applied sciences

MDPI

Article

Application of Neural Radiance Fields (NeRFs) for 3D Model Representation in the Industrial Metaverse

Lidia Fabra, J. Ernesto Solanes *, Adolfo Muñoz, Ana Martí-Testón, Alba Alabau and Luis Gracia

Instituto de Diseño y Fabricación, Universitat Politècnica de València, 46022 València, Spain
* Correspondence: esolanes@idf.upv.es

Abstract: This study explores the utilization of Neural Radiance Fields (NeRFs), with a specific focus on the Instant NeRFs technique. The objective is to represent three-dimensional (3D) models within the context of the industrial metaverse, aiming to achieve a high-fidelity reconstruction of objects in virtual environments. NeRFs, renowned for their innovative approach, enable comprehensive model reconstructions by integrating diverse viewpoints and lighting conditions. The study employs tools such as Unity, Photon Pun2, and Oculus Interaction SDK to develop an immersive metaverse. Within this virtual industrial environment, users encounter numerous interactive six-dimensional (6D) models, fostering active engagement and enriching the overall experience. While initial implementations showcase promising results, they also introduce computational complexities. Nevertheless, this integration forms the basis for immersive comprehension and collaborative interactions within the industrial metaverse. The evolving potential of NeRF technology promises even more exciting prospects in the future.

Keywords: 3D reconstruction; metaverse; artificial intelligence

check for updates

Citation: Fabra, L.; Solanes, J.E.; Muñoz, A.; Martí-Testón, A.; Alabau, A.; Gracia, L. Application of Neural Radiance Fields (NeRFs) for 3D Model Representation in the Industrial Metaverse. *Appl. Sci.* **2024**, *14*, 1825. https://doi.org/10.3390/app14051825

Academic Editor: Shana Smith

Received: 4 February 2024
Revised: 17 February 2024
Accepted: 20 February 2024
Published: 23 February 2024

1. Introduction

In the industry, virtual spaces are increasingly being utilized for worker training, process simulations, and quality control. However, a major challenge associated with these spaces is that, at first glance, they may captivate attention but often seem to have more resemblance to a video game than real-life scenarios.

To enhance the realism of these virtual environments, advanced three-dimensional (3D) models capable of accurately reproducing textures and details of real-world objects are essential. While various mesh-based 3D modeling techniques demonstrate efficacy for simpler objects, challenges arise when dealing with more complex models involving reflections and flexible elements. The precision required to model these sophisticated features presents a substantial obstacle.

To confront this challenge, there is an urgent need to advance 3D modeling techniques tailored for virtual industrial environments. The difficulty in accurately capturing complex characteristics underscores the necessity for research and development in more sophisticated 3D modeling technologies. This pursuit aims to achieve a more precise and detailed representation, ultimately enhancing the practicality and effectiveness of virtual spaces in industrial applications.

1.1. Motivation

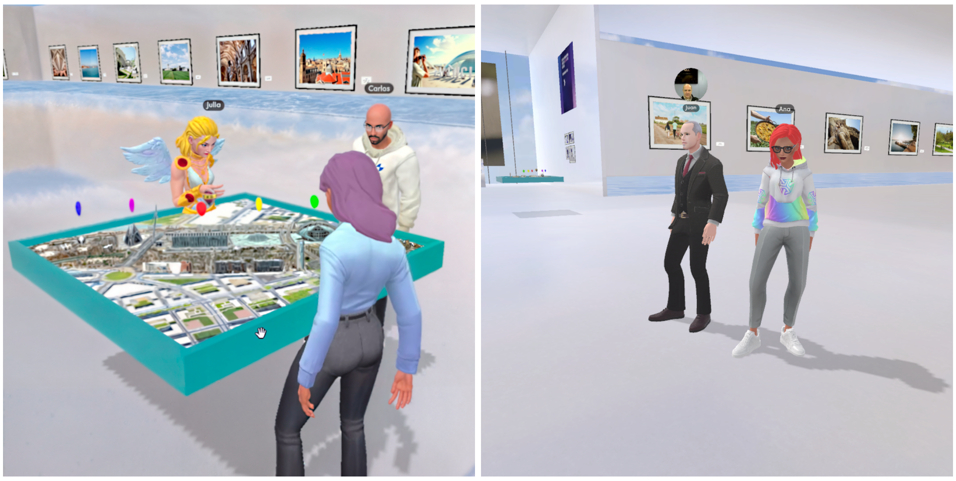
The implementation of six-dimensional (6D) models based on Neural Radiance Fields (NeRFs) in an industrial metaverse presents itself as an interesting choice for various reasons. Firstly, it represents an opportunity for technological innovation in various fields, such as industrial, heritage, or educational. By utilizing advanced technology, a detailed and realistic representation of objects and scenes in the environment is achieved. This has a significant impact on design and simulation processes, enhancing the way objects are represented and visualized.

