BoidVR: An Agent Simulation Environment Based on Freehand and Virtual Reality

Nicola Capece, Ugo Erra, Gilda Manfredi, and Rocco Di Bello
University of Basilicata, Department of Mathematics, Computer Science, and Economics

***Abstract*—Freehand-based Interaction Techniques (FITs) are an emerging technology of importance for effective Virtual Reality (VR). In this regard, most modern Head Mounted Displays (HMDs), such as the Oculus Quest, are equipped with inside-out cameras for egocentric hand tracking and natural hand gestures recognition. Following this technological evolution, we have developed an agent-based modeling and simulation VR application, called BoidVR, focused on freehand-based interaction with simulated agents. Our system enables users to view and interact with both the agents and the virtual environment using FIT-based tools. We illustrate a number of specific gestures that allow users to manipulate the virtual environment and affect agent behaviors. Finally, we conduct a user experiment in which we evaluate the usability and user sentiment of our BoidVR application.**

■ **AGEND-BASED** simulation is a common way to implement autonomous characters or individuals to represent crowds and bird-like objects in coordinated group motion generally referred to as "boids" [1], [2]. In this type of simulation, a boid has a local behavior model and moves by coordinating with the motion of other boids. That is, each individual must take decisions according to the behavior of only its neighbors, and so must be able to identify these individuals among all others in its virtual world.

Digital Object Identifier 10.1109/MCG.2022.3208682

In the past, several models have been defined from a mathematical point of view by modeling each organism individually [1], [2]. These models use only local perception and a set of simple behavior rules such as alignment, cohesion, and separation. The rules handled by parameters that regulate the behavior model are capable of providing a realistic-looking representation of flocks of birds and other creatures such as schools of fish or herds of animals. In recent years, simulation of crowds has used agent simulation to predict the behaviors of people in dangerous situations such as fires, earthquakes, terrorist attacks, and many more [3]. In this context, virtual reality (VR) can play a fundamental role [4] because users in

 Published by the IEEE Computer Society

the simulation experience high immersion within the agent environment. However, VR alone limits users' experiences to only visual feedback. The addition of freehand-based interaction techniques (FITs) augments this with tactile sensations, and this higher level of immersion provides a more compelling way to interact with the simulation in real-time. FITs and VR were originally developed in separate contexts, but with the evolution of both technologies, they have begun to converge into a fully immersive VR experience [5]. FITs create a fluid and realistic user interface, which better exploits the natural and innate abilities and skills of the user. FIT domains in many cases involve virtual entities and objects which are associated with real information. Through their direct, natural, and continuous handling, the user can manipulate the information representing such entities. FITs have also proven useful in providing interactions that do not cause the user too much stress in terms of attention and cognitive load. We have developed an application called BoidVR, which allows the user to interact with an agent-based system through the HMD Oculus Rift and Leap Motion Controller. BoidVR enables a fully immersive experience in the 3D agent's scene, where the user assists in the simulation in the first person, moves in the surrounding space, and interacts with the boids. Interaction is enabled through the user's predefined movements and hand gestures. In this way, the user can edit and control the design of the virtual environment and directly influence the collective behavior of virtual agents. BoidVR allows users to interact with the agent environment in multiple ways. BoidVR features fall into two categories. The first category allows modification of parameters which characterize a single agent. This category allows the user to vary the weight of the fundamental rules which manage the agents, so that one or more rules can dominate over the others. It is also possible to modify the range of interaction between agents, to search in the neighborhood, and change their field of view (FOV). The second category of features allows direct interaction between the user and the group of agents, allowing the user to create paths and force the group to follow them. Other features facilitate creating, activating, deactivating, and resizing obstacles, and allowing the dynamic insertion of special agents such as leaders, threats, and victims. BoidVR demonstrates how users can more readily achieve a fluid and dynamic VR experience through the use of freehand interaction techniques.

THE AGENT SIMULATION MODEL

any animal groups such as fish schools and bird flocks clearly display structural order, with the behavior of the organisms integrated so that, even though they may change shape and direction, they appear to move as a single coherent entity. Individual-based computer simulations are a very useful analytical tool for studying such groups and demonstrate that group leadership, hierarchical control, and global information are not necessary for collective behavior. In the collective-behavior model, three rules play a crucial part in the simulations. In the first rule, individuals attempt to maintain a minimum distance between themselves and others at all times. This rule, called *separation*, usually has the highest priority and corresponds to a frequently observed behavior of animals in nature. If individuals are not performing an avoidance maneuver to maintain a minimum distance, they tend to be attracted toward other individuals (to avoid being isolated) using a *cohesion* rule, and also tend to align themselves with neighbors using an *alignment* rule. BoidVR uses an agent schooling model inspired by the works of Couzin [1] and Reynolds [2]. In this model, each individual has a strictly local perception of the space it occupies. None of the creatures in the group has full knowledge of the entire group. The decisions of every individual take into account only local neighbors that are perceived within its field-of-view (*FOV*). Groups are composed of individuals. At every moment, each individual has a position, a direction, a speed, and a maximum turning rate. Individuals simultaneously determine a new desired direction of travel by considering neighbors within two behavioral zones. The first zone, called the *zone of repulsion*, has a local *interaction range*. Each individual attempts to avoid collision between itself and another individual in this zone by turning away. If neighbors are not detected in the zone of repulsion, the individual tends to align with neighbors in the second zone, called the *zone of alignment and cohesion*. This zone

