Abstract

Virtual Reality (VR) technology has the potential to provide a highly immersive experience when navigating through densely populated environments, such as macro-molecular models. However, current solutions lack the ability to provide both a sense of immersion and a clear overview of the environment. In this study, we propose improvements to the existing systems for an automatically generated guided tour through a molecular model that address these limitations and enhance the overall experience. We propose a sparsification technique based on implicit quadric equations, providing seamless transitions between a closed, enveloping experience and an open, spacious one for improved spatial awareness. Additionally, we give users more control over the tour to alleviate feelings of being overwhelmed in a crowded environment. Our subjective, qualitative results indicate that these methods can improve the overall VR experience compared to existing solutions and provide a more enjoyable tour. However, more research is needed to further enhance the experience and provide an even more engaging and informative guided tour through densely populated molecular environments within VR.

Keywords: sparsification, virtual reality, dense environment, molecular model

1 Introduction

The COVID-19 pandemic has increased public awareness and interest in molecular assemblies. Scientists now have better ways to depict biological structures using 3D modeling approaches, but the use of scientifically accurate models in educating the general public remains limited. One of the reasons being that without proper explanation, the resulting images or videos are often confusing for those without scientific knowledge and can only be fully grasped with some sort of guidance from an expert. This highlights the need for a new approach to science communication, one in which the user can visualize, investigate, and ultimately understand a complex scene, such as a molecular model, in a non-expert setting. To address this issue, researchers have recently developed new systems [6, 1] that allow the creation of interactive scientific documentaries using multi-scale, multi-instance, and dense 3D molecular models.

These systems provide a journey through a very crowded environment filled with proteins, lipids, fibers, genetic material, and other structures of various colors, shapes, and sizes. For instance, one of the models shows a SARS-CoV-2 virus with all its parts, such as the spike proteins, the envelope, the membrane, and the RNA, floating in a "box" filled with blood plasma and all its components. As the camera guides us to key locations within the model, a voice-over explains what these structures are and what functions they perform.

While both systems provide a unique, interesting, and educational experience of a guided tour through a molecular model, there are some issues regarding both of them that were pointed out during a user study conducted by the authors of Nanotilus [1]. One key distinction between the two, from a user experience perspective, is the way each system handles a crowded environment, specifically the sparsification of it. Molecumentary [6] uses a vertical clipping plane set at a particular distance from the viewer and thus avoiding collisions with any objects. Nanotilus on the other hand uses a camera-centric sparsification technique that removes instances in an ellipsoid shape around the user. The different sparsification techniques can to some extent be seen in Figure 1.

Users reported that with Molecumentary, they did not see a huge added value in being inside a VR as opposed to seeing a video play on a screen. The clipping plane sparsification made the scene seem flat. In contrast, Nanotilus provided a more immersive experience within a 3D model, but at the cost of compromised spatial awareness. Besides that, users also reported feeling overwhelmed with all the information presented to them at the same time and not having enough control to look at and explore the space around them.

In this paper, we present our efforts towards improving
the VR experience inside a guided tour through a crowded molecular model, namely:

- A novel sparsification approach based on quadric surfaces that combines the immersive experience with good spatial awareness.
- An innovative “landing plane” sparsification technique to better connect the real-world standing environment with the VR flying experience.
- Limited user interaction to mitigate the overwhelming effects of excessive information.

In Sections 2 and 3, we introduce the field, outline the foundational systems on which our research is based, and present an analysis of a user study that highlights their limitations. In Section 4, we elaborate on the new techniques we have created to overcome these limitations. We then evaluate the effectiveness of these methods in Section 5, by analyzing their impact on the perception of crowded environments and the user experience of navigating through them. Finally, in Section 7, we provide a summary of our work and propose potential areas for future research and improvement.