Another relevant aspect is the ability to promote collaboration and communication among teams and professionals. By having a shared virtual representation of objects and scenes in an industrial metaverse, remote collaboration and joint decision-making are facilitated. This overcomes geographical limitations and allows different users to interact with and visualize the same data, which is particularly useful in industrial environments where collaboration between teams is essential.

1.2. Technological Context and Related Research

Virtual environments, computer-generated 3D simulations, enable the exploration and manipulation of digital worlds, offering advantages such as access to information from anywhere and the development of skills like creativity and critical thinking. Despite their benefits, it is crucial to consider potential negative effects, such as vision problems or software dependence, requiring careful design and attention to updates to ensure compatibility with specific goals [1]. In the industrial context, these virtual environments play a crucial role in training and decision-making, providing professionals with a secure and controlled environment for studying complex processes [2].

The concept of the metaverse, a shared virtual space for the creation and experience of interactive virtual worlds, has sparked controversy. Its increasing relevance in areas such as entertainment, education [3], and industry is highlighted [4]. In Figure 1, an example illustrates what a metaverse looks like and how avatars are visualized. Current advancements are crucial for the future of the metaverse, as mentioned in [5]. The emphasis lies in enhancing aspects such as reducing latency for seamless interactions, acquiring high-quality data to ensure an immersive and precise experience, and crafting virtual worlds that are more realistic and interactive, among other considerations.



(**a**) (**b**)

Figure 1. Example of interaction in the metaverse: virtual visit to Valencia city in the following. https://www.spatial.io/s/Visit-Valencia-639b11ab10f4070001b4a87b?share=8520243597075764882 (accessed on 19 February 2024). (**a**) Users interacting with elements in the metaverse. (**b**) Customized avatars.

In the industrial context, the metaverse offers advantages such as product visualization, worker training, and effective communication, but high-quality virtual models that accurately reproduce the physical and visual characteristics of real objects are required [6].

Currently, 3D technologies have experienced exponential growth, reshaping various sectors of modern society, including entertainment [7], education [8], medicine [9], and culture [10]. These technologies have fundamentally altered how humans perceive and engage with the world around them. The creation of 3D models, which includes detailed information about the geometry, textures, and physical properties of 3D objects, presents a notable challenge in the realms of computer graphics and artificial vision. Choosing the most fitting 3D modeling technique to fulfill the specific requirements of an application is not always a straightforward task. Despite the apparent simplicity of crafting a basic 3D model, producing an accurate and photorealistic computer model of a complex object continues to demand substantial effort [11]. This complexity amplifies when attempting to capture reflections and intricate details of the model.