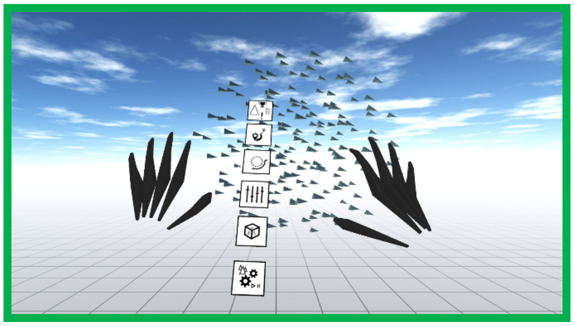

Figure 1. The basic BoidVR app appearance. The image shows flock of boids, one or two hands, and the main UI controls.

is an annulus of inner radius and outer radius around the individual. The global behaviour of each individual is a mix of separation, cohesion, and alignment rules which are weighted using three parameters to balance their orientation and attraction preferences.

BOIDVR

In BoidVR, agents are dynamic entities located in space and time; they interact with the surrounding virtual environment and the user's actions, see Figure 1. A video showing our system in action is available at [[IEEE: please include URL to video hosted in the IEEE CSDL here]]. Users interact with the virtual simulation through their hands. The left hand selects a 3D object, and the right hand interacts with the selected object using user interface controls. To create a more realistic experience, the users and their hands are represented inside the scene and considered by the agents as obstacles to avoid. We also distinguish special agents such as leaders, threats, and victims from the normal neutral agents.

Movements and Interaction Gestures

BoidVR implements several movement patterns and well-defined configurations using predefined finger and hand gestures. A first gesture is touching the virtual object using the hand or part of it, such as a finger, as shown in the first two images in Figure 2 (green border). A second gesture may be used to grasp the virtual objects (red border). This pinch gesture is performed by closing the tips of the index and thumb fingers. When this gesture ends, the grabbed object will be dropped. Another important gesture is scaling, which can be performed by using the grasping gesture performed through both hands on the desired object to be scaled, as shown in the yellow-bordered images of Figure 2 The object's scale will increase when the user separates the hands and decrease when the user closes the hands.

Interaction Components

To interact with the simulation, we implemented two types of components. The first type of components belong to the visible interface and are useful for activating and deactivating features and modifying simulation parameters. The second type includes all the virtual objects that the user can manipulate through natural gestures to control the collective behavior. We designed each component with a form, dimension, and color, allowing them to be combined and made proportionate in the virtual environment, to reduce the user's cognitive stiffness. All the designed components respond to user input through visual feedback, each with its own mechanism.

Basic components of the user interface

A basic component of our user interface is the button, which has two states, enabled and disabled, which can be modified through a simple finger touch (Figure 3 blue border). The current state of the button is presented visually by changing the border color or varying the texture opacity. To allow the user to change some parameter values in a well-defined range, we implemented a graphics slider component (Figure 3 red border). The slider consists of a vertical bar and an indicator, which changes color based on user interaction in three different states: when the user performs the grasping gesture, the slider's indicator becomes green; if the distance between the indicator and the hands is below a threshold, the indicator changes to yellow; finally, when the user does not interact with the slider, its indicator remains black. The user can build and add a new object inside the scene through a generator feature that generates anchoring objects. As shown in Figure 3 (yellow borders), the user pinches the first object, moves, and releases it at a generic point in virtual space. In the releasing phase the next object will become the current object and a new object will be created; otherwise, the object

