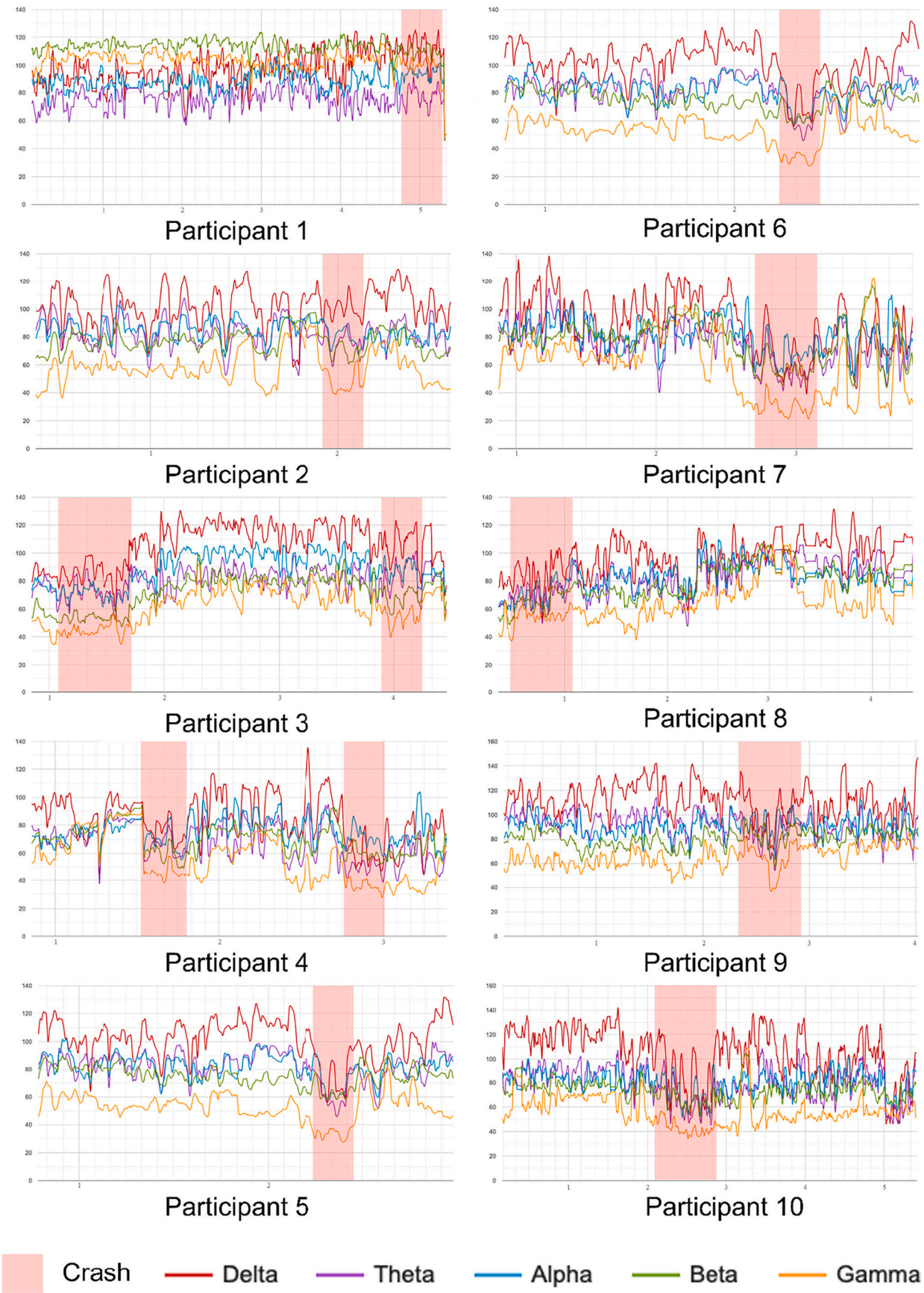


Fig. 18. EEG brain wave data for human participants. The waves have varying wavelengths and have been normalized to a 0–100 scale on the Y-axis. The X-axis represents relative time in the experiment.

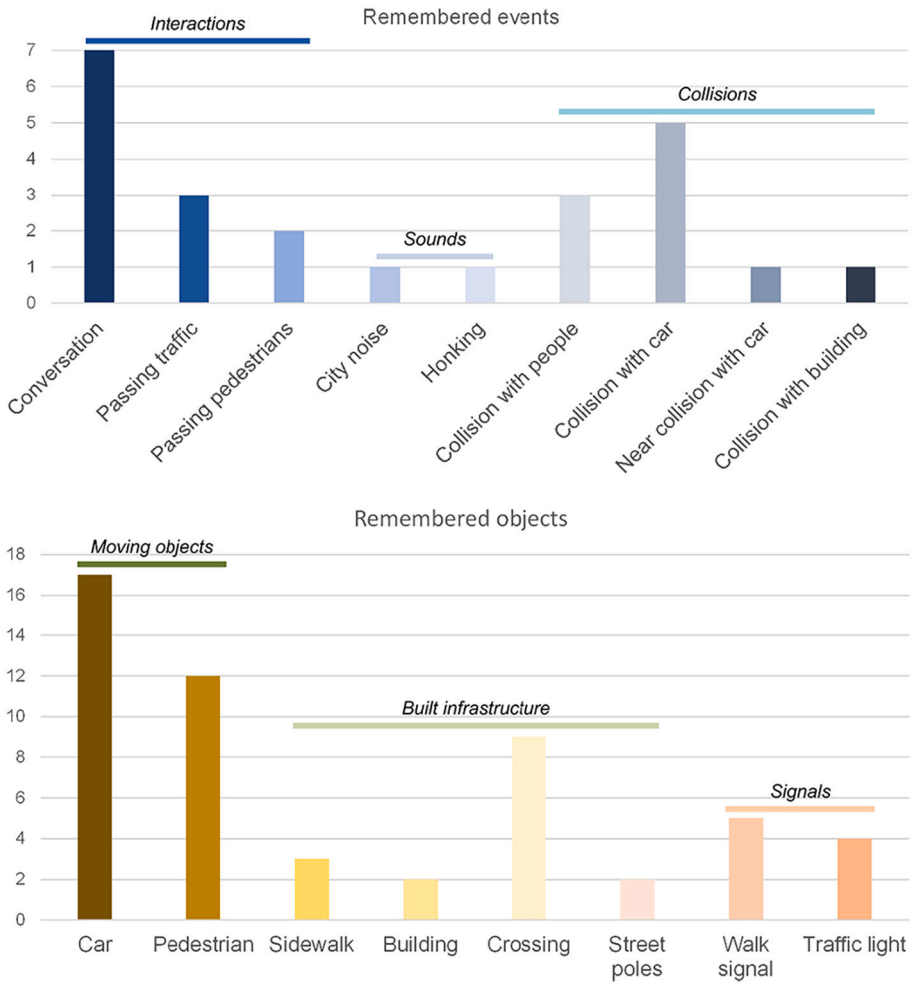


Fig. 19. Events and objects that user-participants recalled from their time in the simulation.

