MEinVR: Multimodal Interaction Paradigms in Immersive Exploration

Ziyue Yuan Yu Liu Lingyun Yu*

School of Advanced Technology, Xi'an Jiaotong-Liverpool University

2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) | 978-1-6654-5365-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISMAR-ADJUNCT57072.2022.00026

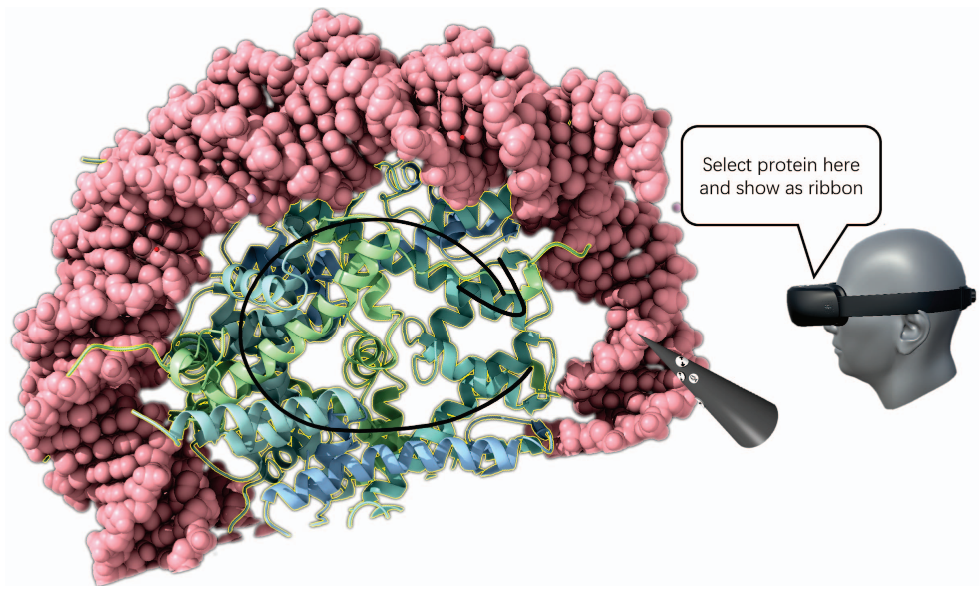

Figure 1: Multimodal interaction paradigm: the user interacts with the 3D molecular data through voice commands and VR controller.

Abstract

Immersive environments have become increasingly popular for visualizing and exploring large-scale, complex scientific datasets because of their inherent features: immersion, engagement, awareness, etc. Virtual Reality has brought rich opportunities for supporting a wide variety of novel interaction techniques, such as tactile and tangible interactions, gesture interactions, voice commands, etc. Multimodal interaction refers to that users are equipped with multiple modes for interacting with data. However, it is still important to determine how these techniques can be used and combined as a more natural interaction metaphor. In this paper, we aim to explore interaction techniques combining VR controller with voice input for a novel multimodal experience. We present MEinVR, a multimodal interaction technique that enables users to manipulate 3D molecular data in the virtual environment. Users can use VR controller to specify location or region of interest, and use voice command to express the tasks that they intend to perform on the data. This combination can serve as an intuitive means for users to perform complex data exploration tasks in immersive settings. We believe that our work can help inform the design of multimodal interaction techniques that incorporate multiple inputs for 3D data exploration in immersive environments.

Index Terms: Human-centered computing—Interaction techniques

1 Introduction

Recent advanced technologies have led to increasing attention on exploring scientific data in immersive virtual reality environments.

*e-mail:lingyun.yu@xjtlu.edu.cn

Hardware for supporting immersive visual exploration, such as the Oculus Quest, HTC Vive, Samsung Gear VR and Google Cardboard head-mounted displays (HMDs), are becoming progressively affordable and powerful. The rapid development allows researchers from both science and visualization fields to rethink how to make use of the advantages such as immersion, engagement and awareness, provided by immersive VR environments. These advantages can potentially enhance user perception and benefit the manipulation of 3D information content.