In the realm of virtual reality, an innovative concept has captured the scientific community's attention. This concept is the NeRF, a promising technique for generating highly detailed and realistic virtual content (see an example in Figure 2). This method represents a scene using a fully connected deep neural network, known as a multilayer perceptron (MLP). Its input is a single continuous coordinate in five dimensions, i.e., spatial location (x, y, z) and viewing direction (θ, ϕ), and its output is the volumetric density and emitted radiance (light intensity) dependent on the view at that spatial location [12]. In other words, it is a technique that employs Artificial Intelligence to achieve highly realistic and detailed results by calculating the camera position and the light intensity it receives. Thus, by capturing images from different angles and positions of the object or scene to be modeled, an extremely realistic virtual recreation can be achieved.

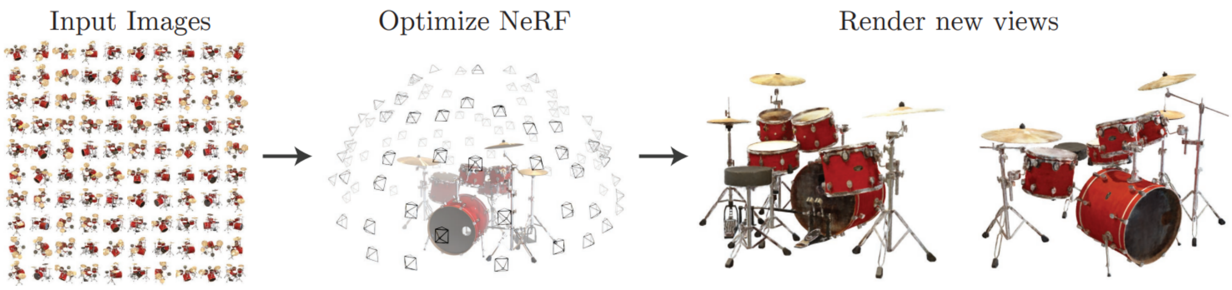


Figure 2. Graphical example of the NeRF method's operation [12].

Applying the NeRF technique in an industrial context within a metaverse opens up opportunities for more accurate representation of models than traditional meshing methods, for machinery, products, and manufacturing processes in a shared virtual space. One of the key features that sets this technique apart from traditional ones is its ability to capture reflections on materials. However, it is important to note that using NeRFs can be computationally intensive, resulting in significant processing overhead and requiring powerful hardware.

Several studies aim to enhance NeRF's capabilities in different areas. For example, in [13], efforts are made to improve the NeRF technique to capture scenes in motion, while, in [14], the goal is to achieve high-quality results in less time than the traditional NeRF method. As the NeRF technique is a developing technology, it remains a work in progress, implying that its effectiveness and applicability may improve over time. This leads to ongoing monitoring and adaptation as its limitations are investigated and addressed, with the potential emergence of new, similarly performing techniques in the near future.

The current literature highlights the increasing use of virtual environments in industrial settings, but it underscores the necessity for more advanced 3D modeling techniques to accurately represent complex objects and scenarios. While traditional methods have limitations in capturing reflections and intricate details, the emergence of Neural Radiance Fields (NeRFs) offers a promising solution.

1.3. Objectives and Main Contributions

The main objective is the generation of 6D models using the NeRF technique and artificial vision, with the aim of integrating realistic models into an industrial metaverse. By introducing these 6D models into an industrial metaverse, interaction and collaboration among users are encouraged. Multiple users can explore and manipulate the models simultaneously, enabling a more immersive experience and facilitating joint decision-making.

Furthermore, the implementation of this technology contributes to the efficiency and optimization of industrial processes. These models allow for a more precise representation of objects and their interaction with the environment compared to classic mesh-based 6D modeling, making it easier to identify improvements and conduct realistic simulations. This can have a significant impact on the productivity and competitiveness of companies.

Lastly, the use of NeRFs as a representation of 6D models opens the door to future applications and developments. This technology is constantly evolving, providing the opportunity to lay the foundations for the creation of new products, services, or tools in the industrial field through research.