The first experiment tested users' neural activity in the system. Ten participants (five male and five female, aged 21–45 years, median of 24) were recruited by convenience sampling at a New York office building. None of the participants reported any physical disabilities. Six participants had little or no experience using VR before. Our field work in Brooklyn indicated an average crowd density of 39 people per block, which we used to populate the VGE/VRE streetscape with simulated pedestrians. On average, users spent ~5 min in the simulation. Participants in the first experiment wore the HMD and a *Muse* EEG headband (see Section 4.1).

In the second experiment, we recruited 33 participants by convenience sampling from among a set of staff, students, faculty, and postdoctoral researchers in a New York University academic department. 14 of the participants were male and 10 were female (nine participants chose not to report these data). The majority of participants held graduate degrees. Only one participant reported a physical disability. We did not record EEG data for the second experiment.

4.1. Collecting electroencephalogram (EEG) data from human participants

In experiment 1, we used electroencephalogram (EEG) measurement to assess the levels of concentration, alertness, and attention exhibited by human participants in the VGE/VRE. EEG measures the relative strength of a series of brain waves, as weak electrical signals emanating from the skull. We relied on a *Muse* headband, operating at a sampling rate of 500 Hz, to capture these weak signals. The headband hosts a series of electrodes and a Bluetooth chip for communicating those data wirelessly to our data listeners. The headband is worn unobtrusively by participants. The device is capable of distinguishing between gamma (32–1000 Hz), beta (13–32 Hz), alpha (8–13 Hz), theta (4–8 Hz), and delta (0.5-4 Hz) brainwaves. The *Muse* headband has been studied by other researchers and shown to be useful for research-level brainwave detection (Krigolson, Williams, Norton, Hassall, & Colino, 2017). We used four sensors: placed in AF7, AF8 (both front of head, between the pre-frontal and frontal lobes), TP9, and TP10 (both back of head, between the temporal and parietal lobes) scalp positions according to the International 10–20 EEG Electrode Placement Standard (Kabdebon et al., 2014).

4.2. Spatial analysis

For both experiments, we subjected users' data streams from the simulation run-time to spatial analyses. These included movement statistics: (1) minimum and maximum scale, (2) path length, (3) mean step size, and (4) number of moves. We calculated the fractal dimension of users' trajectories, scaled from step-by-step stride through to their entire path through the VGE, by (5) fractal dimension (Nams & Bourgeois, 2004), (6) fractal mean (Nams, 2006b), and (7) VFractal (Nams, 1996). We also calculated autocorrelation of their sinuosity, using (8) mean cosine of turning angle between successive steps, (9) probability of turning in the same direction per step, and (9) correlation among the cosines of successive turn angles (Caldwell & Nams, 2006; Nams, 2006a). For comparative validation, we assessed the performance of

Table 1
Fractal and trajectory results for 33 human participants in the VGE/VRE experiments.

ID	FD	MFD	Scale (min)	Scale (max)	Path length	Mean step size	No. of moves	Mean cosine of turning angle	Probability of turning in same direction	Correlation among cosine of successive turn angles	Survey age	Survey sex	Crash?	Moved straight?
A	1.2	1.1733	1	17.51	152	0.414	367	0.525	0.5671	−0.0016	22	F	N	Y
B	1.1266	1.1011	1	15.23	83.5	0.435	192	0.555	0.5421	0.0009	28	F	Y	Y
C	1.2358	1.1885	1	16.61	127	0.067	1905	0.369	0	−0.1243	23	M	Y	Y
D	1.1071	1.0864	1	14.05	78.3	0.056	1403	0.428	0.004	−0.1278	24	F	Y	Y
E	1.4217	1.4076	1	10.82	89	0.053	1667	0.26	0.0172	−0.1573			Y	Y
F	1.2728	1.1616	1	13.36	65.7	0.043	1534	0.243	0.0142	−0.1056			Y	Y
G	1.3388	1.312	1	10.96	140	0.043	3234	0.292	0.019	−0.1329	26	F	N	N
H	1.3178	1.2616	1	10.84	71.2	0.029	2435	0.397	0.0179	−0.1224	25	M	Y	N
I	1.233	1.2162	1	12.95	117	0.43	272	0.606	0.6556	0.0057	25	F	Y	Y
J	1.1574	1.1354	1	14	91.6	0.043	2141	0.387	0.0053	−0.1145	23	M	Y	Y
K	1.1735	1.1673	1	10.28	70.5	0.043	1646	0.364	0.0086	−0.1138	22	F	Y	Y
L	1.1368	1.0972	1	14.71	101	0.044	2267	0.366	0.0085	−0.1128	22	M	Y	Y
M	1.2951	1.239	1	10.47	94	0.04	2347	0.274	0.0102	−0.1449	26	M	Y	Y
N	1.1098	1.0898	1	13.52	135	0.053	2545	0.82	0.6577	0.0042	23	M	Y	Y
O	1.153	1.1154	1	13.71	80	0.043	1852	0.916	0.7252	0.007	30	M	N	N
P	1.1588	1.1336	1	14.42	83.9	0.056	1507	0.325	0.0073	−0.119			N	Y
Q	1.1907	1.1591	1	14.81	89.8	0.058	1542	0.416	0.0149	−0.1394	24	M	N	Y
R	1.1713	1.1501	1	15.63	105	0.071	1478	0.294	0.0212	−0.1258	23	F	N	Y
S	1.1751	1.1279	1	13.25	78.9	0.041	1935	0.245	0.0226	−0.1616	21	M	N	N
T	1.2141	1.1851	1	13.46	100	0.071	1420	0.21	0.0228	−0.1636	23	F	Y	N
U	1.253	1.1995	1	12.06	134	0.057	2328	0.161	0.0432	−0.1683			N	Y
V	1.2219	1.1869	1	15.8	182	0.039	4661	0.458	0.03	−0.2894	24	M	Y	N
W	1.1616	1.1142	1	13.69	131	0.048	2761	0.278	0.0346	−0.1527			N	Y
X	1.1347	1.0967	1	13.84	79.2	0.052	1524	0.282	0.018	−0.168	24	M	Y	Y
Y	1.0955	1.0647	1	13.15	79.7	0.049	1612	0.244	0.0431	−0.1617			Y	Y
Z	1.2727	1.2276	1	14.18	131	0.062	2112	0.185	0.0527	−0.172			N	Y
α	1.1809	1.1525	1	13.12	118	0.06	1965	0.24	0.0323	−0.1803			N	Y
ß	1.2527	1.2135	1	14.36	131	0.055	2380	0.235	0.0319	−0.1743	24	M	Y	Y
Γ	1.1427	1.1027	1	13.24	78.4	0.046	1702	0.248	0.0276	−0.1586	23	F	N	Y
Δ	1.1239	1.0777	1	13.22	77.8	0.052	1508	0.26	0.042	−0.1917	23	F	N	Y
ε	1.1755	1.1415	1	13.69	88.1	0.057	1552	0.292	0.0157	−0.1533	24	M	Y	Y
ζ	1.2886	1.2733	1	14.6	142	0.061	2319	0.247	0.016	−0.1585			Y	Y
θ	1.1642	1.135	1	17.46	118	0.048	2436	0.299	0.0255	−0.1468	28	M	N	Y

user participants in the simulation against different real-world walking and urban environments in the United States and Japan (Torrens, 2012), and standard movement routines used in mathematical and physics models (Torrens et al., 2012).