VR controller, as the main interaction technique provided by VR devices, has been widely used in 3D manipulations [4, 24] and data explorations [16]. However, data exploration with VR controllers has also been questioned for multiple reasons including being inefficient [19], requiring high learning costs for novice users [13], and fatigue issues in long-period manipulation [15].

However, many issues related to VR controller have been noticed and considered. These issues may affect the practical use of VR controller for data exploration. First, switching among different modes (such as sorting, selection, rotation, etc.) in VR is not trivial. In order to fully understand the structure and details of scientific data, domain experts often need to switch exploration modes, such as rotation, translation and selection. However, switching modes in VR cannot be easily performed. One possible solution is to create a 2D menu in 3D virtual environment. However, interacting with a 2D visualization widget in VR is not suggested in general [19]. Another possibility is to use triggers/buttons on the VR controller, but it may increase cognitive loads since users are required to remember the functions of the buttons [2]. These issues greatly affect the practical use of VR controller when multiple operations need to be performed in data exploration. Second, selecting data in 3D environment is not straightforward. Although VR environment provides an additional dimension of input which possibly offers more degrees of freedom in interacting with 3D data, 3D interaction is remarkably different

2771-1110/22/$31.00 ©2022 IEEE

DOI 10.1109/ISMAR-Adjunct57072.2022.00026

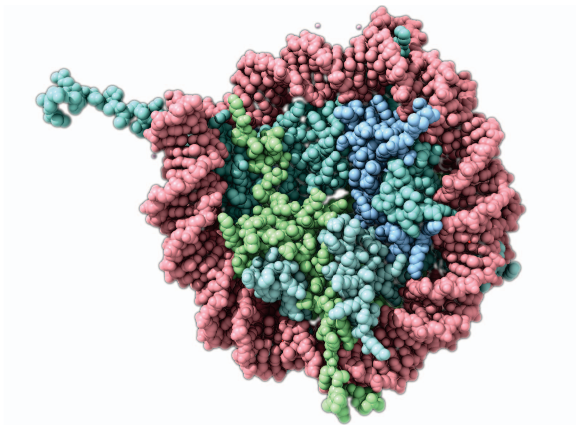

Figure 2: Three-dimensional molecular data, a Crystal structure of the nucleosome core particle at 2.8 Å resolution [17].

from the usual methods provided in traditional 2D systems. For instance, with 2D display, people can select data based on what they see on the screen. Moreover, people can draw a rectangle or a lasso to enclose interesting data. However, selecting 3D data becomes difficult in VR since users may not be able to see the hidden structure due to the occlusion. Furthermore, precise interaction is not easy to be performed because people may not be able to make an accurate judgement about spatial structures and inherent features of 3D data in VR. These challenges inspire us to consider the possibility of combining multiple interaction techniques (e. g., VR controller, voice commands, gestures) to complete data exploration tasks.

In this paper, we propose a multimodal interaction metaphor to explore 3D molecular data. The 3D molecular visualization (Fig. 2) shows a histone protein octamer and a superhelix with 146 base pairs of DNA. For exploring such a complex 3D data, our multimodal interaction provides the following advantages. First, it is an intuitive and natural interaction technique for data exploration. Our technique can potentially reduce users' learning costs and cognitive efforts. Users can *point at* the region of interest (*the data*), *tell* what they intend to do on the data (*the task*) and how to perform the task. Second, it can increase the efficiency of data exploration, especially when the data structure is complex and multiple interactions are required. Finally but importantly, the requirement of the accuracy of input can be greatly reduced. Through the integrated interactions, users' intention for the data exploration can be identified. This is particularly useful in exploratory visualization when data features and structures are unknown to users. To explore unknown data, it is common that people do not have a clear idea about the data structures and the location of interesting features. With the multimodal interaction metaphor, people can express their needs and intention in a more "general" and "abstract" manner and use the advantages of multiple interaction techniques to make the instruction more clear.

The final goal of our study is to understand the roles of individual interaction techniques in performing a complex exploration task. We hope that our research can further guide the design of multimodal interaction techniques for exploring complex data in immersive environments.

