

# Neuro-TARDIS: Visualization of a Spiking Neural Network in Virtual Reality

Fabio Cassisa  
970826-T293  
cassisa@kth.se

Harsh Kakroo  
970830-7799  
kakroo@kth.se

Alfred Berg  
900922-3158  
alfrberg@kth.se

Alessandro Iop  
960916-3234  
aiop@kth.se

**Abstract**—In Computational Neuroscience, researchers and academics have so far been relying on 2D representations of neural activity. Representing simulations of artificial neural networks in virtual reality may aid explorative analysis but this area is yet to receive substantial attention. In the presented work, we aimed at providing additional modalities and interactions to an already existing application where short-term memory formation in a Spiking Neural Network is visualized in virtual reality. The main goal was to provide spatial audio cues to help neuroscientists track the propagation of neural signals through the network and to deliver improved interactions. Evaluating the interface with subject matter experts provided support for the notion that VR representations can assist analysis.

**Index Terms**—Information Visualization; VR; Spatial Audio; Spiking Neural Networks; Multimodal Interaction; Unity; SteamVR

## I. INTRODUCTION

At the forefront of Computational Neuroscience, a major factor in making the large amount of data that is being processed accessible – i.e. legible and understandable – is providing an effective visualization of the information which is also interactive for the researchers themselves. Several applications of Information Visualization in this specific field of Neuroscience were focused in recent years on enabling their users to not only visualize the Neural Networks and their activity, but also explore them and interact with them with a more immersive approach [1], [2], [4], [14], [18]. In order to identify patterns of activation in Spiking Neural Networks (SNNs) with greater ease, researchers and/or academics may

want to be able to identify by sight what kind of neurons are being activated over time, when exactly they are being activated and where they are placed in the hierarchical sub-division of the networks into semantic groups. In particular, two types of neurons and three hierarchical structures are most often considered when analyzing such patterns: Pyramid cells, which display an excitatory kind of behaviour, and Basket cells, whose behaviour is of inhibitory kind, are the former; the single neurons, the so-called Minicolumns (MCs) the Pyramids are grouped in and the more comprehensive Hypercolumns (HCs) are the latter.

Virtual Reality (VR), along with a 3D representation of data structures in space, can provide an important support to the immersiveness of the visualization for researchers and academics when analyzing and interpreting such information, as supported by Leigh et al. [13] and by Kasabov et al. [11], [12]. The previous version of the project presented in this paper, developed by Vincent Wong in 2018<sup>1</sup>, attempted at representing the aforementioned types of neuron and hierarchical relationships through an implementation of a navigation system within a Virtual Environment (VE) using the data retrieved from an existing simulation of an SNN. In the 2017 paper by Fiebig and Lansner, which the previous as well as the current versions of the project are based on, entities such as Pyramids, Baskets, MCs and

<sup>1</sup><http://www.cawards.se/project/c-awards-2018-vr-visualization-of-simulated-neural-nets>

HCs lie at the basis of their claims about how the working memory is formed within a human brain [9] (see § II for a more exhaustive insight): 30 Pyramid cells form a single MC, while twelve MCs, together with 24 Basket cells, constitute a single one of the 16 total HCs. In order to keep track of their activation patterns throughout the simulation, the authors have so far been relying on the two-dimensional rasterplot that comes as output of the simulation itself. It maps the spikes associated of the different individual neurons in the network ( $y$  axis) with their timestamp ( $x$  axis), as shown in Figure 1, while associating a different color to each of the 12 words given as input. Three main phases in the simulation can be distinguished: (1) the noise phase, in which neurons are activated randomly; (2) the stimulation phase, where the words are given as input; (3) the consolidation phase, corresponding to the period of time where the words are being remembered – i.e. memorized. The bottom part of the rasterplot, depicted in grey, specifically maps the activity of the Basket cells over time. Despite the clear chromatic distinction between the stimuli, mapping all the information that can be useful for researchers and academics in a two-dimensional image can prevent it from being truly legible, the users from easily distinguish the different neurons as well as their groupings, the patterns from being tracked. The project’s main goals in this regard are: bringing the visual structures and their animations in a 3D VR environment, which would greatly improve the user experience; introduce audio as an additional modality; remap the controls for an improved interaction with the system.

This paper is structured as follows:

- In § II (*Background*) previous research in Information Visualizaton, Virtual Reality and multimodal interfaces is presented, along with more details about Vincent Wong’s version of the project.
- In § III (*Design and Development*) the different aspects – i.e. modalities – that constitute our work are described, namely

*Interaction and immersiveness, Audio and Vision*. For each, the most relevant design choices, technical details about the implementation, and/or challenges faced are presented.

- In § IV (*Results and Evaluation*) the focus is on what we were able to achieve by the end of the project, given the time, resources and skills at our disposal. The evaluation of the application by Subject Matter Experts (SMEs) as well as Media Technology Master’s students is also introduced, as their feedback was sometime divergent.
- Finally, in § V (*Discussion and Future Work*) possible improvements for the current version of the project, along different directions, are hypothesized, in accordance to the results obtained and feedback received in the evaluation phase. A brief digression on cyber sickness as a design concern is also included.