4.3. Pre- and post-experiment survey questionnaires

Participants in both experiments were asked to complete survey questionnaires. We used a pre-experiment survey to query participants' background, including demographic information (age, sex), education, and profession. We also asked participants if they had any physical disabilities. A set of questions then asked users to rank-scale their VR experiences: whether they had prior experience with the technology and whether VR caused sickness (motion sickness and dizziness are relatively common in VR use). We also surveyed participants' experiences on real-world streetscapes: whether they had confidence in their locomotive skills, whether they regarded themselves as good drivers, and whether they follow the rules of the road as a pedestrian (each on a 9-unit Likert scale).

Post-experiment surveys were administered *immediately after* users completed the experiments to capture participants' reflections on their experiences in the VGE/VRE experiments. Participants were asked to recall all *events* and *objects* that they remembered and to rate the immersive nature of the simulation on a Likert scale. They were then asked a series of free-response questions: did they face any problems in using the equipment, did they understand the experiment instructions, what is their usual strategy for driving in crowded parts of the city, where do they usually pay the most attention while crossing roads as a pedestrian, and do they have any suggestions for improvements to the experiments?

5. Results and findings

5.1. Insights from EEG

For experiment 1, we recorded EEG data during user participation. The brain wave data for each participant are shown in Fig. 18. With the exception of Participants 1 and 2, brainwave activity dropped qualitatively across all bands during the period in which they were hit by cars (shown in pink blocks in Fig. 18). Some general commonalities were observed. Data for nine participants exhibited a very marked and sudden drop in the frequency of gamma waves (shown in orange in Fig. 18). (Gamma bands did drop for Participant 1, but toward the last few instances of the crash.)

The results, although cursory, reveal some relevant findings. First, they demonstrate that each participant had a unique experience in the simulation, with individualized brain response to the experiments. Second, the brainwave data, across all bands, indicated that participants were generally alert and focused during the experiments. Gamma-bands correspond to the highest frequency waves in the brain, and they are associated with intense focus and concentration involved in perception (Meador, Ray, Echauz, Loring, & Vachtsevanos, 2002). Third, the sudden drop in gamma waves was associated with collision with a car, but was generally *preceded by a steady decline before collision* (for participants 1–7 and 10). This suggests that participants lost focus *ahead* of the collision. A drop in gamma signal is indicative of a quick shift from concentration to loss of focus and a decline in perceptual awareness.

5.2. Insights from user surveys

On average, participants self-rated their locomotive skills in real-world tasks at 7.29 out of 9. Most felt they were good drivers and

Table 2
Metrics for human walkers on real streetscapes.

Scenario/ environment	FD	MFD	Scale (min)	Scale (max)	Path length (meters)	Number of moves	Mean cosine of turning angle	Probability of turning in same direction	Correlation among cosine of successive turn angles
Salt Lake City (campus)	1.0590	1.0441	0	126.91	682.23	2200	0.721	0.028	−0.436
Salt Lake City (downtown)	1.0356	1.0295	0	951.16	2802.13	35,062	0.923	0.001	−0.477
Salt Lake City (retail)	1.0611	1.0515	0	498.93	4574.14	64,655	0.875	0	−0.482
Shibuya crossing (path 1)	1.1686	1.0876	1	54.23	152.76	226	0.403	0.189	−0.099
Shibuya crossing (path 2)	1.1764	1.1094	1	43.02	144.41	139	0.479	0.301	−0.102
Shibuya crossing (path 3)	1.1534	1.1024	1	43.28	134.71	156	0.293	0.293	0.102
Shibuya crossing (path 4)	1.2985	1.2031	1	40.33	161.47	139	0.47	0.375	−0.076
Shibuya crossing (path 5)	1.3814	1.2629	0	19.62	78.1	116	0.546	0.383	−0.155
Yokohama (path 1)	1.2084	1.184	1	393.54	6122.82	6116	0.361	0.285	−0.154
Yokohama (path 2)	1.1565	1.1037	1	287.09	1204.85	1047	0.463	0.259	−0.25
Tempe (art festival)	1.2251	1.1951	2	418.02	2759.68	3444	0.633	0.197	−0.295
Phoenix (zoo, path 1)	1.1505	1.1197	0	211.21	1589.67	166	0.702	0.494	0.003
Phoenix (zoo, path 2)	1.0927	1.0743	0	200	1214.39	121	0.644	0.437	0.003
Phoenix (zoo, path 3)	1.0935	1.0775	1	204.06	1008.41	117	0.658	0.409	−0.026
Phoenix (zoo, path 4)	1.2468	1.1891	6	56.45	280.36	30	0.714	0.500	−0.035
Phoenix (zoo, path 5)	1.1656	1.1071	4	59.83	179.77	16	0.190	0.286	−0.232
Phoenix (zoo, path 6)	1.1313	1.0656	2	53.93	150.42	18	0.704	0.625	0.093
Phoenix (zoo, path 7)	1.2715	1.2062	1	18.3	61.99	61	0.348	0.2553	−0.027
Phoenix (zoo, path 8)	1.1251	1.0915	1	238.11	1755.58	177	0.731	0.48	−0.012
Phoenix (zoo, path 9)	1.181	1.1502	1	137.03	834.53	1083	0.358	0.353	−0.036

followed the rules of the road as pedestrians (average of 5.74 out of 9, where 9 indicates that participants "strongly agree" that they follow the rules). Participants ranked the immersive experience of the VGE/VRE at an average of 7.29 out of 9 and only one participant reported having difficulties using our system.

The survey free-form question that solicited user-participants' recall of the simulation events and objects yielded narrative results. The results are parsed for simulation-specific events and objects in Fig. 19. Users most conspicuously recalled interaction events with cars and pedestrians, sounds, and collisions. The most recalled event was conversation with simulated pedestrians. User-participants also recalled collisions with simulated pedestrians and with cars with relatively high frequency. Among the objects recalled, moving objects were remembered with most frequency, followed by crossing infrastructure and traffic and pedestrian crossing signals.

The results show that user-participants had a *vivid recall* of objects that they encountered in the VGE/VRE. As we discuss in the results for gaze (Section 5.5), users' recall did also match what the system reported they actually gazed upon when in the VGE/VRE.

5.3. Insights from spatial analysis

We report spatial analysis results in Table 1 for all trajectories of human participants. In Table 2, we present a comparison set taken from real-world walkers on streets in in Phoenix, AZ; Salt Lake City, UT; Tempe, AZ; and Tokyo and Yokohama in Japan. In Table 3, we show comparison results for simulation-generated trajectories on free planes by popular movement algorithms.