2 Related Work

In this section, we briefly review the interaction tasks in 3D data visualization. Also, we discuss the prior research on interaction techniques in virtual environments with a focus on voice-based techniques.

2.1 Interaction tasks for 3D visualization

The tasks for exploring 3D visualizations have been studied extensively. Based on the survey [1], the interaction tasks can be categorized into three groups: volumetric view and object manipulation, defining, placing and manipulating visualization widgets, selection and annotation. The first task focuses on how to support users in observing 3D data. This can be achieved in two ways, either by manipulating the virtual camera to change the position and orientation of the viewpoint, or by manipulating 3D objects, such as translation, rotation and scaling. The second task, defining and manipulating visualization widgets, such as seeding in flow visualization, requires users to define a location in 3D space. However, for unknown datasets, users may not know the exact location of the interesting features. The last task, selection, is one of the most fundamental tasks in 3D visualizations and has been extensively explored. Yu et al. [32] developed three selection strategies based on the density distribution to improve the efficiency and accuracy in selecting 3D point cloud data on 2D display. However, selecting data in VR proposes new challenges. For instance, due to the 3D visualization and occlusion issues, users may fail to know the exact locations of the targets. Moreover, when the data is unknown to users, it becomes necessary to support users switching among different tasks, such as rotation, translation and selection. However, there is no straightforward approach on how these tasks can be switched in VR. These issues have not been well addressed in the previous research.

2.2 Interaction techniques in VR environment

As the most widely used interaction method in VR [5,19,28,29], VR controller is not the most natural interaction interface for 3D data exploration. Most VR controllers come with a set of buttons and triggers which can be used to grab, move and rotate virtual objects. Although VR controllers offer 6DOF interactions, which provide users full flexibility in manipulating virtual objects as what they can do in the physical world [25], users are required to remember buttons and their associated functions.

Mid-air gesture interaction with motion tracking technologies is a more natural approach for VR applications. Gesture interaction mimics the physical actions that people make in the physical world. Because of the advantage of not relying on buttons or triggers, gesture interaction has been widely applied to exploration scenarios that require touchless interactions. For instance, in the medical field, a sterile environment needs to be maintained. Mid-air gesture interaction enables doctors to browse through 2D images and manipulate 3D medical data using only their hands [8]. Except for hand gestures, previous research also proposed interaction techniques based on the movements of other body parts, including head postures for target selection [30, 31] and symbol and text input [12], foot posture for pointing at objects [21], an integrated interaction with foot pedal and eye gaze for area-specific zooming [9], and full-body movements and gestures for multimedia educational contents interaction [7].

Voice input is also a widely used interaction technique and it is regarded as an effective method for switching interaction modes [18]. Voice input is often used in complex tasks or text-based scenarios, such as text entry or annotation. However, voice for direct manipulation is generally discouraged, for the reason that, with only voice commands it is difficult to specify the parameters that would be used for the interactions. For instance, it is not easy for the user to present the distance that the object should move or to describe the position of the rotation center. Nevertheless, voice-based interaction can be used as an additional input for the multimodal interaction paradigm, which will be discussed in the next section.

2.3 Natural User Interface

Natural User Interface (NUI) refers to a system that enables users to operate through intuitive actions [22]. It may be operated in a number of different ways or involve a combination of multiple inputs. Recently, many studies have involved speech in supporting NUI, and we usually call it natural language interfaces (NLIs). One important reason for involving natural language is to improve usability: while users normally know the questions they are concerned with or interested in their data, they can have a heavy burden on learning complex user interfaces and use them to find answers [11]. Cox et al. [3] presented some early works in the field of NLIs for visualization. Their work demonstrated that multimodal input adapts to a richer environment than a single modality [23]. The Articulate system [27] is a machine learning-assisted NLI for visualization. It parses user queries into commands and analysis data attributes to determine and generate suitable visualizations. DataTone [6] is another system that allows users to generate visualizations via typing or speaking queries. This system also keeps track of user corrections which may be used to improve the queries. The Eviza system [26] allows users to have richer interactive conversations with the given visualization than in previous work. Users can revise and update their queries in an iterative way. While existing NLIs for visualization facilitate some level of multimodal input, these systems focus more on responding to user queries in desktop-based environments rather than immersive environments. Our work builds upon techniques presented by prior work and extends them to support complex operations required to explore 3D molecular visualizations. We use the context of a given 3D molecular data visualization as our main use case in this work.