1.4. Structure

The paper's structure is outlined as follows: the proposed application is developed in Section 2. Subsequently, Section 3 provides insights into the interface's usability and additional aspects. Finally, the paper concludes with a discussion and concluding remarks presented in Sections 4 and 5.

2. Development of a Prototype of Industrial Metaverse for Teaching/Learning Activities

2.1. Design Methodology

The methodology proposed in this work for designing an industrial metaverse is graphically depicted in Figure 3, which shows the systematic progression through key phases. It initiates with client specification gathering, involving the establishment of initial contact to comprehensively understand the client's specific needs and vision for utilizing the virtual environment. Following this, the collaborative preliminary design phase unfolds, where, based on the client's specifications, a joint effort is made to conceptually outline each element of the metaverse without delving into programming. Mock-ups are generated during this phase, providing a tangible representation for client feedback before transitioning to the subsequent development stage. After the client's validation, the selection of tools, software development kits (SDKs), platforms, and hardware is carefully undertaken, ensuring technical viability and effectiveness for the industrial metaverse's unique requirements. The implementation phase starts, presenting an alpha version for validation by testers and subsequently generating a beta version for client validation. This iterative process incorporates feedback, addressing necessary adjustments and refinements. The evaluation and delivery phase involves subjecting the beta version to a thorough evaluation during an agreed-upon period with clients and testers, culminating in the delivery of the finalized industrial metaverse. The methodology concludes with a prolonged validation period and continuous improvements, where the final version undergoes extended validation by users within the industrial environment, with ongoing refinements based on user feedback to ensure continuous adaptation and optimal efficiency.

Unity, renowned for its versatility and user-friendly interface, has been selected as the preferred software for this work [15]. The platform, excelling in the development of interactive applications in 2D and 3D, empowers developers to create immersive virtual environments and interactive applications effectively. With a robust game engine and a diverse set of tools, developers can seamlessly integrate innovative technologies like NeRFs to achieve a highly realistic 3D representation in the industrial metaverse under development. The specific version utilized for this work is Unity 2021.3.2f1. This choice over alternatives like Unreal Engine [16] is motivated by Unity's simplicity, efficiency in disk space usage, and compatibility with C# language. Unity's smaller file sizes reduce the demand for disk space and system resources, making it particularly suitable for the requirements of this work, even though Unreal Engine may be more attractive for larger projects with advanced graphics needs. In this specific context, Unity emerges as the most fitting and advantageous choice.

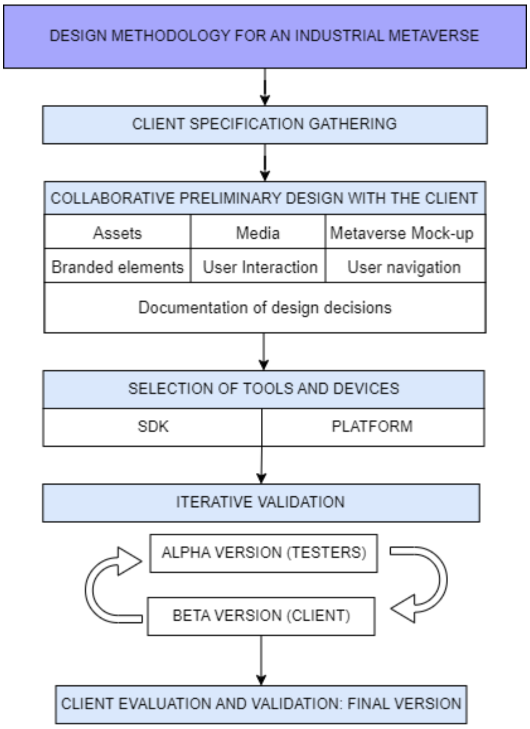


Figure 3. Methodology proposed for designing an industrial metaverse.