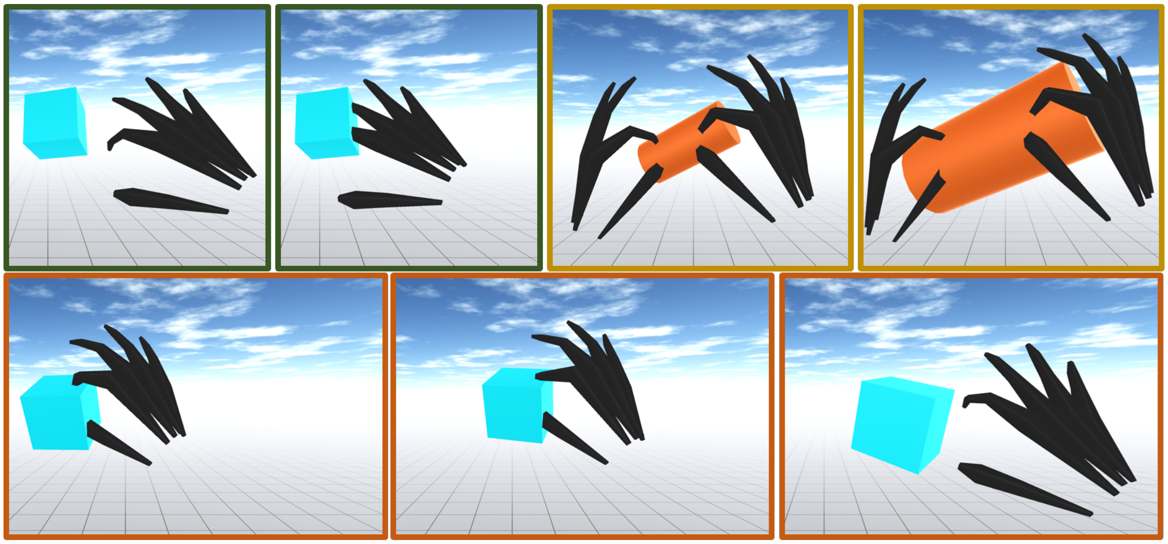

Figure 2. Movements and interaction gestures. The green-bordered images represent object-touching gestures, the yellow-bordered images represent object-scaling gestures, and the red-bordered images represent object-grasping gestures.

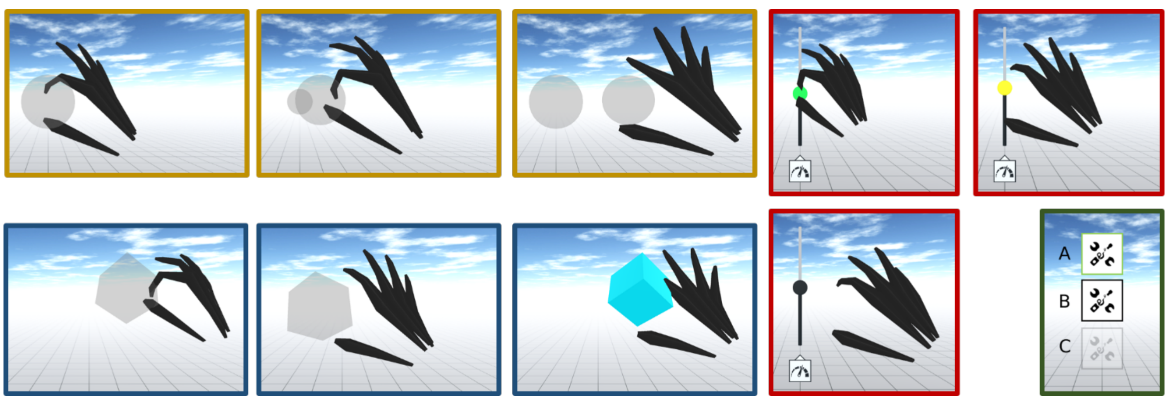

Figure 3. Summary of the basic components of the user interface. From left to right, the yellow-bordered images show the pinching gesture performed on an object, the object's movement by grasping, visualization of the next object, and the object release gesture. The blue-bordered images show the disabled obstacle, the gripped and moved obstacle, and the enabled obstacle after it is touched. Red-bordered images show interaction with the slider and the texture button. The green-bordered image shows the appearance of enabled and disabled buttons.

will be returned to initial point. It is also possible to delete objects through a trash feature. We used a trash mesh that catches the objects in a well-defined area, so that the user can delete an object by grasping and releasing it in the trash area. The visual feedback in this case is represented by diminishment of the deleted object when it is in the catching area.

Handled components Through the BoidVR simulation, the user can use and handle other types of virtual objects. A first object is the obstacle shown in Figure 4 (green borders), which represents an element that the agents will avoid. Each obstacle is deactivated by default and has a spherical interaction zone around it, which has a mesh size value as the radius. The user can interact with obstacles through the index-finger touching and grasping gestures: Index-finger touching