3 MEinVR: Multimodal Interaction in VR

In this section, we propose the design of MEinVR, a multimodal interaction technique for 3D scientific data exploration in VR environment.

3.1 Design Goals

The main objective of our multimodal interaction technique is to allow users to explore and manipulate 3D complex data in a natural and intuitive manner. More specifically, We list three goals to guide the design and implementation.

DG1 Reduce the accuracy requirements of each input. Most scientific data, such as 3D molecular data, has a complex structure and abundant features. It is often challenging for users to select a specific item or define a location in 3D space in two scenarios: first, people know how to describe the target data; however, it is difficult for them to find it or to describe the specific location in words; and second, they know the location of the target data, but it is challenging for them to select the data precisely through VR controllers. Our interaction technology should minimize the accuracy requirement of individual inputs from the user. Furthermore, even when there is no precise input from each interaction, the system can still combine input information obtained from multiple interactions to give the user accurate and fast results.

DG2 Facilitate the prominent advantages of each input. As discussed above, for unknown scientific data, users may not be able to describe what they are looking for or specify accurate regions through a VR controller. A combination of multiple interaction techniques, such as a combination of voice input, mid-air gestures and tangible devices, can be effective to achieve the exploration goal. The reason is that different interaction methods have varying characteristics and can complement each other when combined, thus compensating for their respective limitations. For instance, people can use voice commands to describe the hidden structure that might be difficult to select in the 3D environment; and use VR controller to manipulate data that may be hard to describe precisely with words. In order to maximize the use of hybrid interaction techniques, we first need to identify the prominent advantages of each interaction technique and find out how to combine their powers to complete a complex exploration task.

DG3 Minimize learning costs and cognitive efforts. Users should be able to focus on the data features and tasks, instead of remembering how to perform the interaction methods. This requires our multimodal interaction paradigm to be intuitive and easy to learn.

3.2 Design Overview

Based on the design goals discussed above, we developed MEinVR. An overview of the work process of our multimodal interaction paradigm is shown in Fig. 3. This method allows users to explore 3D scientific data through voice input and VR controller. Users can use a VR controller to define an approximate area where the interesting feature can be found and use voice input to express more abstract instructions. For example, as shown in Fig. 1, the user defines a region with VR controller around the area of interest on the 3D molecular data, and at the same time, says "select protein here and show as ribbon". In this case, the user knows an approximate area where protein may exist, but he/she does not know the exact position. Moreover, in order to distinguish protein from the surrounding atoms, the user wishes to change their rendering style. Note that, all inputs are not required to be precise; however, the user's intention is clearly identified because of the combined information from the multiple inputs. For instance, the user can use ambiguous words like "here" or "there" and point at an approximate area with the controller, and they can still obtain accurate results.

There are some interesting questions that require to be considered in the design. First, how different interaction inputs (such as voice command, VR controller and gestures) can be combined into a multimodal interaction technique. Second, how the multimodal interaction technique can be used to perform a complex exploration task. Furthermore, compared to the traditional interaction techniques, whether the multimodal interaction technique would be more effective in real data exploration. To answer these questions, we analyze the exploration tasks that users often perform on the data and investigate the prominent advantages of individual interaction input. Based on the findings, we design the multimodal interaction paradigm for exploring 3D molecular data.

3.3 Interaction Interface

Natural Language. The top row of Fig. 3 presents an overview of the pipeline for implementing NLIs that support users in manipulating 3D molecular data through natural language queries. Users' queries are captured by the microphone of the VR headset. The system transmits the audio source to the Google Speech-to-Text (STT) engine and the engine returns the text data of the commands which are further processed by the NLP engine. Based on the word vector of the command processed by Word2vec [20], the system compares the semantic similarity of the speech query with the existing queries in the system library using Spacy [14]. The cosine similarity between commands is calculated as follows:

$$similarity(A,B) = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_1^n A_i B_i}{\sqrt{\sum_1^n A_i^2}\sqrt{\sum_1^n B_i^2}} \quad (1)$$

Finally, the system returns the query which obtains the highest similarity with the input query. Furthermore, we also consider the possibility of multiple queries into one single command. When the system recognizes that multiple queries are included in one command, it splits the command into separate queries and performs similarity matching individually.