2.2. Development of the Metaverse Platform

In the development of the metaverse, a template focused on an industrial factory has been employed to shape the virtual space. For an immersive virtual reality experience, the Oculus SDK [17] software (version 62.0) has been chosen, specifically designed for interaction with Oculus virtual reality devices. This decision has enabled the creation of a realistic and engaging virtual environment, enriching user interaction within the virtual space.

Regarding user communication, the API from the Photon PUN2 library [18] has been utilized. This tool facilitates data synchronization and real-time interaction among users, enabling the creation of multiplayer rooms. Additionally, for voice communication in multiplayer applications using Photon, Photon Voice [19], a dedicated plugin facilitating real-time communication between users, has been integrated. In summary, the combination of Oculus SDK and Photon PUN2 has facilitated the development of an industrial metaverse with visually appealing and functional environments, supported by specialized tools for interactivity and real-time communication.

2.3. Development of Realistic Industrial Objects Based on NeRFs

In Figure 4, a detailed workflow is presented for creating immersive industrial metaverses in Unity using NeRFs. This groundbreaking methodology enables the generation of realistic 3D models from captured images, offering a novel approach to virtual representation.

To achieve a seamless integration of physical models into the virtual space, the workflow has been structured into *six steps* as detailed below, each crucial for the overall success of the process.

(S1) Capture of model images: Record a video around the object from different perspectives, followed by the extraction of the individual images for use in the NeRF model generation process.

(S2) NeRF Generation: Create a folder structure in Instant NeRF [20] and generate JSON file based on captured images for 3D reconstruction. This step culminates in generating the NeRF 3D model.

(S3) Exporting NeRF to Unity: Transfer the NeRF model from Instant NeRF to Unity. During this process, resolution is configured to ensure a visually adequate representation of the model in the Unity environment.
(S4) Creation of 3D texture: Assemble a mosaic of model images. This mosaic is imported into Unity as a 2D texture, which is then transformed into a 3D texture to be applied to the NeRF object.
(S5) Incorporating NeRF Object into Unity Scene: Create a material based on the 3D texture and apply the 3D texture to the NeRF object. A GameObject is configured to effectively represent the NeRF model within the virtual scene created in Unity.
(S6) Metaverse integration: Develop scripts that allow real-time interactions with the NeRF model. Specific components are configured to share the object in a metaverse, utilizing Photon PUN2 to synchronize and enable interaction among different users in the virtual environment.

This comprehensive workflow offers a step-by-step guide for seamlessly transitioning physical models into immersive industrial metaverse. It is important to note that the NeRF representation within Unity serves as a virtual copy, meticulously capturing the form of the real NeRF based on the acquired images. Leveraging this NeRF technology opens new possibilities for virtual representation, collaboration, and exploration in industrial environments.

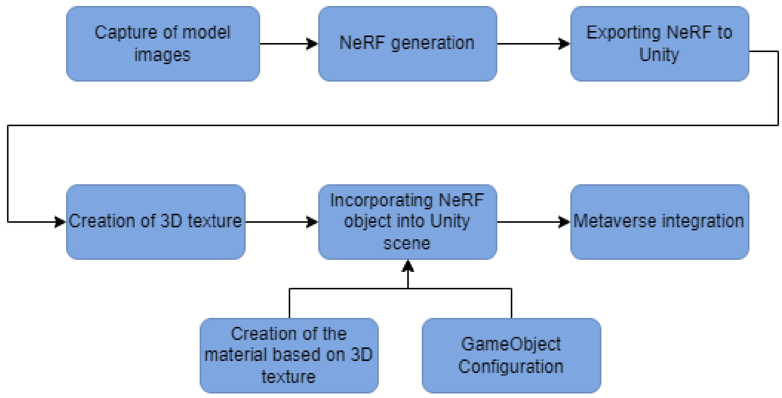


Figure 4. Workflow for creating a NeRF-based industrial metaverse.

Case Study: Obtaining a NeRF in Unity

This case study provides a practical demonstration of the step-by-step process outlined earlier for obtaining a NeRF in Unity. By following the detailed steps discussed, this case study offers a hands-on example that illustrates the seamless integration of NeRF into the Unity development environment.