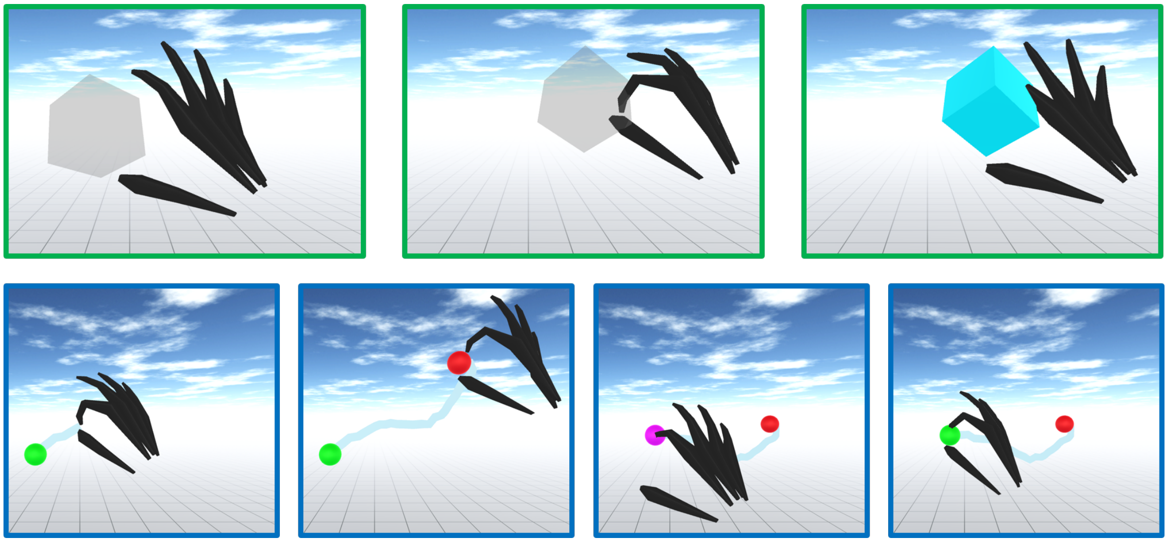

Figure 4. Handing other virtual objects. From left to right, the green-bordered images show a deactivated obstacle (opaque gray), an obstacle grabbed and moved in the simulation space, and an activated obstacle. The blue-bordered images show the start and end, movement, rotation, and selection of the path to follow.

allows the user to change the obstacle's state when the index-finger falls within the interaction zone for a defined time interval; grasping allows the user to rotate the object regardless its state. Furthermore, it is possible to resize the object and its interaction zone. Figure 4 (blue borders) shows our path-follow component, which represents a path that the agents can follow in groups until the destination is reached. The start and end of the path are represented by green and red spheres, respectively, while the path itself is shown as a string. The path is created by moving the pinching hand for as long as it takes to create the desired path. Each path can be moved and rotated through the grasping gesture at its initial point by moving or rotating the hand. In addition, path-follow can be activated or deactivated: If it is activated, then the initial point is purple, otherwise it remains green. Finally, the most important handled components are the special agents. These agents have a similar 3D shape to neutral agents, but they are different colors according to their role and behavior. A *leader agent* is followed by other agents; it has a random behavior because its trajectory is not conditioned by any other agent. The *threat agent* is an agent whose aim is to capture the other agents by performing a behavior that consists of following the closest neutral agent. The last special agent is the *victim agent*, which represents "food" for the neutral agents. The victim agent's behavior is to escape from the closest neutral agent. Each victim agent has a lifetime, at the end of which it disappears, simulating its death. The lifetime is shown by a counter, which decreases when ten boids fall within the victim's search area. Special agents do not respect the basics of collective behavior and do not interact with each other. Threat agents have a search area with a fixed radius, and each neutral agent searches around it for special agents within a certain distance equal to the radius of the neighborhood search. General behavior of the neutral agents is to follow the closest leader, avoid the closest threat, and follow and capture the closest victim. If an agent finds a leader and victim together in its search area, its priority will be to keep the victim by ignoring the leader.

BoidVR Features

A key BoidVR feature is control of the agents' behavior through the alignment, cohesion, and separation parameters. These parameters define the weights that characterize the three rules of the agents. In addition, the user can define the inner and outer radius of their FOV. As shown in Figure 5 (yellow border), when the palm of the left hand is in front of the user, an agent appears over the hand. The simulation can be handled

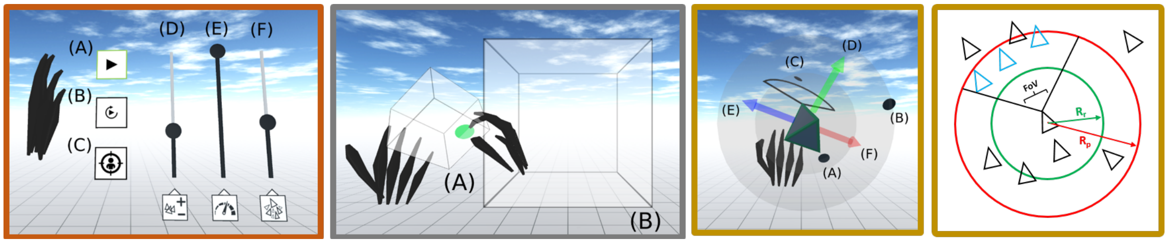

Figure 5. Control of agent behaviors. The orange border: icon (A) starts and stops the simulation and (B) restores the scene. Agents were brought toward the user through icon (C). The number of neutral agents number is controlled using the slider (D). Animation's speed and rules configuration are controlled through sliders (E) and (F). The gray border: containment menu values can be increased or decreased through the slider (A) and viewed in the spatial volume (B). The yellow borders: (A) and (B) buttons allow the user to enlarge and shrink the inner and outer spheres' radii, and button (C) allows the user to change the FOV; red (F), green (D), and blue (E) arrows are useful to modify the simulation weights.

by the user through manipulation of the agents' group, as can be seen in Figure 5 (orange border). The containment menu in Figure 5 (gray border), allows the user to define the volumetric spatial amount in which the simulation takes place.