VR Controller. Similar to most VR applications, users can use the VR controller to rotate and translate an object in the VR environment. When the user presses the button and enters the "query"

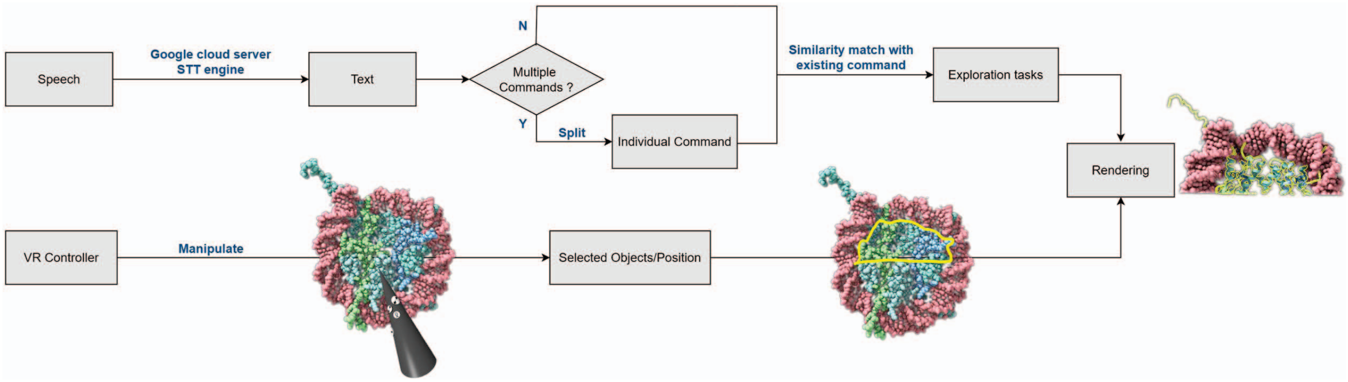

Figure 3: An overview of the working process of multimodal interaction paradigm.

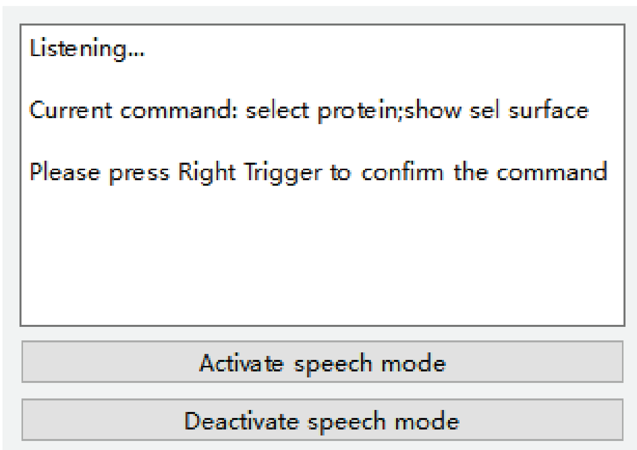

Figure 4: Users can view the execution of the command in the user interface.

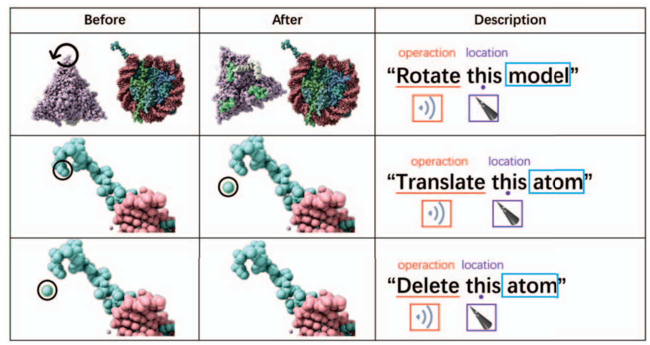

Figure 5: Simple tasks are supported in our systems with some examples.

mode, the system starts recording the real-time position of the VR controller. Combining the input information from the VR controller (for instance, the approximate area defined by the controller) and voice commands (the abstract instructions given by the commands), a specific exploration task (e.g., rotation, translation) will be determined and performed on the data (as shown in the bottom row of Fig. 3). The visualization will be updated accordingly.