The most relevant technologies employed during the development of the VR application were Unity (v. 2019.2.4 and 2019.2.5)<sup>2</sup>, a HTC VIVE headset<sup>3</sup>, SteamVR (v. 1.9.16)<sup>4</sup>.

## II. BACKGROUND

Three-dimensional representations of data was seen as a promising area of research during the early days of information visualization. Humans naturally perceive the world in three dimensions so there was a widespread belief that adding depth to visually complex representations would presumably only benefit analysis. The advent of modern desktop computers made 3D representations viable to a large audience, but interest faded after several studies failed to find support for the hypothesised utility of a third dimension and 2D visualisations of abstract data have remained the dominant paradigm ever since [6]. However, recent advances in virtual reality and similar technologies have sparked new interest in the potential

<sup>2</sup><https://unity.com>

<sup>3</sup><https://www.vive.com>

<sup>4</sup><https://store.steampowered.com/app/250820/SteamVR>

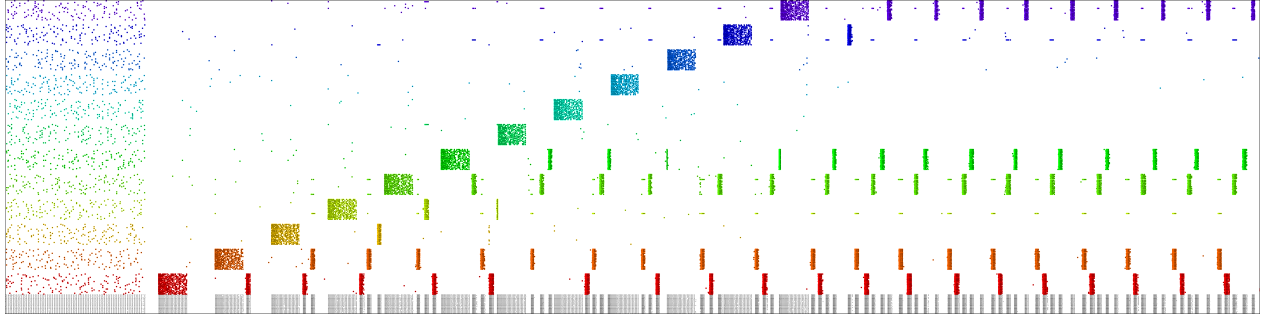


Fig. 1. Rasterplot of the neural activations in the SNN over time.

of 3D visualizations. The binocular presentation and motion tracking afforded by modern head mounted displays (HMDs), like the *HTC VIVE*, enables natural depth processing as well as embodied interaction, thus giving the user a radically different experience compared to the same 3D visualization if it was displayed on a flat surface. As a result, information visualization researchers have once again begun looking into the potential benefits of 3D visualizations. Marriott et al. [16] reviewed several studies that compare the relative effectiveness of 2D and 3D visualizations and found that the binocular 3D vision and embodied interactions associated with modern HMDs seemed to unlock many of the previously hypothesised advantages of 3D visualisations. Holistic presentations of multi-dimensional and spatial data (notably 3D representation of networks) seem to be particularly promising areas.

While the old version of 3D representations (displayed on monoscopic, flat surfaces) quickly fell out of favor within the information visualization community, they were widely adopted within the life sciences where the third dimension enabled natural mappings of biological structures [16]. To this day, much of the interest for 3D visualizations stems from the life sciences and several studies have investigated the use VR to aid scientific analysis. For example, Marai et al. [15] reported several instances where analysis of brain connectivity was aided by 3D visualisations in VR. Nowke et al. [17] used VR to present researchers with

synchronised output of 2D and 3D data views. The main benefit being that the user can switch between macro and micro scale views quickly which greatly assists workflow. Informal studies about the utility of VR have produced positive results from scientists but few if any quantitative studies have been performed to date.

Three related concepts seem to permeate every study on information visualization in VR: embodiment, immersion, and interaction. Immersion describes a system’s ability to make the user feel present in the virtual world. A strong feeling of being present increases engagement with the material, which has both direct (e.g. effort levels) and indirect (e.g. memory) benefits [5]. Embodiment is the sense of having one’s own physical body translated into the virtual world which benefits immersion as well as interaction. The possibility of interacting with data in a natural, physical way is one of the most interesting aspects of information visualization in VR and the simple addition of head-tracking has already proven to increase the utility of such visualizations [16]. This more “natural” way of interacting with data also presents interesting ways to integrate non-visual modalities, which we view as perhaps the most exciting possibility afforded by modern VR HMDs. Given that natural interaction with data seems to be one of the main drivers for information visualization in VR, multimodal interfaces are indispensable for the future utility of such systems since our cognition has evolved

to process multi-sensory information. However, relatively few studies and applications seem to have explored the benefits of multimodal interaction, and most of these are concerned with the integration of vision and haptics. Spatial audio could, if implemented correctly, provide valuable redundant and complementary output to aid spatial processing as well as identification and immersion. Sonification has been used in scientific analysis for many decades, but there seems to have been few attempts to apply sonification in VR visualizations [20].