In the initial phase of this case study, the process began with the capture of model images (S1). A video was recorded around the object from various perspectives and then individual images were extracted from the video. As shown in Figure 5, the images are placed in the "images" folder, which should be created within a directory where Instant NeRF is located.

Transitioning to the next phase (S2), the process involves obtaining a JSON file containing camera parameter values for each captured image. Colmap [21] is used in this work to generate this file. The subsequent step unfolds as Instant NeRF takes center stage, utilizing the acquired JSON file to adeptly reconstruct the NeRF model within a matter of seconds, as showcased in Figure 6.

Proceeding to the next step (S3), the subsequent phase involves acquiring the 3D texture for a Unity-ready representation of the obtained NeRF. In the Instant NeRF tools window, a pivotal option labeled "save PNG sequence" is available. This option facilitates the exportation of NeRF slices, which are cross-sectional views of the 3D volume. Prior to extraction, modifications can be made to tailor the NeRF to specific requirements, including volume trimming and resolution adjustments to meet the desired specifications for integration into Unity. These images serve as the foundation for creating a 3D texture within Unity, effectively producing a replica of the real NeRF volume in the Unity environment.

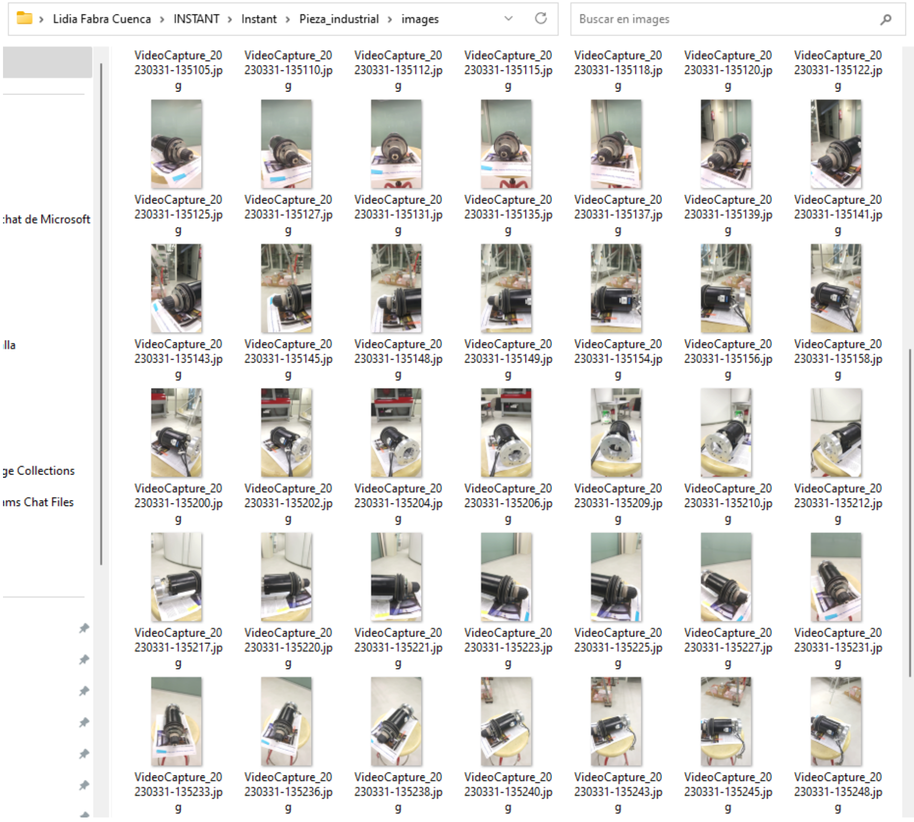


Figure 5. Folder with the images of the model.