3.4 Implementation

We developed the system using the Oculus Quest 2, a standalone VR headset with 6 DOF tracking. The headset is equipped with an 1832×1920 pixels screen (per eye). The system enables wireless VR connectivity to SteamVR via Oculus Air Link.

We implemented the system based on the open-sourced UCSF Chimera X [10] and visualize the data in an immersive VR environment. ChimeraX is the next generation interactive visualization program from the Resource for Biocomputing, Visualization, and Informatics (RBVI). Fig. 4 shows the user interface, which can be viewed both on the desktop screen and in VR environment.

4 Use Case

In this work, We implement a set of exploration tasks for exploring 3D molecular data. In this section, we demonstrate the feasibility and usability of MEinVR through two use cases: the simple use case is used to demonstrate the flexibility of switching among different visualization tasks, while the complex use case shows the intuitiveness and effectiveness of the multimodal interaction technique in the data exploration.

4.1 Use Case One: Simple Tasks

Fig. 5 lists the exploration tasks that are supported in our system. These tasks are the common tasks in data exploration, which have been integrated into desktop-based 2D molecular visualization systems. However, data exploration in the VR environment is not easy. Due to the unfamiliarity, users may need to switch among different exploration modes (translation, rotation and lighting) in order to gain a general idea about the data. However, in case users want to rotate a part of the data, they first need to switch to the "selection" mode to define a part of the data required to be rotated, then they should switch to the "rotation" mode to rotate the selected part around a certain axis. However, switching between these modes in VR is not simple. One possible solution is to switch modes through a dedicated menu and users are required to select options from the menu. However, system controlled mode is not suggested since users would need to remember the current mode. Moreover, 2D visualization widget is discouraged in the VR environment. Another possibility is to switch modes through buttons on the VR controller. However, users will need to remember the button functions, which will increase users' memory load and affect users' experience.

Our interaction technique supports users to switch among the modes directly through voice commands. People can directly give voice commands, for instance, "rotate this model" and "delete this atom". As an example, *"rotate"* is regarded as an exploration task and *"model"* indicates the object that the user is intended to manipulate. In addition, when *"this"* is captured, the object located at the

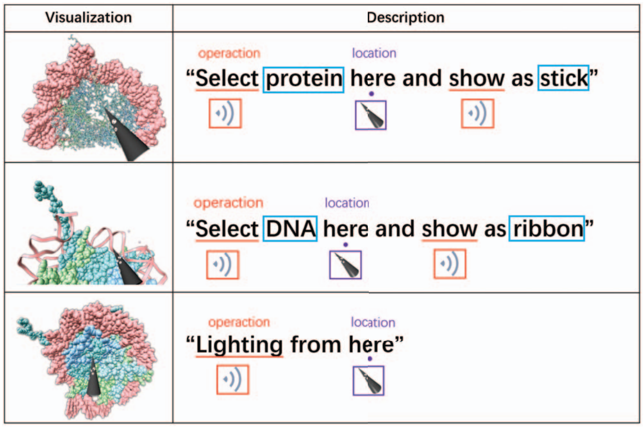

Figure 6: Complex tasks are supported in our systems with some examples.

position of the VR controller is identified.

4.2 Use Case Two: Complex Tasks

As mentioned above, to explore unknown data in the immersive environment, users may not know where the interesting features are or whether they even exist. Because of the occlusion, it becomes even more challenging to select the features precisely with VR controller. In this case, speech is a more effective and accurate approach for expressing abstract commands.