Agents collectively avoid static or dynamic obstacles using a classical force field approach. In this approach, a discrete force field surrounds every object present in the environment, and upon approaching an obstacle, the forward vector of the individual is summed with the vectors of the force field, and the individual feels a growing opposing force on its path toward it. As shown in Figure 6 (light-green border), there are three types of obstacles with different shapes and colors. The path-follow feature can be activated and handled by the user through a menu as shown in Figure 6 (dark-green border). The path can be translated and rotated by the user if no menu item is selected. In addition, Figure 6 (purple border) shows the management of special agents. Through this feature, the user can add to the scene all desired special agents for each expected type through the grasping gesture.

Behavioral Functions

BoidVR allows one to define a number of neutral agents whose positions and orientations are chosen in a random way and included in a region near the user's virtual position, as shown at the top of Figure 7 (yellow border). In this phase, the agent's behavior is to follow simultaneously the alignment, cohesion, and separation rules and specified containment as well as avoid the virtual positions of user's body and hands. During a BoidVR session, the agents will try to join and adjust their trajectories to form a collective behavior, as shown in the second row of Figure 7 (yellow border). Based on the behavior observed by the user, it is possible to define or resize the containment space as the first operation (see Figure 7, last yellow border). As shown in Figure 7, when a path to follow is created and positioned as desired, the user can command the agents to follow this path. The agents follow the path and simultaneously respect the fundamental rules. The last three brown-bordered images in Figure 7 show the special agents; note that the group changes its behavior according to the special-agent type.

Through the agents' configurator, it is possible to change the main parameters of behaviors. In Figure 8, the last three (blue-bordered) images show changes to the weight of the separation and cohesion values representing the agent's group behavior while respecting the predominant rule's value. The last central (blue-bordered) image in Figure 8 shows the agents' behavior without an alignment rule. For example, in Figure 8 (pink border), there are different obstacles in the scene, and the agents try to avoid the pink obstacle by adjusting their trajectories.

User Study

BoidVR's usability and sentiments are evaluated through a user study experiment. This experiment aims to answer two questions:

Q1. How learnable and usable is the BoidVR

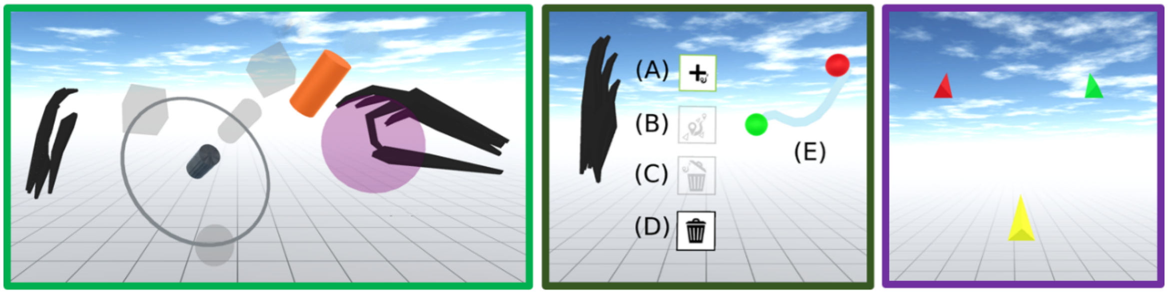

Figure 6. Manipulation of obstacles. The green border: the three types of obstacles which can be enabled (colored) or disabled (grayed). Dark-green border: (A) creates new path-follow, (B) tells the agents to follow the path, (C) removes a selected path-follow, and (D) removes all path-follows. Purple border: the red, yellow, and green tetrahedrons represent special agents: the threat, the leader, and the victim respectively. During the simulation, it is possible to grasp a special agent and either place it at a different point or delete it.