Concerning the version of the project we inherited from Vincent Wong, some of the visual structures and interaction modalities that we ended up including in our work had already been implemented, but were not as “polished” as we would have liked, therefore we decided to improve most of them in addition to implementing new ones. To make a few relevant examples, at the beginning there was no sound associated to the activation of the neurons at all, Basket and Pyramid cells were placed randomly in the HCs (and MCs) – which could have wrongly suggested to the user that their position reflected the real one in the SNN, it was impossible to keep track of time during the animation since there was no 2D rasterplot to be accessed, and there was only one view mode – the “transparent” one, where the Pyramids can be seen within the MCs, which are in turn almost completely transparent. The last feature, in particular, was problematic when the whole environment was being visualized from far away, because it was almost impossible to distinguish between different neurons (and therefore track patterns) when they were being activated.

### III. DESIGN AND DEVELOPMENT

In this section, the most relevant aspects of the development process for our project are presented, divided by general topic – i.e. by modality. Despite being currently distinguished as visual, aural, and that related to the interaction with the system, such modalities are strictly related to each other: their functions

sometimes overlap; they have at times redundant impacts on the perception of the information and at times complementary ones; they influence, support and enhance each other. How – and most importantly *why* – their related features were introduced in the project is also introduced and contextualized in this section, while the more reflective aspects that concern them will be introduced in § IV.

In the context of the project, potential cybersickness of the users was a concern at all times, considering that the application is meant to be used extensively and the animation can last a long time if the playback speed is relatively low. We focused in particular on keeping a sufficient framerate in the application in order to prevent any kind of discomfort, by always employing technical solutions that would not impact the number of frames displayed per second, given also the fact that in the starting version of the project the framerate varied between 35 and 60 fps. More about cybersickness will be presented in § V.

#### A. Interaction and immersiveness

Much focus was put on embodied interaction during the design and development of the VR visualization, as it enhances the expressiveness and impact of the information that is being represented as well as increase the speed at which it is processed. Improving interaction and movement within the VE was therefore a priority, in particular: (1) correctly mapping the user’s physical movement to the virtual one to allow them to cover the entire surface of the network representation; (2) providing a fairly intuitive and quick way of controlling their own position, the playback of the animation and some of the aspects of the visualization itself through the controllers – whose control mapping is displayed in Figure 2; (3) associating acoustic stimuli to events that happen over time in specific locations within the VE (a design aspect that will be discussed in § III-B).

Concerning the playback of the animation, which represents the activation of the different Pyramid/Basket cells over time, in addition to



Fig. 2. Mapping of the controls onto the VIVE controllers.

keeping the play/pause and fast-forward (i.e. increasing the playback speed) functions, our intention was to develop a “rewind” control as well as a “stop” button. Given the need to implement an inverted animation in the former case, and the difficulties in manipulating the data structures employed in the previous version of the project in both cases, we ultimately settled with just adding a “speed decrement” function which is currently placed on the trackpad of the left controller along with the already-existing controls.

Redundancy of input modalities plays, in our opinion, an important role in ensuring usability across different usage context for the application. We have implemented, for example, some of the controls that can be activated through the controllers on the keyboard as well: movement in space can happen through the classic WASD key combination but also by walking around the room and touching the trackpad on the right controller; for changing from the transparent view mode, where Pyramid cells are visible, to the opaque one, where the Minicolumns are visible, either the  $\odot$  key or the button above the trackpad on the right controller can be used

(more on this in § III-C).

Similarly to graphic attributes and spatial movement (which is somewhat limited when used a VR set like the VIVE, considering that it is cabled to the workstation), the way visual information is perceived by the user is greatly influenced by the scale of the 3D objects that are visible in the VE. We therefore focused on providing a mean for looking the whole representation of the network from a more distant perspective than being directly *in* it, by improving the previously-existing zoom in/out functions along with the vertical movement of the user. The first allows them to keep perceiving the spatial audio stimuli while the animation is playing, while the second is subjected to a linear rolloff that prevents cluttering of audios at a certain distance from the 3D objects; both are controlled through the trackpad on the left controller.

Additionally, as mentioned previously, when designing the interaction with the VE, we chose to prevent the navigation to make the user feel sick at any point, by carefully choosing the values for parameters like movement speed and rotation of the plane where the 3D objects lie,

while keeping the full range of movements at the user’s disposal. We decided, for instance, to rotate the aforementioned plane instead of the camera itself, in order to decrease the discrepancy between the actual movement of the head and what was being seen in the headset.