2 Related Work

Visualization in scientific communication

Visualization plays a critical role in scientific communication, used for experiment validation and data exploration. Many people believe that its significance will increase even further in the future, with the potential to lead to new breakthroughs in research [2]. For many years, visualization has been recognized as an important tool for promoting understanding of science [5], particularly in presentations of the cosmos- or micro-scale world and associated processes, as people can rely only on illustrations, interactive simulations, and animations as visual aids. Given that, virtual reality (VR) has emerged as a powerful tool that is more than simply a better version of 3D visualization [16] as it enables us to form a conception of and understand things that 2D or even desktop 3D graphics cannot. As a result, VR has been applied in various fields such as biology [18, 19, 12], medicine [9, 15], and other areas [13]. Although classical VR applications offer a valuable learning experience by enabling users to examine objects from a different perspective, they often lack the necessary guidance to enhance their understanding of complex structures. Tools such as the one described in [14] have been developed to allow the creation of VR applications that incorporate storytelling elements, either through written or spoken narration. However, creating a narrated story manually requires substantial time and expertise, leading to the development of new methods like the ones our work is based on and are further discussed in Section 3, which aim to automate the creation process for narrated VR tours.

Occlusion management in VR

As we delve into the microscopic world through molecular models, we discover that the environment is densely populated with elements. To realistically depict this information in our models and VR experiences, while also giving the user some breathing room and preventing collisions, we need some sort of occlusion management. While there are numerous different approaches to this, falling into five design patterns as identified by Elmqvist and Tsigas [4], researchers most often deal with it using sparsification, which involves removing some of the instances inside the model. One way to manage occlusion is to use a cutting plane that divides the scene into visible and hidden parts. However, to keep important objects visible, an extension that exempts certain objects from being cut away was used by authors of Molecumentary in this and previous [7] research. Le Muzic et al. [10] proposed a technique called the visibility equalizer, which also considers
the type of instance being removed, enabling the removal of more abundant objects. A downside to using clipping planes is that they can remove large portions of visualization which reduces the sense of crowdedness in one part of the environment. To avoid this, smart sparsification methods that remove individual instances using different importance measures have been proposed, such as the one presented in [17] by Lesar et al. Another technique for managing occlusion was proposed by Elmqvist [3], who suggested distorting space with a spherical force field that repels objects around the 3D cursor. Similarly to this, Nanotilus uses a centralized sparsification technique that removes instances in concentric zones around the user. In addition, Nanotilus utilizes an automatic visibility equalizer that guides the sparsification process using heuristics.

No single method is inherently superior to the other. The method we choose depends entirely on our use case. For certain visualizations, a clear cutting plane may provide a better model overview, while in other cases, we may prefer to experience the sense of crowdedness. As a guided tour is dynamic, we encounter various scenarios within it, which may call for different sparsification techniques depending on what we aim to see. To address this, we propose a technique that combines both a selective clipping plane and local centralized sparsification.

3 Background

We based our work on two recently published systems, both providing a VR experience of narrated documentaries of molecular models.

Koufil et al. introduced the concept of adaptable documentaries in their work on Molecumentary [6]. They presented a system comprising of real-time visualization, automated exploration, and synthetic commentary to create an adaptable and automated narrated tour. The system consists of two steps: story graph foraging and narrative synthesis.

The first step involves automatically compiling information about the biological model into a story graph that holds model elements, their relationships, and verbal descriptions. The second step utilizes the story graph to generate a sequence of story elements with corresponding commentary using text-to-speech technology. Subcomponents of the model are brought to life using camera animations and occlusion management. The order of model elements shown is determined by either an algorithmic approach that produces a self-guided documentary or by following a storyline supplied as a written text input, making the moleculary more responsive to user choices. The authors’ approach automates the entire process without the need for a domain expert’s involvement.

Alharbi et al. published Nanotilus [1] as a follow-up to previous research, where they aimed to address the limitations of existing visibility and occlusion management strategies for navigating dense 3D structures. They introduced a new sparsification method that offers an endoscopic inside-out perspective, maintaining the immersive quality of virtual reality, instead of the previously used outside-in view. Since this is the aspect that we focused more on in our paper, the mechanics of it are described in more detail in Section 4.1.

The authors also changed the journey planning part of the pipeline, which roughly replaces the story graph foraging in Molecumentary. They included the use of void spaces as the navigation through them minimizes the number of instances to sparsify, preserving the model realism and increasing the participants’ immersion. The process of journey planning starts with identifying the instances that the journey should visit and then generates a path that connects them while traversing through the void spaces.

A user study comparing the two systems was conducted with 29 participants who viewed guided tours of mesoscopic biological models of SARS-CoV-2 and HIV using both of them. The primary aim of the study was to evaluate the user experience regarding various sparsification and navigation techniques. User feedback was collected on engagement and overall user experience, spatial understanding of the displayed structures, traversal of the virtual environment, and the ability to orient oneself in the environment and follow the story.