On average, users in our experiments traversed 104.41 (sinuous, i.e., moving back and forth) meters in the simulation (standard deviation 27.36 m) with an average of 1895.42 discrete moves. This path length is commensurate with the traversal records that we collected for real-world streetscapes. The high number of moves recorded is a by-product of the exquisite tracking detail that we have access to within the simulation, which has a refresh rate of 90 Hz. We note that our calculations of fractal dimension are *scaled*, meaning that the high-detail of the simulation is considered in the measurement, as discussed in Nams (1996) and in Torrens et al. (2012). We also calculated mean fractal dimension (Nams, 2006b) to address this detail concern.

The average fractal dimension (FD) was 1.2, and the average mean fractal dimension (MFD) was 1.17 for our user-participant experiments. These were a match to walking recorded at a crowded arts festival in Tempe, AZ (FD = 1.23, MFD = 1.2), to a crowded zoo tour in Phoenix, AZ (FD = 1.17, MFD = 1.21), and to walking in a retail district of Yokohama (FD = 1.21, MFD = 1.18). Turning angle data from also matched the Yokohama and Phoenix examples. Our experimental average for mean cosine of turning angle was 0.35. The values for our observations in Yokohama were 0.361 and for Phoenix were 0.348. Our experimental results for correlation among cosine of successive turn angles (average of −0.13) were a match to observed data for Tokyo (−0.1546) and Yokohama (−0.1536).

The average experimental probability of turning in the same direction was 0.113, which was not a good match to any of our real-world observations. This is likely due to the instructions that we gave to

Table 3
Metrics for algorithm-driven walkers on featureless planes.

Simulated scenario	FD	MFD	Scale (min)	Scale (max)	Path length (meters)	Number of moves	Mean cosine of turning angle	Probability of turning in same direction	Correlation among cosine of successive turn angles
Relocation by hopping	1.5743	1.4398	4	68.24	8611.72	5098	0.931	0	−0.514
Brownian (long-range)	1.2812	1.2621	0	16.15	1253.9	37,188	0.845	0.636	0
Brownian (medium-range)	1.2293	1.224	0	11.51	327.97	10,031	0.839	0.638	0
Brownian (short-range)	1.1606	1.1085	0	3.31	171.067	8872	0.907	0.6482	0
Lévy flight (long-range)	1.1362	1.1309	0	57.14	3467.08	37,188	0.003	0.504	0
Lévy flight (medium-range)	1.1221	1.1104	0	53.2	1182.99	10,031	0.001	0.508	0
Lévy flight (short-range)	1.1016	1.1085	0	67.28	1035.21	8973	0	0.498	0
Random walk (long-range)	1.9983	1.9252	1	54.07	5489.01	37,188	0.703	0	−0.387
Random walk (medium-range)	1.8501	1.8053	1	18.79	1254	10,031	0.752	0	−0.415
Random walk (short range)	1.9089	1.8692	1	22.13	1126	8973	0.75	0	−0.409
Social force (simulated crowd) (path 1)	1.4462	1.2618	1	19.58	70.71	3470	1	0.728	0.900
Social force (simulated crowd) (path 2)	1.1433	1.0897	1	31.26	85.7	3470	1	0.763	1.108
Social force (simulated crowd) (path 3)	1.5139	1.3427	1	9.47	74.8	3470	1	0.884	1.032
Social force (simulated crowd) (path 4)	1.2110	1.1469	1	23.84	72.84	3444	1	0.657	1.132
Social force (simulated crowd) (path 5)	1.1152	1.0616	1	28.52	70.99	3470	1	0.790	0.784

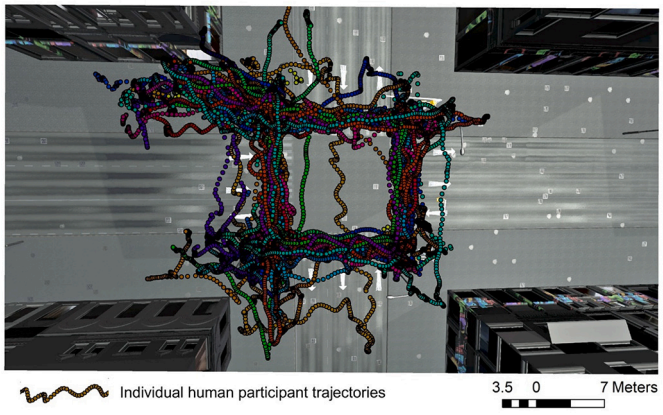


Fig. 20. Trajectory results for all 33 participants in experiment 2.

Fig. 21. Crash location map.

participants to cross all four sides of the intersection if feasible (which induces four large ninety degree turns).

No significant differences in fractal dimension or turning were noted by age or sex among user-participants in the experiments. Female participants had a median FD of 1.17, median MFD of 1.16, median cosine of turning angle of 0.33, median probability of turning in the same direction of 0.025, and median correlation among cosine of successive turning angles of −0.13. For male participants, the results were FD = 1.18, MFD = 1.14, median cosine of 0.333, median probability of 0.02, and median correlation of successive turning of −0.14.

Broadly speaking, our user-participant data shows a poor match to algorithm-driven data by popular agent-based methods). *This is a promising finding for our work, because it shows that user-participants can produce realistic walking within immersive simulations, in ways that algorithmic movement (such as agent-based models and social force models) cannot.* In other words, the movement of human users in our system is a match to recorded movement of real people walking in real-world contexts, and less well-matched to mathematically-driven agent models. Put simply: the real behavior of real users in our system comes across as *quantitatively real*. For example, we point out that all values of correlation were *negative* in our experiments, as they were in our real-world observations, but this is not the case for simulation-based movement using popular algorithms (which generally *inaccurately over-produce straight-line movement*).

Maps of user trajectories (Fig. 20) illustrated that most users traversed the sidewalk and pedestrian crosswalks. A few users veered into

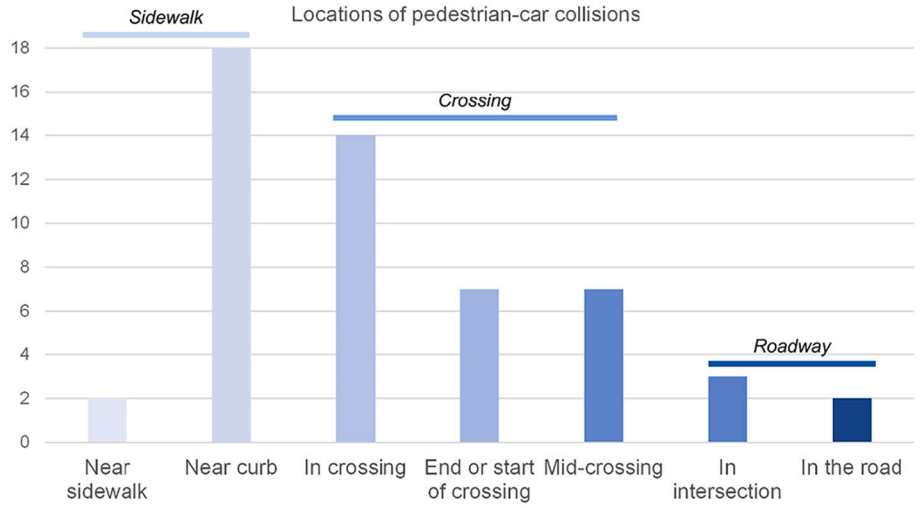


Fig. 22. Crash locations relative to the streetscape.