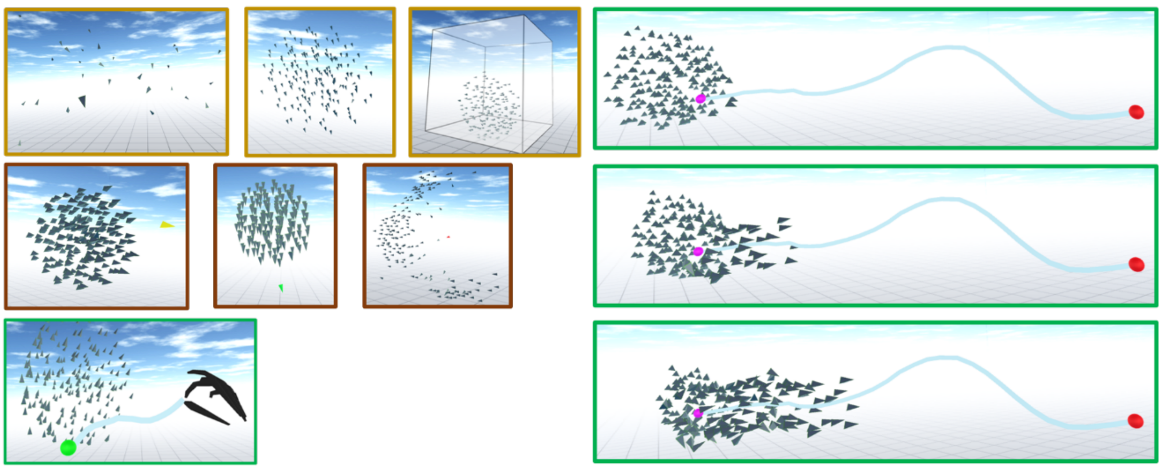

Figure 7. Controlling different types of agents. From left to right, the yellow-bordered images show the initial state of the simulation, the classic collective behavior with an increment of neutral agents, and the containment box. The brown-bordered images show the effect caused by the leader, victim, and threat special agents. The green-bordered images show the creation of a path to follow and the agent's movement over this path.

application?
Q2. What is the sentiment and feeling of the users interacting with the BoidVR application?
To answer these questions, we involved 18 computer science students at the University of Basilicata, Italy, by asking them to perform a single task using the BoidVR application. The user study was structured in four phases: (a) a preliminary survey, a questionnaire in which the users had to provide some demographic information such as gender, age, educational level, videogame and VR experience, use of glasses, *etc.*; (b) a training phase, where the users were free to use and interact with the virtual environment and boids, exploring the features provided from the BoidVR application; (c) a testing phase, in which the users had to perform the task of *path follow definition*, placing three obstacles around it and its subsequent activation; (d) a summary survey that consisted of answering a series of usability and emotional response questions.

Results

Based on the preliminary questionnaires, the majority of recruited participants had experience with video games (78%), and 56% had experience with VR. Additionally, the majority of respondents wore glasses (78%). We considered

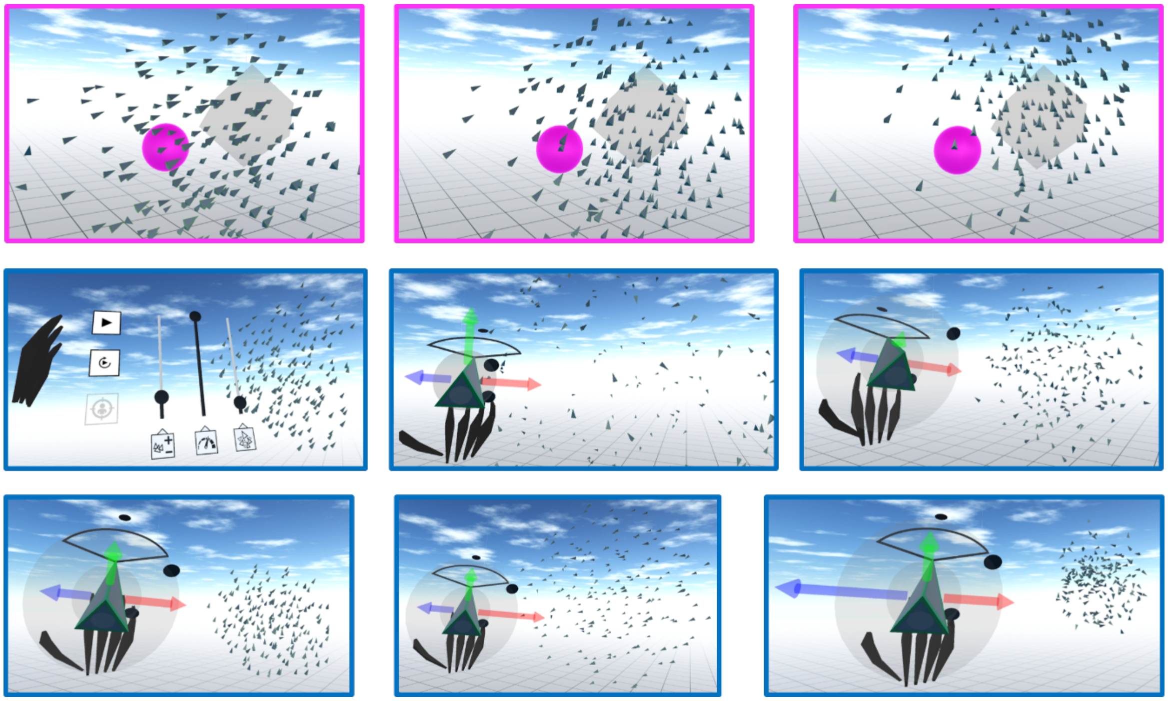

Figure 8. Simulated behaviors. The pink-bordered images show the obstacles and behaviors of agents with respect to active and deactivated obstacles. From left to right: the first blue-bordered row shows the sliders of the simulation. The second blue-bordered row shows the behavior of agents when the predominant value of the separation, alignment and the cohesion rules are selected

the videogame and VR experience as factors that might affect the learnability of the BoidVR and, wearing glasses, a factor that makes the head-mounted display (HMD) uncomfortable which can affect the usability of the tool.