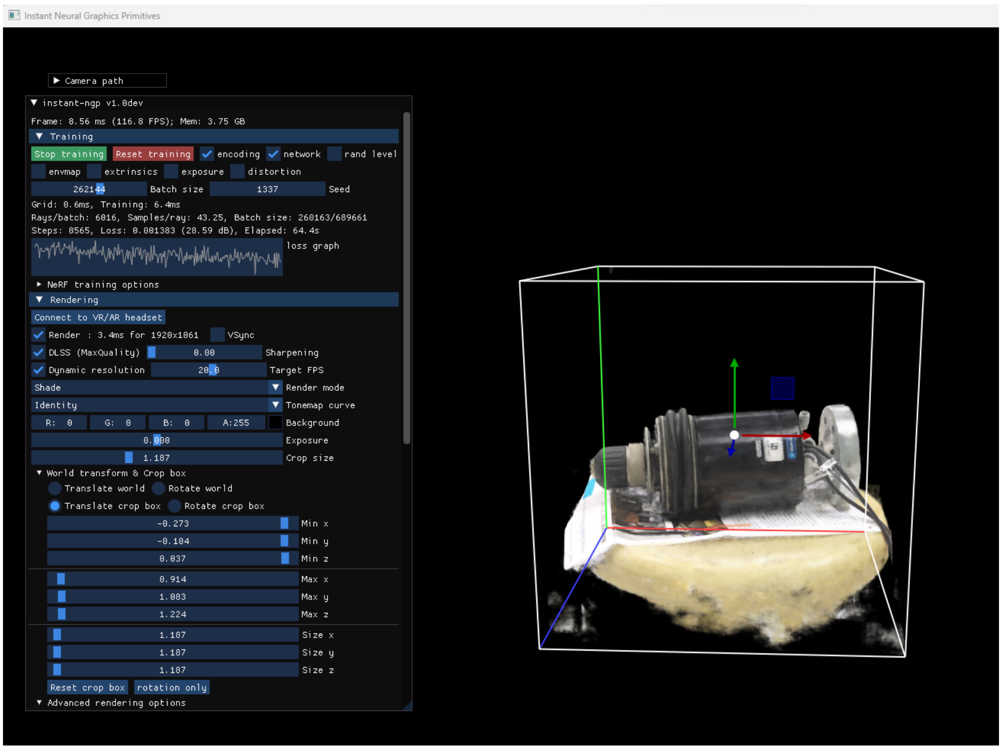


Figure 6. NeRF generated in Instant NeRF.

In the fourth phase of the process (S4), the creation of the 3D texture is started from the previously exported slices. The initial step involves generating a mosaic of the exported slices, shown in Figure 7. This mosaic will subsequently be introduced to Unity as a 2D texture, serving as the foundation for creating the required 3D texture for the NeRF object representation. Following the introduction of the mosaic to Unity, necessary adjustments such as size and alpha channel can be made, resulting in the desired 3D texture for the NeRF object, see Figure 8a,b. To generate the mosaic, Image Magick [22] has been employed, although any image editor can be utilized.