Participants in the user study preferred the inside-out animation of Nanotilus for its immersive quality, while also noting that it could be difficult to see the whole structure at times by commenting "[...] its more difficult to understand how all the viral components are positioned." and "Sometimes in the inside view it is too close to see the whole structure.". On the other hand, Molecumentary was criticized for lacking immersiveness with comments such as "I honestly feel that this experience doesn’t add much to what you would get just looking at a video on a good screen." and "[...] it was not that an immersive experience as the inside-out animation, [...].". However, users appreciated Molecumentary’s outside-in approach, which provides an overview of the virus and displays each component clearly. To balance these strengths, we aimed to create a sparsification method that incorporates both approaches at appropriate points during the tour.

4 Improving the VR experience

The primary area of possible improvement that we focused on was sparsification, as it was one of the key distinctions between the two systems and was a significant point of feedback in the user study. Additionally, we tried to add a bit more control over the tour to the user in order to mitigate feelings of being overwhelmed, which was another issue mentioned in the study.
Quadratic equations are a type of algebraic equation that defines a space that can be defined by a quadratic equation. A quadric surface is a type of surface in three-dimensional space that will be described in detail next.

4.1.1 Quadric surfaces

A quadric surface is a type of surface in three-dimensional space that can be defined by a quadratic equation. Quadratic equations are a type of algebraic equation that can be written in the form:

\[ ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0. \]

where \( x, y, \) and \( z \) are the coordinates of a point on the surface and \( a, b, c, d, e, f, g, h, i, \) and \( j \) are constants. The constants in a quadric equation are real numbers and at least one of them has to be non-zero. These surfaces can be classified into different types depending on the values of the constants in the equation, such as spheres, cylinders, cones, etc.

Representing a capsule shape that was already used in Nanotilus was straightforward, as we just used the implicit equation for an ellipsoid:

\[ \frac{x^2}{A^2} + \frac{y^2}{B^2} + \frac{z^2}{C^2} = 1 \Rightarrow Ax^2 + By^2 + Cz^2 - J = 0. \]

For representing the flat clipping plane, we decided on a paraboloid, because the parameters in the equation can easily be set up such that the resulting surface resembles a plane:

\[ z = x^2 + y^2 \Rightarrow Ax^2 + By^2 - Iz = 0. \]

Both surface representations can be seen in Figure 3.

The logic behind choosing instances that had to be removed remained unchanged from Nanotilus. Individual instance memberships were determined through a simple calculation: if the surface equation with the instance’s position as \( x, y, \) and \( z \) resulted in a value equal to or less than 0, the instance was considered part of the shell; otherwise, it was not. In Figure 4, we show the instances inside a simple molecular model, colored based on their corresponding shell membership. The innermost shell is colored green, the middle shell is blue, and the outermost shell is red.

When using the ellipsoid shape, what we call the capsule mode, we kept the three shells with sparsification ratios of 0, 0.33, and 0.66. During the plane mode when we used the paraboloid shape, we only used one shell with a sparsification ratio of 0, meaning everything inside was sparsified.

Figure 2: The illustration shows the sparsification procedure, where instances are sparsified based on both importance of the type (cyan or red) and their shell membership (source: [1]).

4.1 Sparsification of crowded environments

The authors of Nanotilus focused on developing a local, camera-centric sparsification procedure that allows the viewer to go through a densely packed scene without colliding with the instances. It reduces structural occlusion and provides the user with endoscopic views that convey scene crowdedness.

The sparsification is controlled with three nested and concentric shells surrounding the camera, each with its own visibility percentage. The innermost shell is responsible for hiding instances that may collide with the camera and is set with a visibility rate of 0.0. The middle and outer shells have their default visibility rates set to 0.33 and 0.66 respectively, making some instances visible and some removed. A mathematical description of the shell geometry is used to determine whether an instance in a scene is a member of a shell or not.

The sparsification is then done in two phases. First, the visibility values of instances that should be hidden are updated based on the visibility percentage of the shells. The algorithm then checks for the number of invisible instances inside the shells and assigns a weight to each visible instance, which represents its priority to be hidden. The weight is affected by the distance from the instance to the camera and the importance of the type. The instances with the smallest weight are hidden until a certain threshold is achieved. In Figure 2, we can see an example of selecting instances for sparsifying.