### *B. Audio*

Hearing is arguably the modality that stands to gain the most from representing SNN-simulations in VR. This is because the auditory system has the ability to identify and locate sound sources in all directions, whereas the visual system is limited to activity within the visual field. Traditional screen and speaker configurations will therefore support most visual tasks just as well as VR headsets but severely impair the spatial aspects of auditory output - this is also true for surround systems since they lack head-tracking capabilities. Sound can, of course, provide useful redundant as well as complementary information in traditional interfaces, e.g. by linking sounds to different colors or by binding activation frequencies to the relative volume of sound cues, but these and many similar tasks could be presented and processed visually if necessary. Therefore, spatial information is perhaps the most interesting aspect of introducing sound in this project and spatial audio was one of the primary goals in development.

Initially, the idea was to turn every Pyramid and Basket into a sound source that would emit a short noise every time the object was activated, that is, every time the artificial neuron fired. In this way, the sounds would complement the visual output by informing the user about activity outside of their visual focus or activity occluded by other 3D structures. It was likely that this solution would result in too many sounds playing simultaneously, and that it would be preferable to aggregate the sounds in some way, but the actual experience was virtually impossible for us to predict before trying it in VR so we decided to cast a wide net and narrow the number of audible objects as needed. Our first task was to identify parts

of the code that activated the visual animations and program these to coincide with the playback of an audio clip. Having figured this out, we had to find an audio software development kit for Unity capable of performing a head-related transfer function, which simulates how sound waves interact with human ears to give a fully binaural sense of direction. We initially used Resonance Audio (v. 1.2.1)<sup>5</sup> but later switched to Steam Audio (v. 2.0-beta.17)<sup>6</sup> for reasons explained below. Unfortunately, using every Pyramid as an audio source caused the framerate to drop to unplayable levels so we decreased the number of sound sources by moving the audio component from the Pyramids to the MCs. In this version, the MCs would emit a sound whenever a nested Pyramid was activated which meant that the number of sounds remained the same while the framerate reached acceptable levels again. Of course, at larger scales, this creates a discrepancy between the aural and visual location of the sound source because sounds always originate from the center of the MC, but these effects were hardly noticeable (see § V for further detail).

Knowing what we could expect to achieve in terms of spatial information left us with the secondary task of conveying identity information through the sound cues. In theory, a listener will be able to identify sounds among a huge number of competing sources as long as the sounds are salient enough [3]. This meant that we could experiment with large numbers of simultaneous sounds, but that we would have to find a balance between the number of different audio sources and properties like pitch, timbre, duration, roll-off, and simulation speed. Naturally, the audio sources have to be attached to the most meaningful structures possible and since individual Pyramids put too much demand on the CPU, we settled for a solution where the MCs emit sounds according to their associated color, i.e. 12 sounds in total. Our initial idea was to find sounds that varied

<sup>5</sup><https://resonance-audio.github.io/resonance-audio/>

<sup>6</sup><https://valvesoftware.github.io/steam-audio/>

considerably in pitch and timbre so that each sound would be distinct and easy to identify. One solution was to use single notes from 12 different instruments in an orchestra, which would have the additional benefit of presenting the user with familiar stimuli. For example, a user might know that violin sounds correspond to activity in blue MCs, which facilitates processing and shortens the time they need for learning to associate colors and sounds. We soon discovered that differences in timbre made some sounds much more salient than others and, having tried compression and normalizing without positive results, we decided to control for these differences by synthesizing our own sounds.

At this stage, we also experimented with the duration of the sounds. Since multiple identical sounds triggered from the same audio source had an amplifying effect, we had an idea about using longer sounds to generate a continuous soundscape; increased activity would highlight certain areas whereas decreased activity would subtract the volume from other areas, resulting in an auditory equivalent of a landscape where peaks and valleys represented fluctuations in activity. To create smooth transitions, we applied fade-in and fade-out effects to the audio clips but performance-related limitations on the number of simultaneously active audio sources caused abrupt deactivation of sounds which was perceived as clicking noises. Our best option was therefore to use as short audio clips as possible but ensure that they were long enough for the pitch to be processed. Having found a sound with a stable waveform, 12 sounds of increasing pitch were produced with Ableton Live (v. 10.0.1)<sup>7</sup> and edited into short stereo clips of 0.36 seconds in Audacity (v. 2.3.2).<sup>8</sup> We also applied fade effects to mitigate possible clicking noises described above. With these factors set, the utility of our project would be determined by calibrating the volume roll-off distance to not overload the users auditory

bandwidth and suit their task at hand.

As mentioned above, we dropped Resonance Audio in favor of Steam Audio. We had two reasons for doing so: first, the framerate increased; second, the scale of our virtual environment was always so small that the simulated dissipation of sound waves - the loss of energy that decreases their amplitude over range in nature - never came into effect. Steam Audio let us use Unity's artificial range limiters to control volume roll-off. We could then fine-tune the audible range of sound cues to maximize the user's chances of locating the sound source. Since we would not be able to provide our users with advanced audio controls, we created two presets: one for exploring the network at the neuron level and one for exploring from an overview position. When inside the network, the scale was large and the volume decreased linearly between 0 and 16 meters; when outside the network, we wanted the scale of the VE to be small so that the user could explore the network through moving physically through space. We therefore lowered the range to 6 meters. In this way, the user would be able to hear activity from the whole network but increase the saliency of sounds in certain areas by "listening-in".