Our interaction technique supports users to manipulate objects directly with voice and VR controller, which can significantly simplify the interaction steps. For example, the user says “select DNA here and show it as ribbon”, and at the same time, he uses the VR controller to define a region in the VR environment. Our approach splits the detected commands into two queries: “select DNA here” and “show it as ribbon”. *“Select”* and *“show”* are exacted as exploration tasks, *“DNA”* and *“ribbon”* are regarded as the objectives, ‘textit‘here” is captured to indicate that the location of the target is at the position of the VR controller. Therefore, the location of the VR controller will be determined.

Moreover, depending on the tasks, users are also supported to define a region with VR controller. Then, all objects within the region are selected/manipulated. Note that, it is also possible that the objective which is indicated in the speech cannot be found within the defined region. In this case, no action will be performed on the data. For instance, when people say “select DNA here” while pointing to the data, however, no “DNA” can be found around the pointed area. Then nothing will be selected. Fig. 6 lists the complex exploration tasks supported by MEinVR.

In short, through the two scenarios we demonstrate how the proposed multimodal interaction technique can be used for data exploration in VR environment, with a focus on simplifying user interaction steps and improving interaction efficiency.

5 Conclusion

In this paper, we present MEinVR, a multimodal interaction technique that integrates inputs from VR controllers and speech to explore 3D molecular data in the VR environment. In the current prototype, VR controller is used to identify regions of interest in the data, and voice commands are used to specify abstract exploration tasks. We expect that a combination of multiple interaction techniques can amplify the advantages of each and support users to complete exploration tasks more effectively.

Further studies are required to explore the prominent advantages of each interaction input in different exploration tasks for various data. Moreover, a comprehensive user study needs to be conducted to evaluate the usability and effectiveness of our method in data exploration.

Acknowledgments

The work was supported by NSFC (62272396) and XJTLU Research Development Funding RDF-19-02-11.

References

[1] L. Besançon, A. Ynnerman, D. F. Keefe, L. Yu, and T. Isenberg. The state of the art of spatial interfaces for 3d visualization. In *Computer Graphics Forum*, vol. 40, pp. 293–326. Wiley Online Library, 2021.

[2] F. Chen, N. Ruiz, E. Choi, J. Epps, M. A. Khawaja, R. Taib, B. Yin, and Y. Wang. Multimodal behavior and interaction as indicators of cognitive load. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(4):1–36, 2013.

[3] K. Cox, R. E. Grinter, S. L. Hibino, L. J. Jagadeesan, and D. Mantilla. A multi-modal natural language interface to an information visualization environment. *International Journal of Speech Technology*, 4(3):297–314, 2001.

[4] J. Falah, S. Khan, T. Alfalah, S. F. M. Alfalah, W. Chan, D. K. Harrison, and V. Charissis. Virtual reality medical training system for anatomy education. In *2014 Science and Information Conference*, pp. 752–758, 2014. doi: 10.1109/SAI.2014.6918271

[5] J. Frommel, S. Sonntag, and M. Weber. Effects of controller-based locomotion on player experience in a virtual reality exploration game. In *Proceedings of the 12th international conference on the foundations of digital games*, pp. 1–6, 2017.

[6] T. Gao, M. Dontcheva, E. Adar, Z. Liu, and K. G. Karahalios. Datatone: Managing ambiguity in natural language interfaces for data visualization. In *Proceedings of the 28th annual acm symposium on user interface software & technology*, pp. 489–500, 2015.

[7] M. Gelsomini, G. Leonardi, and F. Garzotto. Embodied learning in immersive smart spaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020.

[8] H. Glas, J. Kraeima, P. van Ooijen, F. Spijkervet, L. Yu, and M. Witjes. Augmented reality visualization for image-guided surgery: a validation study using a three-dimensional printed phantom. *Journal of Oral and Maxillofacial Surgery*, 79(9):1943–e1, 2021.

[9] F. Göbel, K. Klamka, A. Siegel, S. Vogt, S. Stellmach, and R. Dachselt. Gaze-supported foot interaction in zoomable information spaces. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, pp. 3059–3062. 2013.

[10] T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, and T. E. Ferrin. Ucsf chimerax: Meeting modern challenges in visualization and analysis. *Protein Science*, 27(1):14–25, 2018.