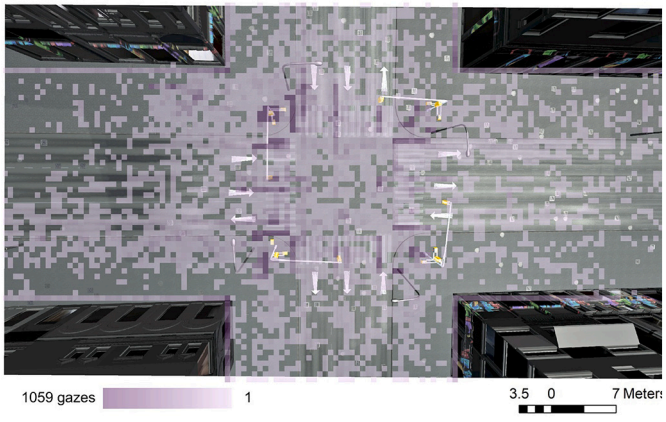
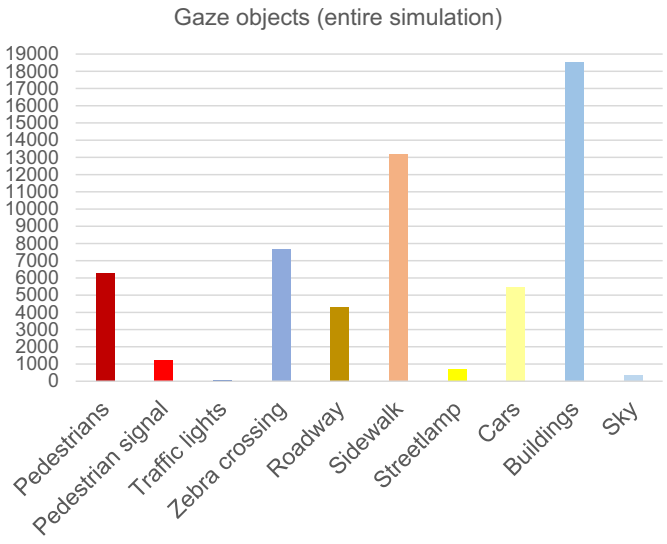


Fig. 23. Map of gaze results.



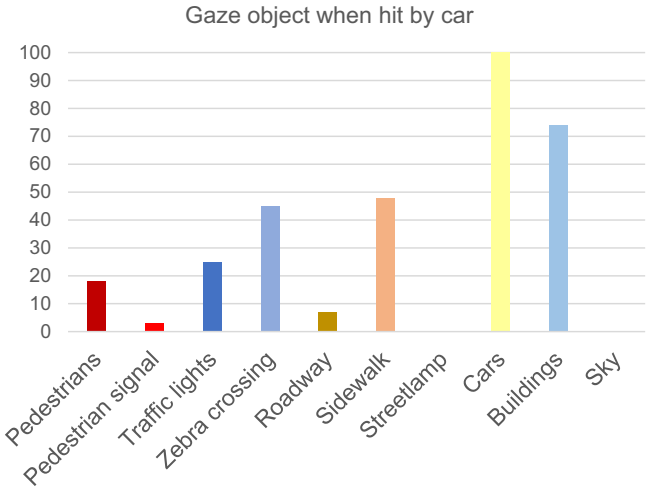


Fig. 24. Gaze objects.

the roadway at the sidewalk edge and wandered beyond the bounds of the crosswalk. Mapping revealed that most users (82%) made their first crossing attempt in a straight-ahead direction at the intersection. Specifically, most users chose to move straight ahead when walking. This follows theoretical and observational work from urban studies that has suggested that walkers in cities with linear designs tend to preserve paths that afford them far-reaching viewsheds of the streetscape ahead of them (Penn, 2003). Hillier (2007), for example, suggested that this was due to what he termed to be "configurational theory".

5.4. Insight from crash results

The data listeners that we established for the system reported when user-participants collided with a simulated car in the system (as well as all of the conditions of the simulation before, during, and after that collision). Fig. 21 maps the locations in the streetscape in which those collisions occurred (across all participants) and we provide a general overview in Fig. 22.

Across the experiments, 20 participants (60.6%) collided with a car at least once. Note that none of the cars in the simulation were able to run red lights. Seven participants collided with cars once, while 13 were in multiple collisions during their (single) exploration of the simulated streetscape (nine were hit twice, two were hit three times, and two were hit four times). Anecdotally, when asked about this at the end of the experiment, most users that collided with cars described that they wanted to see how that part of the simulation worked. This could be

Fig. 25. Our implementation of the CARLA driver behavior model with geographic automata pedestrian simulation in the system. The pedestrian is jaywalking and the vehicle has detected this and come to an abrupt stop outside of its usual traffic signal event.

addressed in future experiments by tasking participants with specifically avoiding collisions, rather than the open-ended and task-free approach that we employed.

We compared measures of fractal dimension for user-participants that collided with cars and those that did not. For users that did collide with cars, their FD was higher (1.2 vs. 1.17 for those that did not crash) and their MFD was higher (1.18 vs. 1.14 for those that did not crash). Participants that crashed also had more movements (median of 2204) than those that did not crash (median of 1893.5). This reveals that users that ended up in collisions *moved more erratically* than those that did not. (Results for mean cosine of turning angle, probability of turning in the same direction, and correlation among cosine of successive turning angles were not markedly different among the comparison groups.)

5.5. Insight from gaze results

Maps of the gaze object locations (Fig. 23) illustrate that users' gaze was largely focused on the crossing area of the simulated streetscape, particularly at the *crossing edge* and *curb* of crosswalks. The frequency of gaze per VGE object across all participants are shown in Fig. 24. Users gazed most frequently at buildings, followed by the sidewalk and zebra crossings. This reveals that much of users' gaze while in the experiment was focused on the *built environment*. Note that this is *not* what they recalled after the experiment. In their free-form surveys, participants recalled mostly *streetscape objects* in the simulation. This perhaps suggests that users took the built surroundings for granted.

Moving entities (cars and simulated pedestrians) were the second-most gazed upon objects in the simulation, and this result does match user recall from the survey data. Results for gaze demonstrate that user-

Fig. 27. AR implementation on an *Apple iPhone*.

participants tended to look at the road and cars more than at crossing signals. This makes sense: once a crossing signal indicates that it is safe to cross, a walker may usually conclude that the signal holds for a period of time (e.g., for 20 s of crossing permission) and then turns their attention to collisions with other pedestrians, as well as staying within the bounds of the pedestrian crossing.