### C. Vision

As mentioned in § III-A, two different view modes were considered during the development of the application, which are shown in Figure 3: one where Pyramid cells can be seen and the MCs are almost completely transparent, so each one of former had an independent animation when it was activated, and one where the MCs are completely opaque and change color instead of the Pyramids within them. The reason of this choice was based on the team's best guesses at hypothesizing what neuroscientists could be possibly looking for in the visualization; it provides a useful mean for aggregating the 3D elements when the distance makes it impossible to distinguish them, or keeping track of certain patterns of activation with greater ease throughout the environment.

<sup>7</sup><https://www.ableton.com/en/shop/live/>

<sup>8</sup><https://www.audacityteam.org/>

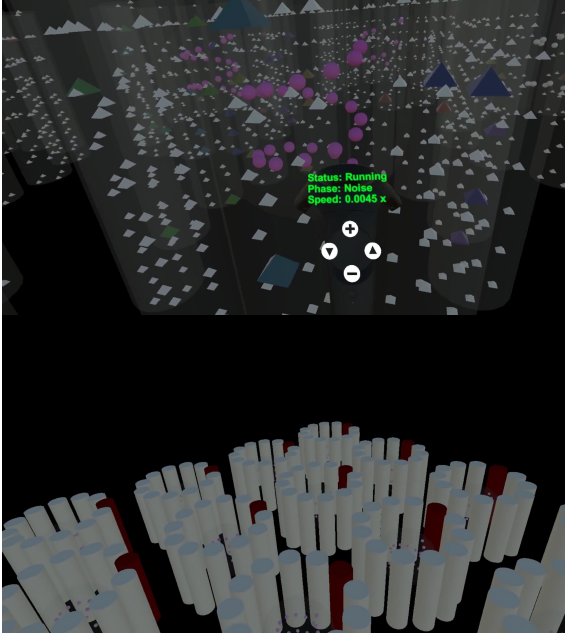


Fig. 3. Screenshots of the VR experience from a close-up perspective with the transparent view mode toggled (above) and from a distant perspective with the opaque view mode toggled (below).

Another hypothesis that we made concerned the positions of the different kinds of neurons in the previous version of the project, which were randomized but seemed like they could bear some kind of meaning. We opted for placing them in a sorted manner to clarify that their relative position does not really reflect their actual collocation within the neural network, but rather their semantic grouping into MCs and HCs. In particular, the Pyramid cells were distributed along the height of each MC in squares of four, and the Basket cells were organised in a spiral-like shape, in order to evenly displace them without cluttering the volume of their own container (Bigbasket).

A relevant concern of ours was allowing the users to keep track of the passing time while the animation was being played, so to identify the different phases of the Neural Network simulation with more ease. In a previous version of the project, different from the one we started our work from, a 3D canvas representing the 2D graph of the simulation had been placed in a fixed position within the VE, precisely to

one side of the HCs, forcing the user to turn away from their focus of attention to read it most of the time. In the current version of the application, it is instead placed slightly above the left controller and can be toggled on and off according to the user’s need. This gives the graph usability and portability at the same time: it can be brought to the user’s attention just by raising one hand, without having to move away from or outside of the current visualization of the environment that they are exploring. Additionally, we tried to implement a sort of “pointer” (a simple vertical bar placed on top of the canvas with the 2D graph) that would allow an easier tracking of the point in time the visualization corresponded to while the animation was running, but due to lack of time along with technical challenges the idea was not developed further.

#### IV. RESULTS AND EVALUATION

Having the opportunity to have an evaluation session with SMEs, we decided to deploy three different playable demos with distinctive changes in order to test the effectiveness of each implementation. The neuroscientists called for this session were Professors Anders Lansner and Pawel Herman, from the University of Stockholm and at KTH Royal Institute of Technology. We also tested this VR application with a few Media Technology Master’s students to have a different perspective and see what they could have thought. The only difference concerned the audio: the first demo was a large-scale VE with long volume roll-off, the second demo had a small-scale VE with short volume roll-off and the third had no sounds at all. However, the muted demo was never used due to lack of sufficient time.

We conducted an informal heuristic evaluation asking them specific questions whilst they could talk out loud regarding precisely what they found relevant and what not. The test consisted of first training session in which, after having explained the controls, they had to explore the environment without knowing about it. They were asked to underline what they liked





Fig. 4. Anders Lansner tests the VR application with the HTC VIVE.

and disliked. By leaving them space to explore the virtual world without guiding the evaluation strictly on what they should have looked for, we were able to grasp what were their primary needs (see Figure 4). Then, we tested them by giving more in-depth and concise tasks to perform to have their opinion on the improvements we made. One of the first things we wanted to assure was how much effective the sonification applied to all the spiking neurons in the VE could have been.

