Daz Studio to VRM pipeline: a discussion on the use of realistic VRM 3D models in upper body motion tracking

Roberta Fernandes[1]**, Ismar Frango Silveira**[2]

[1]Escola de Engenharia
Programa de Pós Graduação em Engenharia Elétrica e de Computação
Universidade Presbiteriana Mackenzie

[2] Faculdade de Computação e Informática
Programa de Pós Graduação em Computação Aplicada
Universidade Presbiteriana Mackenzie

`robertas@id.uff.br, ismar@mackenzie.br`

***Abstract.** **I**n recent years, especially after the pandemic, virtual reality avatars and virtual counterparts have experienced a surge in popularity, due to their usefulness in settings where people don't, or can't, see each other frequently. However, there are still very few ways to actually create such avatars that offer realism and ways to customize that are low-cost and available to the general public. This paper explores the process of converting photorealistic models created in DAZ Studio to VRM and achieves an end result that can be used in virtual reality, the metaverse, and live motion tracking applications.*
***Keywords** DAZ Studio, VRM format, Diffeomorphic Plugin, Blender, Virtual Reality (VR), Face Tracking, Finger Tracking, Vtubing, Avatars.*

1. Introduction

The COVID-19 global pandemic has brought significant changes to people's behavior, especially when it comes to online content consumption. Online platforms such as YouTube and Twitch offer a wide variety of content, from live-streams and gaming, to art performances, and were heavily used during lockdown periods [Machado 2023]. The adoption of virtual reality (VR) head-mounted displays (HMD) has also seen a boom, as working, learning, medical care, and communication activities in general started becoming long-distance [Ball et al. 2021]. In these 3D virtual environments, avatars are our digital counterparts. Their appearance and presentation matters to the user, as well as the other people sharing the same virtual space and using their own avatars [Latoschik et al. 2017]. 3D avatars have been increasingly used in various applications. These avatars, also referred to as digital humans, have drawn growing attention from researchers in an effort to develop and improve techniques related to fidelity, realism, empathetic response, and interactivity [Vilchis et al. 2023].

Facial animation is a method for virtual characters to include both facial posture and movement. Graphical visualization is commonly used in this approach, particularly in fields such as gaming, filmmaking, and human-computer interaction, in response to the growing demand for engaging interactions, advanced facial expressions, and realistic

motion [Noor et al. 2018]. The blendshapes of a face model are a group of morphed versions of the model [Wang et al. 2023]. They behave like the muscles of the human face and are indispensable features for facial animation. However, they are tricky and laborious to make, like the rigging bones of the model, which prompts many people to find shortcuts to this process. DAZ Studio is a freeware media design software developed by Daz 3D. It's a 3D scene creation and rendering application that produces images and videos using the render engines 3Delight or Nvidia Iray. It is compatible with most 3D file formats and has a system of humanoid figures called Genesis, which come pre-rigged and allows the user to use morphs to change the shape and appearance of the character [DAZStudio 2025]. Blender is a free and open-source 3D creation software that offers a comprehensive suite of tools for modeling, animation, simulation, rendering, compositing, and video editing [Blender Foundation 2025a]. One of its key features for facial animation is the use of Shape Keys—also known as morph targets or blend shapes—which allow objects to be deformed for use in various animations [Blender Foundation 2025b]. The Diffeomorphic DAZ Importer is a Blender add-on designed for importing models from DAZ Studio into Blender with enhanced compatibility for animation and rigging workflows. The software consists of two components—one on the DAZ Studio side and the other on the Blender side—ensuring a seamless export and import process [Diffeomorphic 2025]. The CATS Blender add-on is a tool designed to shorten steps needed to import and optimize models into VRChat (which uses the VRM format). A great variety of softwares and file formats are compatible, such as models from Unreal Engine, MMD, Source Engine, DAZ Studio and potentially more [Quantum 2023]. The VRM Add-on for Blender adds VRM import, export, and editing capabilities to Blender. It also adds features such as importing and exporting VRM, adding VRM Humanoid and setting MToon shaders. It supports Blender versions 2.93 to 4.4 [Saturday-06 2024]. Warudo is a comprehensive, feature-packed 3D VTubing software designed for both casual and professional users. It offers support for a wide variety of motion capture systems. It also offers a wide range of features, including posing your character with 500+ idle animations, integration with livestreaming platforms, a node-based visual scripting system, integration with Unity and other capabilities [Warudo 2025].