Most users were staring at either cars or buildings at the moment that they collided with cars in the simulation (Fig. 24). In the case of buildings, this perhaps suggests that users were looking in the distance when in a crossing. In the case of cars, it reveals that users noticed cars at the last moments before a collision (recall that cars would honk to warn user-participants of impending collision). Because most of the users that did collide with cards were struck in the crosswalk, the results also indicate that participants were gazing at a variety of crossing objects (and not necessarily traffic) when struck, including the crosswalk (zebra crossing), traffic lights, and pedestrian signals.

6. Limitations and future work

The overarching aim of the work that this paper describes was to enhance the level of realism in urban modeling, principally by bringing real human users to parity with simulations of urban pedestrian dynamics. We approached this using a scheme for inverse augmentation. Chasing realism is always a daunting tasks, and our approach has some limitations that can be improved upon in future work.

The first relates to the *realistic system behavior* of the geographic automata. Our application is to pedestrian behavior, but commensurate driver behavior would be a useful counterfoil. We have been experimenting with coupling the CARLA simulator (Dosovitskiy, Ros, Codevilla, Lopez, & Koltun, 2017) with our system (Fig. 25). CARLA offers some sophisticated driving dynamics, but currently lacks matching pedestrian dynamics. Building connections between the two could be

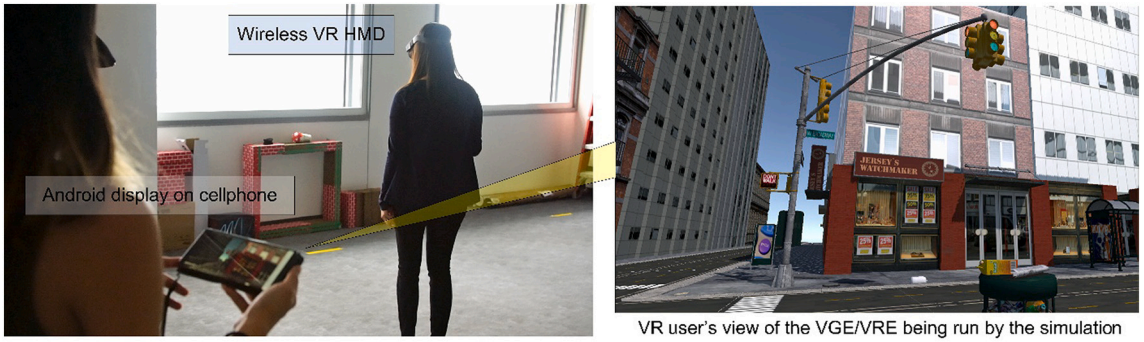


Fig. 26. Wireless implementation of the system on the *Oculus* HMD and a remote *Android* viewscreen.

useful.

The second relates to the *plausibility of the VGE/VRE setting for user participants*, relative to their real-world experiences. That so many users actually were struck by cars in the simulation is obviously unrealistic. An experimental design, e.g., games with a purpose (von Ahn, 2006), that rewards users for avoiding collisions could straightforwardly resolve this. We are also exploring whether it is feasible to support multiple human users in the simulation at once, which would open-up the possibility of placing real and synthetic confederates in the model to establish plausible dynamics of mimicry and peer behavior.

The third relates to *congruence*. Reliance on a tethered HMD can be limiting, although no users mentioned so in our experiments. Recently, we have ported the system to a wireless HMD (Fig. 26), although wireless systems currently lack significant computing resources to run high-detail simulations.

Fourth, while we regard our use of encephalography to empirically assess user congruence with the system as promising (we are unaware of any other urban modeling literature that has attempted this), we have some work ahead of us to develop EEG as a diagnostic simulation tool. In particular, we hope to do more exhaustive EEG testing to study which simulation and user events correspond to. This will invariably require that we collect matching data for real-world scenarios.

Fifth, we experimented with porting the system to augmented reality (AR) platforms. An early prototype is shown in Fig. 27, wherein our entire system was run directly via *Unity* on an *Apple iPhone* using *AR Kit*. The "VGE" in this case is a real street in Brooklyn, which can be docked to our system through map-matching. The use of an augmented approach, alongside our inverse augmentation, would support richer comparisons between the real-world and our simulations.

7. Conclusions

In this paper, we contend that approaches from AR can be inverted to emphasize human users in participant roles within simulations. Ultimately, this leads to a new class of VGE which we might consider to be "Augmented Virtual Environments" (AGEs), which combine VGEs, geosimulation, and GIS with real-world and user-driven dynamics. We demonstrated that emphasizing human users in simulation (*alongside* synthetic model processes such as agency) allows simulation scenarios to be brought to closer parity with real-world behaviors of pedestrians than might otherwise be possible with purely agent-driven models (our tests against ABMs in Table 3 show this empirically). We regard this finding as significant in drawing simulation experiments into more useful alignment with theory, and perhaps with decision support systems.

Key, in pulling this off, is to accomplish parity with a broad sense of realism, which we considered along three dimensions of presence, plausibility, and congruence. To address presence, we fused VGEs, VREs, and geosimulation to establish a dynamic and lively substrate for immersive representation of urban scenes. To tackle plausibility, we parameterized that substrate with high-detail models, going as far as to include data from fieldwork and studio-based examination of locomotion. To support congruence, we introduced the concept of geographic passthrough as a way to map simulated spaces and phenomena to real-world physicality of walking and looking around.

We highlight some key details involved in building these capabilities. First, we developed a coupled VGE with realistic behavior-driven agent models for vehicles and pedestrians. This allowed us to populate simulated urban scenes with realistically dynamic activity. Agents were designed with very high-fidelity detail and implemented as geographic automata for efficient functioning within VGE systems. Agent locomotion was calibrated to real human users' motion capture data. All agent characters in the model reacted naturally to human participants as they would in the real world, including gaze and interest, velocity-based steering, avoiding collisions, and engaging in conversation. Second, the model allowed real human participants to directly immerse themselves in the simulation in real-time using VR HMDs. Rather than simply standing in front of a screen, users were able to move physically in the real world and their behavior was directly reflected in the VGE/VRE and adapted to by its simulated characters. Third, we developed a series of data listeners to continually monitor the simulation and export massive troves of information, in real-time, as the simulation unfolds. This allowed us to build very detailed records of cause-and-effect type relationships during simulation-assisted experiments. We showed how these data can be used to perform comparative behavioral analyses in the simulation, and that the outputs can be paired to real-world observations. Key in this design was implementation of a gaze system that allowed us to note every single thing that a user looked at during their time in the simulation, as well as where they were when they did so, and what was going on around them. This opens up the possibility that the simulation environment can be used as a proxy for real-world context. We showed, for example, that it can be used to assess the fleeting decisions that users make when crossing the road. Fourth, we performed experiments to monitor the brain activity of users while engaged in the experiments using sensor-assisted electroencephalography. While preliminary, this work suggests that users' neurological appreciation for the simulation has bearing on (and perhaps linkage to) their real-world brain activity of perception, action, and cognition.