The last aspect we pointed out was the overall navigation composed by the new controls we provided and the different views we thought they would have found useful. After having run both the evaluations to our two audiences we got different feedback. This happened primarily because they had different backgrounds, but we still considered everything to be relevant for future improvements. Generally, the 3D audio was appreciated because it gave them a good idea of what was going on while they were focusing on another neuron cluster. It seemed very natural and clear for all of our users. Anders Lansner pointed out it would have been even more helpful distinguish each pattern with a definite sound, but he was satisfied by that. Another feedback on the way to map audio surfaced during the discussion, and this was about sonifying the synchronous activity of an individual cell. This could aid pattern recognition. Pawel Herman found sounds engaging and very pleasing, especially while inside a cell investigating them one by one. He assured he

could "feel more" thanks to the audio feedback. From the session the SMEs, we can conclude that spatial audio complemented the visual representation, but that graphical and navigational aspects needed improving. A common criticism was the difficulty to understand, at first glance, how to control the interface; this meant that our controllers' displacement could have been designed better, as it was functional but not accessible for a novice. Even when they knew what to look for, they felt quite disoriented and lost some time due to the lack of orientation. Regarding color mapping, a suggestion from a neuroscience student was to use either white or grey color for the non-firing neurons and lighten them up when they start firing. Our hypothesis on having multiple views turned out to be appropriate since they clarified that there is not a better view when it comes to analyzing a neural network, sometimes a general view is needed and sometimes it is best to have a detailed one. When we interviewed students we understood that there was a general confusion with movement inputs, even though they liked the audio and they got used to the controls faster than the scientists.

## V. DISCUSSION AND FUTURE WORK

The overall results we got from the project pointed us to the fact that the sonification of the Virtual Reality representation for the simulation of working memory works very well to explain the 2D spike graph in a meaningful way. In other words, with the introduction of multi-modal qualities, more information was being communicated. Having said that, there were hurdles that we came across during the time we worked on this project, some we were able to overcome, while for others, *time* was a resource we just did not have. But these said hurdles are definitely something we will look more into, in the days to come.

Having said that, there were some issues that were pointed out to us during the evaluation, e.g. mapping problems on the controller and issues with the movement speed of the user.

These were some of the more realistic dilemmas whose solutions, given the time-frame, were more achievable for us. Given our limited neuroscientific background knowledge, it was difficult to predict the playback speeds and view modes that would generate the most useful data and we were essentially forced to rely on qualified guesswork in this regard. This was, of course, a huge issue when deciding everything from the VE scale to the number of sounds since we lacked the time to provide controls that let the user adjust relevant settings during runtime. As mentioned in § III-C, one important feature that we wanted to implement but could not completely, was a *vertical bar* that moves across the 2D graph, the idea here was to let the user know where they are in time during the simulation.

#### A. Welfare of the User

Another important issue that had to be addressed was the health and welfare of the user that may interact with our application, this did bring up some interesting dilemmas regarding what movement features we could include in our Virtual Reality application, particularly the plane rotation feature, as we found out that some individuals might be more susceptible to *flicker vertigo*, which means that when they are exposed to flickering lights in the range of 8 to 12 Hz, it could ultimately lead to seizures [22].

Although there are various suggestions about what causes cybersickness, to date there have been no definitive proofs about cybersickness. Furthermore, it became quite clear that this is exactly the reason why cybersickness poses a serious threat to the usability of Virtual Environments. For this particular reason, we wanted to improve the rotation feature in our application, as this feature along with the other multimodal qualities, in general, may amplify the effects of cybersickness, but again due to time constraints, we could not completely do that.

#### B. Presentation and Interaction

*Spatially Locating* the user includes providing an indicator to the user which lets the user know their location with respect to the rest of the Virtual Environment. To help with this dilemma we plan to add landmarks around the simulation space for the user so that the user has an easier time locating themselves.

Another hurdle we faced but could not completely get over was *Path Planning*; the idea here was that the user could select the path of the neuron to stop, and to perhaps understand the patterns better. Furthermore, the user would be able to preview the camera motion in advance before committing themselves in the larger immersive Virtual Environment. Another important feature of Path Planning that we plan to add in the future is allowing the user to jump at any particular time of the simulation, which means that the user can go through a moment in the simulation as many times as they want, which is quite important from a research point of view, as pointed out to us during the evaluation.

Another key aspect of improving the visualization features involved in this project, we had to first understand how the Virtual Reality headset, in this case HTC VIVE, works. In other words, how the input data is transmitted between the user and the computer and how that relates to the bandwidth of interaction involved, with the primary goal of improving the User Experience. After some research, it was quite clear that the visual data communicated to the user usually corresponded to the portion of the virtual space viewed, again, from the perspective of the user, but as we also had the spatial audio as a key modality of our project that might not be completely accurate. Concerning the actual bandwidth and data communicated between the computer and each user, designated virtual spaces may have to be defined within the portion of the Virtual Environment data being communicated to the user. There are various ways in which data can be communicated in case of sound and hence different ways in

which auditory processing happens. In the case of our project, sounds were assigned to different points of origin within the Virtual Environment, the MCs, and a sound control was used to then send the data to the User for simulating the origin of the sound within the virtual space [8].