In our implementation, we took the main logic from Nanotilus that is described above, however, the shell shape and implementation of it were modified to facilitate utilization for both capsule-type sparsification and the clipping plane. This was achieved using quadric surfaces, which will be described in detail next.

Figure 3: Ellipsoid (a) and paraboloid (b) drawn using implicit quadric equations.
The smoothness of the visual experience in VR is essential
to provide a comfortable and natural-feeling environment.
This is because changes in the real world happen gradu-
ally, not abruptly, and sudden changes in the surroundings
can result in discomfort. To recreate this natural prog-
ression in VR, we implemented smooth transitions between
the two sparsification modes. Reflecting on how the trans-
sitions feel to the user, we refer to the transition from cap-
sule to plane mode as opening up and the reverse transi-
tion from plane to capsule as closing down. Our imple-
mentation of the surface shapes with parametric implicit
equations made calculating the steps between them triv-
ial. For each of the parameters, we divide the range be-
tween the start and end values into the desired number of
steps. Combined with a gradual fade-out of instances as
their membership changes, this produces a smooth and vi-
sually appealing transition between the shapes.

In reviewing the results of the user study and conduct-
ing our own testing of the systems, we tried to determine
the optimal points for transitions and when certain sparsifi-
cation modes were most appropriate within the document-
ary. We concluded that while the user is moving through
the scene from one target to another, using the capsule
mode makes for a more immersive experience. When stop-
ping at the target and entering a so-called focus scene,
the plane mode provides a better overview of the scene and
gives the user a good perception of the environment. We,
therefore, implemented opening up the sparsification im-
mediately after stopping at the target of the focus scene.
The sparsification then closes down when the user starts
moving toward the next target. Another transition hap-
pens at the beginning of the tour when we want to intro-
duce the model that we’re showing to the user, so we start
with the plane mode and transition to the capsule mode
before moving forward. Although we found that the cap-
sule mode is preferable while moving in most cases, we
observed that when the camera moves backward, the close
proximity of the instances on the left and right and new
instances popping into the scene in front of the user can
cause some discomfort. To address this, we experimented
with opening up the sparsification after we detect that the
camera started moving in reverse.

4.1.3 Landing plane

Another feature that we wanted to test for increasing the
enjoyment inside VR is to add a horizontal plane as a floor,
providing the user with a sense of stability and grounding.
Our hypothesis is that adding a floor plane when the user
is standing still would help bridge the gap between virtual
reality and real-world experiences, as most of the time in
reality we are standing on a two-dimensional plane. By
creating a more realistic virtual environment, we aim to
increase the user’s enjoyment and make the virtual expe-
rience feel less alien.

In order to maintain the density of the scene and avoid
adding unnecessary, false or misleading elements to the
model, we came up with a solution that utilizes existing
instances in the scene to create the floor. This way, the
floor landing plane is just another form of sparsification,
achieved by removing instances that are located above a
horizontal plane and were previously already selected to
be removed through earlier phases of sparsification.

The floor under your feet makes sense only when you are
standing still, not when you are moving around in the
capsule. Therefore, we integrated the floor sparsification
with the plane mode as can be seen in Figure 5, mean-
ing that it toggles on and off with a fading effect whenever
there is a transition between the sparsification modes.
Since the transition takes place whenever we arrive to the
target instance inside a capsule mode, the floor sparsifica-
tion then has a feeling of landing that capsule. That’s why
we refer to it as a landing plane.

4.2 Bringing more control to the user

Being in a VR environment, where everything is con-
stantly moving and the user has no control over their di-
rection or speed, can be quite disorienting. To enhance
the user’s comfort and overall experience, it’s crucial to give
them some degree of control.

As a simple but effective solution, we added the option
to pause the guided tour at any point during the experience.
This allows the user to take a break, look around, gather
more information about what interests them, or simply
pause for a moment. They can also switch to plane mode
if they feel lost, claustrophobic, or overwhelmed while in
capsule mode. Once they’re ready to continue the tour, the
sparsification automatically returns to the previous mode
and the guided tour resumes.