[11] L. Grammel, M. Tory, and M.-A. Storey. How information visualization novices construct visualizations. *IEEE transactions on visualization and computer graphics*, 16(6):943–952, 2010.

[12] J. P. Hansen, A. Alapetite, M. Thomsen, Z. Wang, K. Minakata, and G. Zhang. Head and gaze control of a telepresence robot with an hmd. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, pp. 1–3, 2018.

[13] M. Holly, J. Pirker, S. Resch, S. Brettschuh, and C. Gütl. Designing vr experiences–expectations for teaching and learning in vr. *Educational Technology & Society*, 24(2):107–119, 2021.

[14] M. Honnibal and I. Montani. spacy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing. *To appear*, 7(1):411–420, 2017.

[15] S. Jang, W. Stuerzlinger, S. Ambike, and K. Ramani. Modeling cumulative arm fatigue in mid-air interaction based on perceived exertion and kinetics of arm motion. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 3328–3339, 2017.

[16] B. Lee, X. Hu, M. Cordeil, A. Prouzeau, B. Jenny, and T. Dwyer. Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1171–1181, 2020.

[17] K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, and T. J. Richmond. Crystal structure of the nucleosome core particle at 2.8 å resolution. *Nature*, 389(6648):251–260, 1997.

[18] S. McGlashan and T. Axling. A speech interface to virtual environments. *Swedish Institute of Computer Science*, 1996.

[19] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge. A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In *Computer graphics forum*, vol. 38, pp. 21–45. Wiley Online Library, 2019.

[20] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[21] K. Minakata, J. P. Hansen, I. S. MacKenzie, P. Bækgaard, and V. Rajanna. Pointing by gaze, head, and foot in a head-mounted display. In *Proceedings of the 11th ACM symposium on eye tracking research & applications*, pp. 1–9, 2019.

[22] K. O'hara, R. Harper, H. Mentis, A. Sellen, and A. Taylor. On the naturalness of touchless: putting the "interaction" back into nui. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 20(1):1–25, 2013.

[23] S. Oviatt and P. Cohen. Perceptual user interfaces: multimodal interfaces that process what comes naturally. *Communications of the ACM*, 43(3):45–53, 2000.

[24] S. Reddivari, J. Smith, and J. Pabalate. Vrvisu: A tool for virtual reality based visualization of medical data. In *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, pp. 280–281, 2017. doi: 10.1109/CHASE.2017.102

[25] W. Robinett and R. Holloway. Implementation of flying, scaling and grabbing in virtual worlds. In *Proceedings of the 1992 symposium on Interactive 3D graphics*, pp. 189–192, 1992.

[26] V. Setlur, S. E. Battersby, M. Tory, R. Gossweiler, and A. X. Chang. Eviza: A natural language interface for visual analysis. pp. 365–377. Association for Computing Machinery, Inc, 10 2016. doi: 10.1145/2984511.2984588

[27] Y. Sun, J. Leigh, A. Johnson, and S. Lee. Articulate: A semi-automated model for translating natural language queries into meaningful visualizations. In *International Symposium on Smart Graphics*, pp. 184–195. Springer, 2010.

[28] D. Vogel, P. Lubos, and F. Steinicke. Animationvr-interactive controller-based animating in virtual reality. In *2018 IEEE 1st Workshop on Animation in Virtual and Augmented Environments (ANIVAE)*, pp. 1–6. IEEE, 2018.

[29] D. Wolf, K. Rogers, C. Kunder, and E. Rukzio. Jumpvr: Jump-based locomotion augmentation for virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2020.

[30] Y. Yan, Y. Shi, C. Yu, and Y. Shi. Headcross: Exploring head-based crossing selection on head-mounted displays. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1):1–22, 2020.

[31] Y. Yan, C. Yu, X. Yi, and Y. Shi. Headgesture: Hands-free input approach leveraging head movements for hmd devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(4):1–23, 2018.

[32] L. Yu, K. Efstathiou, P. Isenberg, and T. Isenberg. Efficient structure-aware selection techniques for 3d point cloud visualizations with 2dof input. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2245–2254, 2012.