This actually brings us to something quite interesting, about the involved sonification of the MCs, which contains the pyramid cells. When the individual pyramid cells were being activated, there was the perception of spatial sound for the users. In other words, as the individual pyramids played their spike animation, it appeared to the user that the Pyramid cells themselves were making the spatial sounds, instead of the MCs, but in actuality it was the other way around. We looked more into this phenomenon to understand why the users were perceiving the spatial sounds originating from a different source and how big a role do visuals play in this. Related research suggests that the state of the listener, from expectations about a scene's content to the level of analysis a listener undertakes (in this case, it would mean listening to the MCs), would influence the amount of the perceived content of a graphical object [7]. It can also be said that expectation plays an important role in affecting spatial processing; an audible sound in a mixture with others is not always allocated between the objects perceived in a scene, and can contribute either to multiple objects [10] or to no object [21].

A considerable improvement would be to provide the user with options for filtering sounds depending on the structures they are interested in. We would also like to provide toggles for preset modes, where the user easily can jump between different scales and audio properties being adjusted automatically to best support the current interaction. The ideal scenario would be an application where the user would be able to remain inside the simulation while making changes to the presentation or even accessing external resources. Being able to interact with the represented structures to control activation or change their properties

would allow users to personalize the types of patterns they are interested in.

Viewing of *alternate representations* was something that we were really careful about. As mentioned previously in this work, a “transparent” and an “opaque” view mode were implemented, while a third alternative, comprising of neurons’ spiking potential (i.e. voltage) was considered as an option, but for now it is on our agenda for the future. Such an addition was suggested by Professors Anders Lansner and Pawel Herman, who evaluated the project, as current research on the imaging of spiking potential in neurons is fairly new and still ongoing. Albeit, in reality, we have barely scratched the surface by understanding how voltage in neurons actually work, current research tells us that the actual number of ions that enter or exit the neuron’s membrane is small (less than 105 of the total ions in the cell), but these ions have a large effect on the electric field of the membrane, even briefly reversing its polarity [19]. For now, it has become clear to us that this is a direction we can move in, to further elaborate on the project in the future.

## VI. MULTIMODAL ANALYSIS

Vision and hearing seem to combine in a nice way for the tasks we aim to support with this application, providing both redundant and complementary output for the user and serving to increase immersion. Haptic feedback was considered at the start of the project but put on hold as soon as we realized we would not have time to implement more embodied interaction with the VE, like grabbing and pulling. With these types of interaction, haptic feedback for physically interacting with the virtual structures would definitely be something we would like to add. Our prediction is that this would serve to increase immersion and offer both redundant and complementary advantages, eg. feeling and seeing that a Pyramid was touched or being given haptic feedback that a structure was selected even when looking away before actually “touching” it. We would also like to expand on the proven relationship between vision and

hearing, eg. having spiking intensity correlate with both increased volume and color brightness.

#### ACKNOWLEDGMENT

We would like to thank our project supervisor, Mario Romero, for guiding us throughout the project, giving us practical directions and providing us with useful insights on how to design and evaluate multimodal interfaces. We would also like to thank Vincent Wong for being always available to give us tips and help, especially in the first phases of the project, as well as useful feedback during its evaluation. Last, but definitely not least, thanks to Anders Lansner and Pawel Herman for the valuable critique, suggestions and validations on our work; to Björn Thuresson for granting us access to the laboratory during the weekend; and to Ingemar Markström for the technical support with the HTC VIVE during the development of the application.