Usability and Learnability We addressed question Q1 using the "System Usability Scale" (SUS) defined by [6], a ten-item attitude Likert scale with five responses options for respondents; from *Strongly Agree* to *Strongly Disagree* widely used to evaluate and quantify the perception of software and hardware usability. As explained by Lewis, J. et al. [7], a two-factor orthogonal structure could be extracted from SUS to independently measure the *Learnability* and *Usability* dimensions. The *Learnability* Score (LS) represents how easy the system to learn without needing to rely on technical support to use the software or learn large amount of information before starting. The *Usability* Score (US) refers to how easy it is for a user to make use of the software itself. The overall scores are divided into 5 intervals, each of which is assigned an adjective rating: (< 51) indicates *Awful*; ($51 - 67$) indicates *Poor*; ($> 67 - 68$) indicates *OK*; ($69 - 80$) indicates *Good*; (> 80) indicates *Excellent*.

The mean SUS score for our subjects was *SUS* = 69.0, which indicated good overall usability of the BoidVR application. Analyzing in detail the results based on the relative frequencies of the answers, as reported in Figure 9, we can see that 17% of participants considered the overall SUS score of BoidVR application *Excellent*, for 39% it was *Good*, for 6% it was *OK*. The remaining 39% of participants considered the usability *Poor* and no participants considered the usability *Awful*. Analyzing the learnability (LS) dimension independently, the mean *LS* score was LS = 64.6, which indicates *Poor* learnability. Based on the relative frequencies of the LS answers, 28% of the participants considered the BoidVR learnability to be *Excellent*, 22% considered it *Good*, 6% thought that it was *Poor*, and finally 44% felt it *Awful*. However, we have noticed a surprising negative correlation between BoidVR's

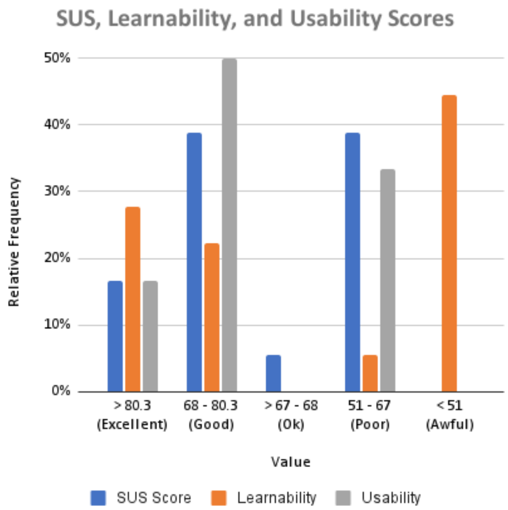

Figure 9. The relative frequencies of SUS Score in addition to the LS and US dimensions individually.

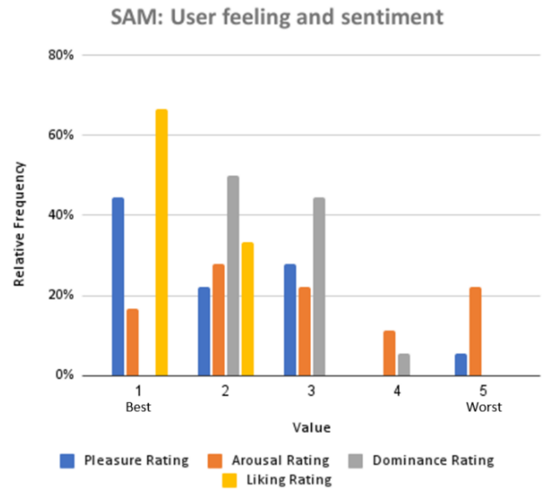

Figure 10. The relative frequencies of the Self-Assessment Manikin dimensions for user feeling and sentiment.

learnability and video game user experience. Indeed, considering only the users with poor or no experience, the propensity to learn the tool increased, with this group of users reaching *LS* = 70.8. On the other hand, users who had more experience with video games and perhaps greater self-assurance tended to not ask the support of a technical person or bother to learn the functionality of the tool and, consequently, proved *less* able to use the tool correctly (*LS* = 58.3). Likewise, if we further take into account users' experience with VR, users with VR experience still gave the system a low learnability score of LS = 58.3, whereas for those with less VR experience rated the learnability even higher LS = 73.2 (Good). Taken by itself, the usability (US) score reached 70.1, improving by one unit compared to the overall SUS score, maintaining its *Good* rating. It appears this dimension benefited from the non-conditioning of the LS, keeping the same overall SUS percentage of 17% *Excellent* and increasing *Good* ratings to 50%. But the US of the remaining participants (33%) still gave a *Poor* usability rating. We also noticed a correlation of the usability of BoidVR with wearing glasses. In particular, considering only the respondents that wear glasses (78%), the *US* = 69.0 maintained *Good* usability, while for the non-glasses wearing participants, the *US* = 74.2 was greatly improved. We infer that wearing glasses does make the HMDs somewhat cumbersome and consequently reduces the usability of the tool.