This article details a workflow using Diffeomorphic, a powerful plugin for importing DAZ models into Blender, with additional add-ons to ensure control over rigging, animation, and customization.

2. Related Works

In a study conducted by Freeman and Maloney, it was observed that participants often chose to present themselves in ways consistent with their offline identities. Although they occasionally experimented with different genders or non-human avatars, many still preferred to reflect their real-world identities in their virtual representations. Participants noted that their experiences in social virtual reality allowed them to discover aspects of themselves they had not been previously aware of [Freeman e Maloney 2021]. Fraser et al. identified in their experiment that the avatars displaying greater facial and bodily expressiveness were those with the most realistic facial and bodily animations, and were also perceived as the most comfortable to interact with [Fraser et al. 2022]. The ability to customize avatars is crucial, and making them more relatable and suitable for various contexts are important mechanisms of self-presentation online

[Freeman e Maloney 2021].

VTubers (Virtual Youtubers) are live streamers that use animated virtual avatars as their way to communicate with their viewers. It is a rapidly growing form of expression and entertainment, especially in East Asia, often associated with gaming and virtual reality. However, setting up an avatar for such activities is still a challenge due to budget, technicalities and other aspects. Live-streaming with a VTuber avatar is still a new, unexplored field, and there is a gap in the literature when it comes to technical aspects [Tang et al. 2021]. Many VTuber systems are phone/webcam-based and do not offer arm/finger tracking, while also relying on hotkeys for expression triggering, which can compromise expressiveness, interaction and even put a cognitive burden on the streamer, that often will be playing a game while controlling the avatar's actions [Tang et al. 2021].

VTubing has seen an exponential growth since 2020, when technology to do so became more accessible and people needed new ways to deal with stress due to the COVID-19 Pandemic. It became a new way to express yourself, communicate and even share music, due to the impossibility of concerts in person [Tambunan e Setiawan 2023]. While previous research has primarily focused on the social and cultural aspects of VTubing, there is a noticeable lack of studies examining the practical challenges VTubers face in creating and operating their avatars [Kim et al. 2025].

3. Methodology

The workflow begins in DAZ Studio, where we first create a Genesis 8 Female model by adding it to the scene and dressing it with clothes and hair to give it the desired aesthetics.
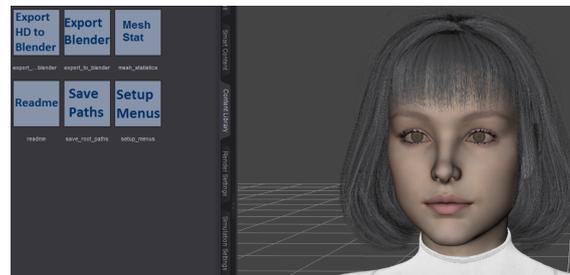
Figura 1. The G8F model in DAZ Studio, ready to be exported using Diffeomorphic.
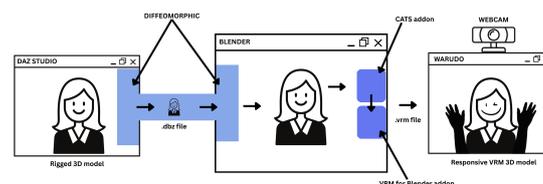
**Figura 2. Flow chart representing the workflow**

Once the Genesis 8 Female model has been prepared, the next step involves exporting it from DAZ Studio using the Diffeomorphic DAZ Importer. This tool was chosen for its enhanced control over export/import parameters, ensuring that all necessary data is transmitted, which leads to a more reliable outcome compared to alternative export

methods. The process generates a .dbz file, a format linked to the .duf DAZ Studio format and that encapsulates all information about the model's structure, textures, and rigging details.