#### VII. VIDEO

The project video is available here: <https://vimeo.com/386795142>

#### REFERENCES

- [1] AGUERA, P.-E., JERBI, K., CACLIN, A., AND BERTRAND, O. Elan: a software package for analysis and visualization of meg, eeg, and lfp signals. *Computational intelligence and neuroscience 2011* (2011).
- [2] ANDERSON, E. W., POTTER, K. C., MATZEN, L. E., SHEPHERD, J. F., PRESTON, G. A., AND SILVA, C. T. A user study of visualization effectiveness using eeg and cognitive load. In *Computer graphics forum* (2011), vol. 30, Wiley Online Library, pp. 791–800.
- [3] BREGMAN, A. S. *Auditory scene analysis: The perceptual organization of sound*. MIT press, 1994.
- [4] BRUCKNER, S., SOLTESZOVA, V., GROLLER, E., HLADUVKA, J., BUHLER, K., JAI, Y. Y., AND DICKSON, B. J. Braingazer-visual queries for neurobiology research. *IEEE transactions on visualization and computer graphics* 15, 6 (2009), 1497–1504.
- [5] BÜSCHEL, W., CHEN, J., DACHSELT, R., DRUCKER, S., DWYER, T., GÖRG, C., ISENBERG, T., KERREN, A., NORTH, C., AND STUERZLINGER, W. Interaction for immersive analytics. In *Immersive Analytics*. Springer, 2018, pp. 95–138.
- [6] CHANDLER, T., CORDEIL, M., CZAUDERNA, T., DWYER, T., GLOWACKI, J., GONCU, C., KLAPPERSTUECK, M., KLEIN, K., MARRIOTT, K., SCHREIBER, F., ET AL. Immersive analytics. In *2015 Big Data Visual Analytics (BDVA)* (2015), IEEE, pp. 1–8.
- [7] CUSACK, R., DECKS, J., AIKMAN, G., AND CARLYON, R. P. Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of experimental psychology: human perception and performance* 30, 4 (2004), 643.
- [8] DURWARD, J., LEVINE, J., NEMETH, M., PRETTIGIANI, J., AND TWEEDIE, I. T. Virtual reality network with selective distribution and updating of data to reduce bandwidth requirements, Aug. 19 1997. US Patent 5,659,691.
- [9] FIEBIG, F., AND LANSNER, A. A spiking working memory model based on hebbian short-term potentiation. *Journal of Neuroscience* 37, 1 (2017), 83–96.
- [10] GAVER, W. W., AND NORMAN, D. A. *Everyday listening and auditory icons*. PhD thesis, University of California, San Diego, Department of Cognitive Science and ..., 1988.
- [11] KASABOV, N., SCOTT, N. M., TU, E., MARKS, S., SENGUPTA, N., CAPECCI, E., OTHMAN, M., DOBORJEH, M. G., MURLI, N., HARTONO, R., ET AL. Evolving spatio-temporal data machines based on the neucube neuromorphic framework: design methodology and selected applications. *Neural Networks* 78 (2016), 1–14.
- [12] KASABOV, N. K., DOBORJEH, M. G., AND DOBORJEH, Z. G. Mapping, learning, visualization, classification, and understanding of fmri data in the neucube evolving spatiotemporal data machine of spiking neural networks. *IEEE transactions on neural networks and learning systems* 28, 4 (2016), 887–899.
- [13] LEIGH, J., VASILAKIS, C. A., DEFANTI, T. A., GROSSMAN, R., ASSAD, C., RASNOW, B., PROTOPAPPAS, A., DE SCHUTTER, E., AND BOWER, J. M. Virtual reality in computational neuroscience. In *Proc. Conf. on Applications of Virtual Reality* (1994), British Computer Society, Wiltshire, UK.
- [14] LIN, C.-Y., TSAI, K.-L., WANG, S.-C., HSIEH, C.-H., CHANG, H.-M., AND CHIANG, A.-S. The neuron navigator: Exploring the information pathway through the neural maze. In *2011 IEEE Pacific Visualization Symposium* (2011), IEEE, pp. 35–42.
- [15] MARAI, G. E., FORBES, A. G., AND JOHNSON, A. Interdisciplinary immersive analytics at the electronic visualization laboratory: Lessons learned and upcoming challenges. In *2016 Workshop on Immersive Analytics (IA)* (2016), IEEE, pp. 54–59.
- [16] MARRIOTT, K., CHEN, J., HLAWATSCH, M., ITOH, T., NACENTA, M. A., REINA, G., AND STUERZLINGER, W. Immersive analytics: Time to reconsider the value of 3d for information visualisation. In *Immersive Analytics*. Springer, 2018, pp. 25–55.
- [17] NOWKE, C., SCHMIDT, M., VAN ALBADA, S. J., EPPLER, J. M., BAKKER, R., DIESRNANN, M., HENTSCHEL, B., AND KUHLEN, T. Visnest—interactive analysis of neural activity data. In *2013 IEEE Symposium on Biological Data Visualization (BioVis)* (2013), IEEE, pp. 65–72.
- [18] OJEDA, A., BIGDELY-SHAMLO, N., AND MAKEIG, S. Mobilab: an open source toolbox for analysis and visualization of mobile brain/body imaging data. *Frontiers in human neuroscience* 8 (2014), 121.

- [19] PETERKA, D. S., TAKAHASHI, H., AND YUSTE, R. Imaging voltage in neurons. *Neuron* 69, 1 (2011), 9–21.
- [20] RAU, B., FRIESS, F., KRONE, M., MULLER, C., AND ERTL, T. Enhancing visualization of molecular simulations using sonification. In *2015 IEEE 1st International Workshop on Virtual and Augmented Reality for Molecular Science (VARMS@ IEEEVR)* (2015), IEEE, pp. 25–30.
- [21] SONNENSCHNEIN, D. *Sound design: The expressive power of music, voice, and sound effects in cinema*. Michael Wiese Productions, 2001.
- [22] STANNEY, K. Realizing the full potential of virtual reality: human factors issues that could stand in the way. In *Proceedings Virtual Reality Annual International Symposium'95* (1995), IEEE, pp. 28–34.