User Feeling and Sentiment We addressed question *Q2* the "Self-Assessment Manikin" (SAM) approach developed by [8], a questionnaire that evaluates persons' affective reaction to a stimulus in terms of *Pleasure*, *Arousal*, and *Dominance* rating. Recently, Kolestra, et al. [9] have added to the SAM a fourth *Liking* dimension. Each of the four SAM dimensions is characterized by a rating scale of 1 (best) to 5 (worst). The first range represents a *Pleasure* dimension; the second range represents an *Arousal* dimension; the third range represents a *Dominance* dimension, which indicates changes in control with changes in the size of SAM; finally, the fourth dimension represents the *Liking* dimension, which measures the participants' preferences as opposed to their feelings. To answer *Q2*, we evaluated the relative frequencies for each SAM dimension, see Figure 10. Considering the first *Pleasure* dimension, 44% of participants (the majority) were very happy with the BoidVR tool, 22% were pleased to use the tool, and 28% had neutral feelings, and only 6% were dissatisfied. The *Arousal* dimension suggests that 17% of the participants were excited during the experiment, 28% were awake and interested, and 22% were neutral, 11% got bored, while the remaining 22% became drowsy. Analyzing the *Dominance* dimension, 50% felt powerful and dominant, 44% were neutral, and only 6% felt powerless and without control. Finally, for the *Liking* dimension, 67% strongly liked BoidVR tool and 33% were inclined to like the tool, overall, an encouraging result.

Conclusions

BoidVR is a software tool in which the user can observe and interact with agents such as a flock of “boid” through a simulation from different fields of view by using Oculus Rift and Leap Motion devices to increase the immersive experience. We have designed a FIT composed of intuitive and natural gestures that are easy to learn, remember, and realize from an ergonomic point of view, and increase user involvement. BoidVR has, among its numerous characteristics, constancy and continuity of interactivity thanks to simulation and graphics components that react responsively to actions, commands, and user inputs. We evaluated the usability of the tool and user sentiment by using SUS and SAM questionnaires. The results show that the tool was sufficiently usable but needs to be explained by a technical person to be used to its full potential. However, usability is also influenced by users’ experience with video games and VR. Our SAM results suggest that the majority of the users were satisfied, awake, and felt powerful during the experiment. Finally, all the participants liked the tool. BoidVR is our first step toward the development of an application that will reuse the freehand interaction components developed here to simulate crowd behaviors in dangerous situations such as earthquakes, fires, and terrorist attacks.

REFERENCES

1. I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, “Effective leadership and decision-making in animal groups on the move,” *Nature*, vol. 433, no. 7025, p. 513, 2005.
2. C. W. Reynolds, “Flocks, herds and schools: A distributed behavioral model,” *SIGGRAPH Comput. Graph.*, vol. 21, no. 4, p. 25–34, aug 1987.
3. J. Ondřej, J. Pettré, A.-H. Olivier, and S. Donikian, “A synthetic-vision based steering approach for crowd simulation,” *ACM Trans. Graph.*, vol. 29, no. 4, Jul. 2010.
4. A.-H. Olivier, J. Bruneau, G. Cirio, and J. Pettré, “A virtual reality platform to study crowd behaviors,” *Transportation Research Procedia*, vol. 2, pp. 114 – 122, 2014.
5. G. Caggianese, L. Gallo, and P. Neroni, “The vive controllers vs. leap motion for interactions in virtual environments: a comparative evaluation,” in *International Conference on Intelligent Interactive Multimedia Systems and Services*. Springer, 2018, pp. 24–33.
6. J. Brooke *et al.*, “Sus-a quick and dirty usability scale,” *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.
7. J. R. Lewis and J. Sauro, “The factor structure of the system usability scale,” in *International conference on human centered design*. Springer, 2009, pp. 94–103.
8. M. M. Bradley and P. J. Lang, “Measuring emotion: the self-assessment manikin and the semantic differential,” *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.
9. S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, “Deap: A database for emotion analysis ;using physiological signals,” *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

Dr. Nicola Capece is Postdoctoral Researcher in the Computer Graphics Lab at the University of Basilicata. Contact him at nicola.capece@unibas.it.

Dr. Ugo Erra is an Assistant Professor at University of Basilicata in Potenza, Italy, where he is the founder and the head of the Computer Graphics Lab. Contact him at ugo.erra@unibas.it.

Gilda Manfredi is a PHD Student at the University of Basilicata and University of Salento. Her research interests include Deep Learning and Computer Graphics. Contact her at gilda.manfredi@unibas.it.

Rocco Di Bello is graduated in Computer Science from the University of Basilicata.