The exported file is then imported into Blender, where the process continues using the same Diffeomorphic DAZ Importer addon, now operating within Blender. The file is imported using the "Easy Import DAZ" feature, and several configuration adjustments are made to this addon to ensure that the model is imported correctly with all the needed information. In this step, attention is given to accurately maintaining the model's morphs, textures, and the alignment of all related components.

In order to convert the morphs embedded within the G8F mesh and G8F eyelashes into workable shape keys, the Diffeomorphic plugin is utilized inside Blender. During this conversion process, the morphs are imported as separate elements, allowing for further individual manipulation. In the "Convert Morphs to Shape Keys" stage, the morphs from DAZ Studio are converted into shape keys, making it possible to control every facial bone in Blender. It is worth noting that this process is also done to the eyelashes and other face adjacent meshes that require moving, otherwise, they won't move as expected, even after the next step.

Subsequently, we integrate two additional addons: VRM for Blender and CATS. The CATS addon is employed to manage and consolidate the various meshes—comprising the body, hair, clothing, and eyelashes. By selecting all the individual meshes and executing the "Join All Meshes" function, the disparate elements are merged into a unified, cohesive model. The merging process is monitored to ensure that the final consolidated mesh exhibits no apparent issues, such as misaligned geometry or lost textures.
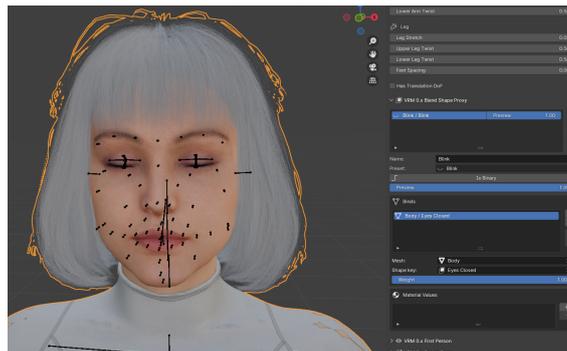
Figura 3. The blend shape proxy section in the VRM for Blender addon, showing the blink/close eye proxy being tested

In the VRM addon for Blender, the model's existing rigging is then mapped to the necessary bones in a VRM 0.x model. The imported morphs and expressions are also "bound" through the addon. This step involves associating key expressions—such as a smile, blink, and mouth open—with their corresponding shape keys, through a feature called "blend shape proxy". This feature enables configuration of all shape keys related to facial expressions and motions, further refining the model's ability to exhibit real-time interactions. Following this, testing of all facial expressions was undertaken to verify that each expression behaved correctly under various conditions, ensuring a seamless and

realistic user experience during live interactions.

Following the bone and shape key assignment, the Check VRM validation tool was run to ensure that the model complies with all VRM format requirements. The validation confirmed that there were no errors or conflicts, affirming that the model is fully prepared for further modifications and eventual export. Once the entire rigging and expression workflow was successfully validated, the model was exported in the VRM 0.x format. It is important to note that, while the VRM Plugin for Blender produces a fully functional VRM file, the current iteration of the plugin does not support integrated compression. As a result, the exported file remains relatively large, a limitation that is acknowledged in the context of its ongoing development.

Figura 4. The final model being tested in Warudo

The final model is now a fully functional VRM avatar, ready for deployment in VTubing and other virtual interaction applications. In the final phase of the workflow, the avatar was tested in Warudo, a widely used VTubing software. The test confirmed that the entire process, from model creation and import to rigging and export, was successful and compatible with the requirements for real-time facial tracking and virtual interactions.

4. Conclusion and Future Work

The conversion of realistic models from DAZ Studio to the VRM format is both possible and highly effective when using Diffeomorphic and Blender as intermediaries. While there are some limitations regarding facial expressions and file size optimization, the final result is a realistic and functional model. It is also important to address the need for improvements in materials and shaders to preserve realism while ensuring compatibility with the VRM format. The fact that many assets from the DAZ Store are paid, as well as comparisons with other conversion methods, still require further evaluation. In conclusion, the workflow using Blender can be quicker and less technical for those looking to an easier way to setup a VRM model.