This feature provides a much-needed sense of control
and comfort, making the overall experience more enjoy-
able and less overwhelming.
5 Results

Our implementation of the system is built upon the Nanotilus platform, which utilizes the Marion library [8]. The application is implemented in C++, OpenGL, and GLSL, leveraging the Qt library, and employs GLSL compute shaders. For the presentation of the resulting guided tours, we utilized both the HTC Vive Cosmos and Oculus Quest Pro headsets. The support for OpenVR was integrated into Nanotilus for use with the HTC headset, but we encountered limitations when using the Oculus. This, along with the deprecation of OpenVR, has motivated us to explore the implementation of OpenXR for improved performance and compatibility in the future.

For developing and testing, we used the SARS-CoV-2 model that was created using a statistical and rule-based modeling approach [11]. The application was run on an Nvidia GeForce RTX 4090 graphics card, AMD Ryzen Threadripper PRO 3995WX 64-core processor, Windows 10, and Qt 5.15.2, with an average framerate of about 25 FPS per eye. Despite being below the typical FPS recommendations for VR, which call for a minimum of about 45 FPS to provide a reasonable experience, and at least 90 FPS for optimal performance, the average framerate was sufficient for us to test the design and evaluate the improvements made to the system. However, we are committed to improving the framerate and overall performance in the near future.

6 Discussion

We tested our implementation and qualitatively evaluated the improvements made to the VR experience.

Our testing revealed that the combination of sparsification modes enhances the overall experience by offering the best of both worlds. The user can fully immerse themselves in the model, feeling connected with the scene as they move from one instance to the next, while still maintaining a sense of their position within the model in the focus scene. We found the transitions between modes visually pleasing, with closing down feeling like you get hugged back into the scene and opening up feeling like taking a deep breath.

We still have to experiment to determine the optimal moments for switching between sparsification modes and when each mode would be most effective during the tour. Additionally, we may want to test out different surface parameters to find the best option. Furthermore, we need to conduct additional tests on the use of the landing plane within the tour to assess whether it evokes the hypothesized feelings when users have the ability to walk around.

We found that when simply being in the scene having a plane underfoot doesn’t necessarily enhance the experience, but it also doesn’t worsen it. The density of the
model is another crucial factor that affects the success of the experience, as less dense models may not provide enough instances to make up the floor, so we may need to investigate this further.

With the additional option of pausing and resuming the tour at any time, as well as the ability to toggle between sparsification modes, the guided tour moves a step closer to becoming a self-exploratory type of documentary. The user has slightly greater control and freedom to personalize their experience, which leads to a more enjoyable journey. Moreover, the pause feature gives the user the ability to take a break, explore the environment, and gather information at their own pace. This creates a more relaxed experience, allowing the user to fully immerse themselves in the VR world.

The control that we give to the user is fairly limited at this point, but it is an important aspect that we will focus on in the future.

Overall, after testing out the system, we have observed that the implementation of the combination of sparsification modes and the added pause feature has indeed had a positive impact on the overall VR experience. However, to verify and validate these findings, we plan to carry out a comprehensive user study that will be conducted on a larger scale. This will allow us to obtain a more accurate and representative evaluation of the impact of our work on the VR experience.

7 Conclusions

In this paper, we proposed improvements to the VR experience for guided tours through crowded molecular models. Our improvements aimed to address the limitations of the existing systems, including the sparsification of crowded environments and the overwhelming effects of excessive information. We proposed a novel sparsification approach based on quadric surfaces, a landing plane sparsification technique, and gave more control to the user during the tour. Our evaluation showed that our methods improved the perception of crowded environments and the user experience of navigating through them.

In the future, we plan to conduct a comprehensive user study to better understand the strengths and limitations of our system. Based on the results of this study, we will focus on improving the technical aspects of our tour implementation to enhance the user experience. This includes optimizing our code to increase the framerate and transitioning from OpenVR to OpenXR to reduce overhead.

Additionally, we want to focus more on adding the exploration aspect to the now completely guided tour as these features have the potential to greatly increase the usability and impact of this type of science communication. To achieve this, we plan to add new and exciting features such as the ability for users to change the story while they are on the tour, explore structures up close by walking around, and select and learn more about the structures that interest them. These enhancements will allow users to have a more interactive and personalized experience while exploring the wonders of science.

References


[9] Rakesh Mishra, Krishna M.D, Giuseppe Emmanuele Umana, Nicola Montemurro, Bipin Chaurasia, and