Declarations of Competing Interest

This research was funded by the National Science Foundation under Grant No. 2027652 and 1729815.

References

von Ahn, L. (2006). Games with a purpose. *Computer, 39*(6), 92–94.
Balsdon, T., & Clifford, C. W. (2018). How wide is the cone of direct gaze? *Royal Society Open Science, 5*(8), Article 180249.
Batty, M. (1997). Predicting where we walk. *Nature, 388*(6637), 19–20.
Batty, M., & Torrens, P. M. (2001). Modeling complexity: The limits to prediction. *CyberGeo, 201*(1). Online 1035.
Benenson, I., & Torrens, P. M. (2004a). *Geosimulation: Automata-based modeling of urban phenomena*. London: John Wiley & Sons.
Benenson, I., & Torrens, P. M. (2004b). Geosimulation: Object-based modeling of urban phenomena. *Computers, Environment and Urban Systems, 28*(1/2), 1–8.
van den Berg, J., Patil, S., Sewall, J., Manocha, D., & Lin, M. (2008). Interactive navigation of multiple agents in crowded environments. In E. Haines, & M. McGuire (Eds.), *Proceedings of the 2008 symposium on interactive 3D graphics and games (I3D), February 15–17, 2008, Redwood City, CA* (pp. 139–147). New York: Association for Computing Machinery.
Birenboim, A., Dijst, M., Ettema, D., de Kruijf, J., de Leeuw, G., & Dogterom, N. (2019). The utilization of immersive virtual environments for the investigation of environmental preferences. *Landscape and Urban Planning, 189*, 129–138.
Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*(2), 103–124.
Blue, V., & Adler, J. (2001). Cellular automata microsimulation for modeling bi-directional pedestrian walkways. *Transportation Research Part B, 35*, 293–312.
Caldwell, I. R., & Nams, V. O. (2006). A compass without a map: Tortuosity and orientation of eastern painted turtles (Chrysemys picta picta) released in unfamiliar territory. *Canadian Journal of Zoology, 84*(8), 1129–1137.
Crookall, D., Martin, A., Saunders, D., & Coote, A. (1986). Human and computer involvement in simulation. *Simulations and Games, 17*(3), 345–375.
Deb, S., Carruth, D. W., Sween, R., Strawderman, L., & Garrison, T. M. (2017). Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics, 65*, 449–460.
Dong, W., Qin, T., Yang, T., Liao, H., Liu, B., Meng, L., & Liu, Y. (2022). Wayfinding behavior and spatial knowledge acquisition: Are they the same in virtual reality and in real-world environments? *Annals of the American Association of Geographers, 112*(1), 226–246.
Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., & Koltun, V. (2017). *CARLA: An open urban driving simulator*. arXiv preprint arXiv:1711.03938.
Fiorini, P., & Shiller, Z. (1998). Motion planning in dynamic environments using velocity obstacles. *The International Journal of Robotics Research, 17*(7), 760–772.
Gamer, M., & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 705–715.
Gipps, P. G., & Marksjö, B. (1985). A microsimulation model for pedestrian flows. *Mathemathics and Computers in Simulation, 27*, 95–105.
Gleicher, M. (1998). Retargetting motion to new characters. In C. C. Yang, & T. M. Murali (Eds.), *Proceedings of SIGGRAPH 98: The 25th annual conference on computer graphics and interactive techniques, Orlando, FL, July 19–24, 1998* (pp. 33–42). New York: Association for Computing Machinery.