5. Acknowledgement

References

Ball, C., Huang, K.-T., e Francis, J. (2021). Virtual reality adoption during the covid-19 pandemic: A uses and gratifications perspective. *Telematics and Informatics*, 65:101728.

Blender Foundation (2025a). Blender manual: History. `https://docs.blender.org/manual/en/latest/getting_started/about/history.html`. [Accessed: Apr. 13, 2025].

Blender Foundation (2025b). Blender manual: Shape keys introduction. `https://docs.blender.org/manual/en/latest/animation/shape_keys/introduction.html`. [Accessed: Apr. 13, 2025].

DAZStudio (2025). `https://en.wikipedia.org/wiki/Daz_Studio`. Wikipedia, [Accessed: Jun. 23, 2025].

Diffeomorphic (2025). Diffeomorphic daz importer wiki. `https://bitbucket.org/Diffeomorphic/import_daz/wiki/Home`. [Accessed: Apr. 13, 2025].

Fraser, A. D., Branson, I., Hollett, R. C., Speelman, C. P., e Rogers, S. L. (2022). Expressiveness of real-time motion captured avatars influences perceived animation realism and perceived quality of social interaction in virtual reality. *Frontiers in Virtual Reality*, 3:981400.

Freeman, G. e Maloney, D. (2021). Body, avatar, and me: The presentation and perception of self in social virtual reality. *Proc. ACM Hum.-Comput. Interact.*, 4(CSCW3):239.

Kim, D., Lee, S., Jun, Y., Shin, Y., e Lee, J. (2025). Vtuber's atelier: The design space, challenges, and opportunities for vtubing. In *Proc. 2025 CHI Conf. Human Factors in Comput. Syst. (CHI '25)*, pages 1–23, New York, NY, USA. Association for Computing Machinery.

Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., e Botsch, M. (2017). The effect of avatar realism in immersive social virtual realities. In *Proc. ACM Symp. Virtual Reality Softw. Technol. (VRST)*, pages 1–10, Gothenburg, Sweden.

Machado, E. C. (2023). Processo de criação de um modelo de vtuber para utilização em plataformas de streaming. TCC (graduação) - Universidade Federal de Santa Catarina.

Noor, N. M., Suaib, N., Ahmad, M., e Ahmad, I. (2018). Review on 3d facial animation techniques. *Int. J. Eng. Technol.*, 7:181–187.

Quantum, A. (2023). Cats blender plugin. `https://github.com/absolute-quantum/cats-blender-plugin`. GitHub repository, [Accessed: Jun. 15, 2025].

Saturday-06 (2024). Vrm add-on for blender. `https://vrm-addon-for-blender.info/en/`. [Accessed: Jun. 23, 2025].

Tambunan, O. N. e Setiawan, A. B. (2023). Factors in the growth of vtuber livestreaming entertainment phenomenon in japanese society. *LITERACY: Int. Sci. J. Soc., Educ., Humanit.*, 2(2):78–85.

Tang, M. T., Zhu, V. L., e Popescu, V. (2021). Alterecho: Loose avatar-streamer coupling for expressive vtubing. In *Proc. IEEE Int. Symp. Mixed and Augmented Reality (ISMAR)*, pages 128–137.

Vilchis, C., Perez-Guerrero, C., Mendez-Ruiz, M., et al. (2023). A survey on the pipeline evolution of facial capture and tracking for digital humans. *Multimedia Systems*, 29:1917–1940.

Wang, J., Qiu, Y., Chen, K., Ding, Y., e Pan, Y. (2023). Fully automatic blendshape generation for stylized characters. In *Proc. IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*.

Warudo (2025). Warudo: Unleash your virtual potential. `https://warudo.app/`. [Accessed: Jun. 23, 2025].