Hart, P. E., Nilsson, N. J., & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics, 4*(2), 100–107.
Henderson, L. F. (1971). The statistics of crowd fluids. *Nature, 229*(5284), 381–383.
Hillier, B. (2007). *Space is the machine: A configurational theory of architecture*. London: Space Syntax.
Jo, H. I., & Jeon, J. Y. (2022). Perception of urban soundscape and landscape using different visual environment reproduction methods in virtual reality. *Applied Acoustics, 186*, Article 108498.
Jung, S., & Lindeman, R. W. (2021). Perspective: Does realism improve presence in VR? Suggesting a model and metric for VR experience evaluation. *Frontiers in Virtual Reality, 2*, Article 693327.
Kabdebon, C., Leroy, F., Simmonet, H., Perrot, M., Dubois, J., & Dehaene-Lambertz, G. (2014). Anatomical correlations of the international 10–20 sensor placement system in infants. *Neuroimage, 99*, 342–356.
Kitazawa, K., & Fujiyama, T. (2010). Pedestrian vision and collision avoidance behavior: Investigation of the information process space of pedestrians using an eye tracker. In W. W. F. Klingsch, C. Rogsch, A. Schadschneider, & M. Schreckenberg (Eds.), *Pedestrian and evacuation dynamics 2008* (pp. 95–108). Berlin: Springer.
Kovar, L., & Gleicher, M. (2003). Flexible automatic motion blending with registration curves. In D. Breen, & M. Lin (Eds.), *Proceedings of the 2003 ACM SIGGRAPH/ Eurographics symposium on computer animation, San Diego, CA, July 26–27, 2003* (pp. 214–224). San Diego, California: Eurographics Association.
Krigolson, O. E., Williams, C. C., Norton, A., Hassall, C. D., & Colino, F. L. (2017). Choosing MUSE: Validation of a low-cost, portable EEG system for ERP research. *Frontiers in Neuroscience, 11*, 109.
Kuipers, B. J., & Levitt, T. S. (1988). Navigation and mapping in large scale space. *AI Magazine, 9*(2), 25.
Kwon, J.-H., Kim, J., Kim, S., & Cho, G.-H. (2022). Pedestrians safety perception and crossing behaviors in narrow urban streets: An experimental study using immersive virtual reality technology. *Accident Analysis & Prevention, 174*, Article 106757.
Latoschik, M. E., & Wienrich, C. (2022). Congruence and plausibility, not presence: Pivotal conditions for XR experiences and effects, a novel approach. *Frontiers in Virtual Reality, 3*, Article 694433.
Lin, H., & Batty, M. (2011). *Virtual geographic environments: A primer*. Redlands, CA: ESRI Press.
Lin, H., Batty, M., Jørgensen, S. E., Fu, B., Konecny, M., Voinov, A., Torrens, P. M., Lu, G., Zhu, A.-X., Wilson, J. P., Gong, J., Kolditz, O., Bandrova, T., & Chen, M. (2015). Virtual environments begin to embrace process-based geographic analysis. *Transactions in GIS, 19*(4), 1–5.
Lobmaier, J. S., Savic, B., Baumgartner, T., & Knoch, D. (2021). The cone of direct gaze: A stable trait. *Frontiers in Psychology, 12* (online).
Loomis, J., Blascovich, J., & Beall, A. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers, 31*(4), 557–564.
Luu, D. T., Eom, H., Cho, G.-H., Kim, S.-N., Oh, J., & Kim, J. (2022). Cautious behaviors of pedestrians while crossing narrow streets: Exploration of behaviors using virtual reality experiments. *Transportation Research Part F: Traffic Psychology and Behaviour, 91*, 164–178.
Marschler, C., Starke, J., Liu, P., & Kevrekidis, I. G. (2014). Coarse-grained particle model for pedestrian flow using diffusion maps. *Physical Review E, 89*(1), Article 013304.
Meador, K. J., Ray, P. G., Echauz, J. R., Loring, D. W., & Vachtsevanos, G. J. (2002). Gamma coherence and conscious perception. *Neurology, 59*(6), 847.
Meir, A., Oron-Gilad, T., & Parmet, Y. (2015). Can child-pedestrians' hazard perception skills be enhanced? *Accident Analysis & Prevention, 83*, 101–110.
Moussaïd, M., Helbing, D., & Theraulaz, G. (2011). How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences, 108*(17), 6884–6888.
Nams, V. O. (1996). The VFractal: A new estimator for fractal dimension of animal movement paths. *Landscape Ecology, 11*(5), 289–297.
Nams, V. O. (2006a). Detecting oriented movement of animals. *Animal Behaviour, 5*(72), 1197–1203.
Nams, V. O. (2006b). Improving accuracy and precision in estimating fractal dimension of animal movement paths. *Acta Biotheoretica, 54*(1), 1–11.
Nams, V. O., & Bourgeois, M. (2004). Fractal analysis measures habitat use at different spatial scales: An example with American marten. *Canadian Journal of Zoology, 82* (11), 1738–1747.
Natapov, A., & Fisher-Gewirtzman, D. (2016). Visibility of urban activities and pedestrian routes: An experiment in a virtual environment. *Computers, Environment and Urban Systems, 58*, 60–70.
Nazemi, M., van Eggermond, M. A. B., Erath, A., Schaffner, D., Joos, M., & Axhausen, K. W. (2021). Studying bicyclists' perceived level of safety using a bicycle simulator combined with immersive virtual reality. *Accident Analysis & Prevention, 151*, Article 105943.
Niehorster, D. C., Li, L., & Lappe, M. (2017). The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research. *i-Perception, 8*(3), 1–23.
Nieuwenhuisen, D., Kamphuis, A., & Overmars, M. H. (2007). High quality navigation in computer games. *Science of Computer Programming, 67*(1), 91–104.
Omer, I., & Goldblatt, R. (2007). The implications of inter-visibility between landmarks on wayfinding performance: An investigation using a virtual urban environment. *Computers, Environment and Urban Systems, 31*(5), 520–534.
Orenstein, D. E., Zimroni, H., & Eizenberg, E. (2015). The immersive visualization theater: A new tool for ecosystem assessment and landscape planning. *Computers, Environment and Urban Systems, 54*, 347–355.
Pelechano, N., Allbeck, J., & Badler, N. I. (2008). *Virtual crowds: Methods, simulation, and control*. San Rafael, CA: Morgan & Claypool.
Penn, A. (2003). Space syntax and spatial cognition: Or why the axial line? *Environment and Behavior, 35*(1), 30–65.
Radiohead. (2016). *Tinker tailor soldier sailor rich man poor man beggar man thief. A moon shaped pool*. London: Warner/Chappell Music Ltd.
Reynolds, C. W. (1999). Steering behaviors for autonomous characters. In Miller Freeman Game Group (Ed.), *Proceedings of game developers conference, San Jose, CA* (pp. 763–782). San Francisco, CA: Miller Freeman Game Group.
Roupé, M., Bosch-Sijtsema, P., & Johansson, M. (2014). Interactive navigation interface for virtual reality using the human body. *Computers, Environment and Urban Systems, 43*, 42–50.
Schwebel, D. C., Gaines, J., & Severson, J. (2008). Validation of virtual reality as a tool to understand and prevent child pedestrian injury. *Accident Analysis & Prevention, 40* (4), 1394–1400.
Shen, Z., & Kawakami, M. (2010). An online visualization tool for internet-based local townscape design. *Computers, Environment and Urban Systems, 34*(2), 104–116.
Shushan, Y., Portugali, J., & Blumenfeld-Lieberthal, E. (2016). Using virtual reality environments to unveil the imageability of the city in homogenous and heterogeneous environments. *Computers, Environment and Urban Systems, 58*, 29–38.
Simpson, G., Johnston, L., & Richardson, M. (2003). An investigation of road crossing in a virtual environment. *Accident Analysis & Prevention, 35*(5), 787–796.
Snape, J., Van Den Berg, J., Guy, S. J., & Manocha, D. (2011). The hybrid reciprocal velocity obstacle. *IEEE Transactions on Robotics, 27*(4), 696–706.
Souza, V., Maciel, A., Nedel, L., & Kopper, R. (2021). Measuring presence in virtual environments: A survey. *ACM Computing Surveys, 54*(8), 163.
Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence Teleoperators and Virtual Environments, 13*(5), 560–571.
Torrens, P. M. (2004). Geosimulation, automata, and traffic modeling. In P. Stopher, K. Button, K. Haynes, & D. Hensher (Eds.), *Vol. 5. Handbook of transport geography and spatial systems* (pp. 549–565). Bingley: Emerald Publishing Group Ltd.
Torrens, P. M. (2012). Moving agent pedestrians through space and time. *Annals of the Association of American Geographers, 102*(1), 35–66.
Torrens, P. M. (2015). Slipstreaming human geosimulation in virtual geographic environments. *Annals of GIS, 21*(4), 325–344.
Torrens, P. M., & Benenson, I. (2005). Geographic automata systems. *International Journal of Geographical Information Science, 19*(4), 385–412.
Torrens, P. M., & Gu, S. (2021). Real-time experiential geosimulation in virtual reality with immersion-emission. In *Proceedings of the 4th ACM SIGSPATIAL international workshop on GeoSpatial simulation* (pp. 19–28). Beijing, China: Association for Computing Machinery.
Torrens, P. M., Kevrekidis, I., Ghanem, R., & Zou, Y. (2013). Simple urban simulation atop complicated models: Multi-scale equation-free computing of sprawl using geographic automata. *Entropy, 15*(7), 2606–2634.
Torrens, P. M., Li, X., & Griffin, W. A. (2011). Building agent-based walking models by machine-learning on diverse databases of space-time trajectory samples. *Transactions in Geographic Information Science, 15*(s1), 67–94.
Torrens, P. M., Nara, A., Li, X., Zhu, H., Griffin, W. A., & Brown, S. B. (2012). An extensible simulation environment and movement metrics for testing walking behavior in agent-based models. *Computers, Environment and Urban Systems, 36*(1), 1–17.
Wilkinson, M., Brantley, S., & Feng, J. (2021). A mini review of presence and immersion in virtual reality. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 65*(1), 1099–1103.
Zou, Y., Torrens, P. M., Ghanem, R., & Kevrekidis, I. G. (2012). Accelerating agent-based computation of complex urban systems. *International Journal of Geographical Information Science, 26*(10), 1917–1937.