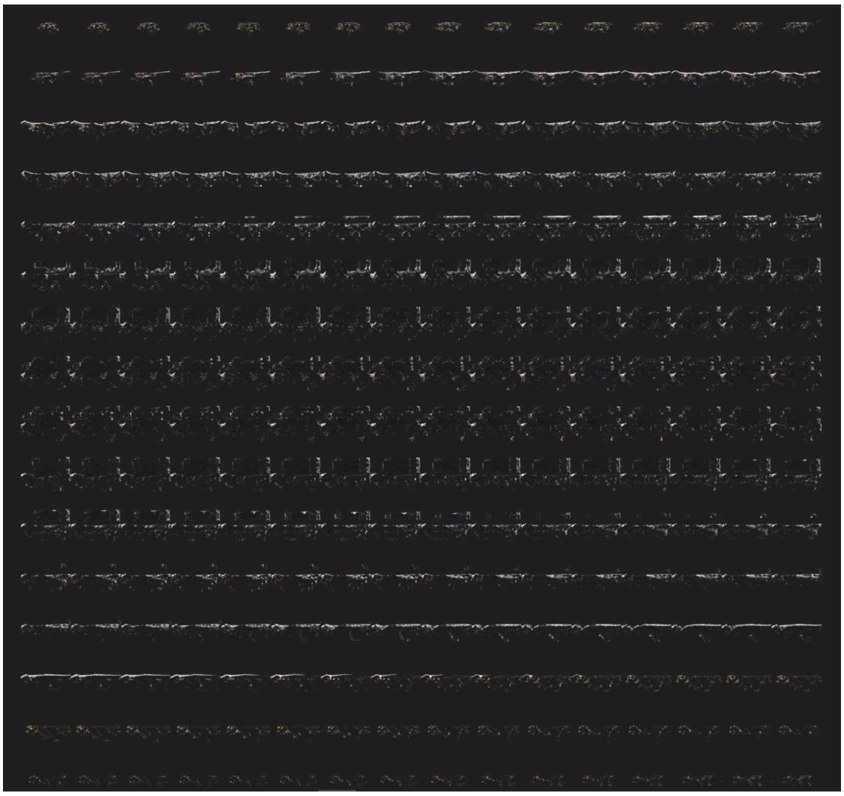
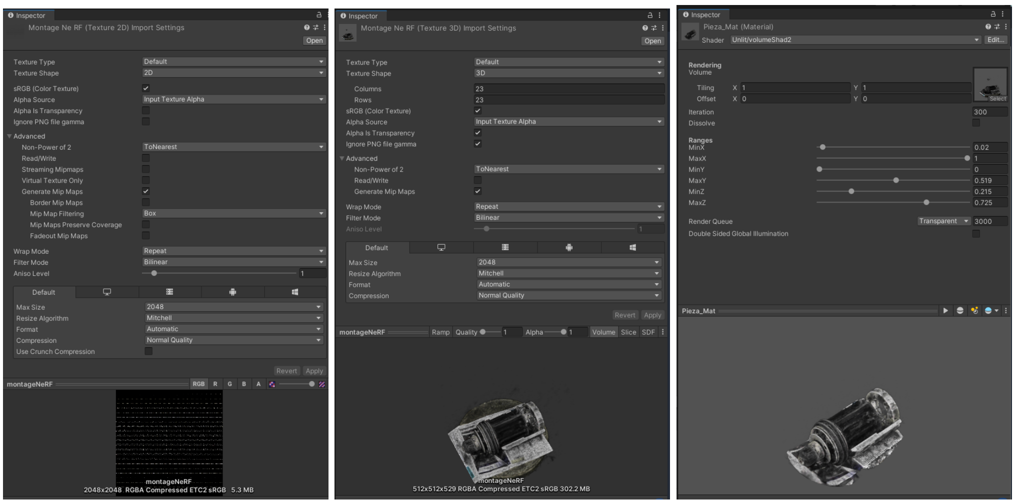


Figure 7. Example of a 512 × 512 image mosaic.



(**a**) (**b**) (**c**)

Figure 8. Inspector window. (**a**) Two-dimensional texture options. (**b**) Three-dimensional texture options. (**c**) Material with 3D texture.

The subsequent step involves the visualization of the object within the Unity scene (S5). To achieve this, it is essential to assign a material created from the 3D texture to a GameObject. The chosen shader for the material is "VolumeShad2" ([23]), selected for its high-quality visualization and straightforward implementation.

Upon material creation, the 3D texture is assigned and values are adjusted as needed until the model appears visually appealing, see Figure 8c. Following this, a new GameObject, specifically a cube, is created and the newly crafted material is assigned to it. Once incorporated into the scene, the NeRF model becomes visible within Unity. The outcome is showcased in Figure 9.

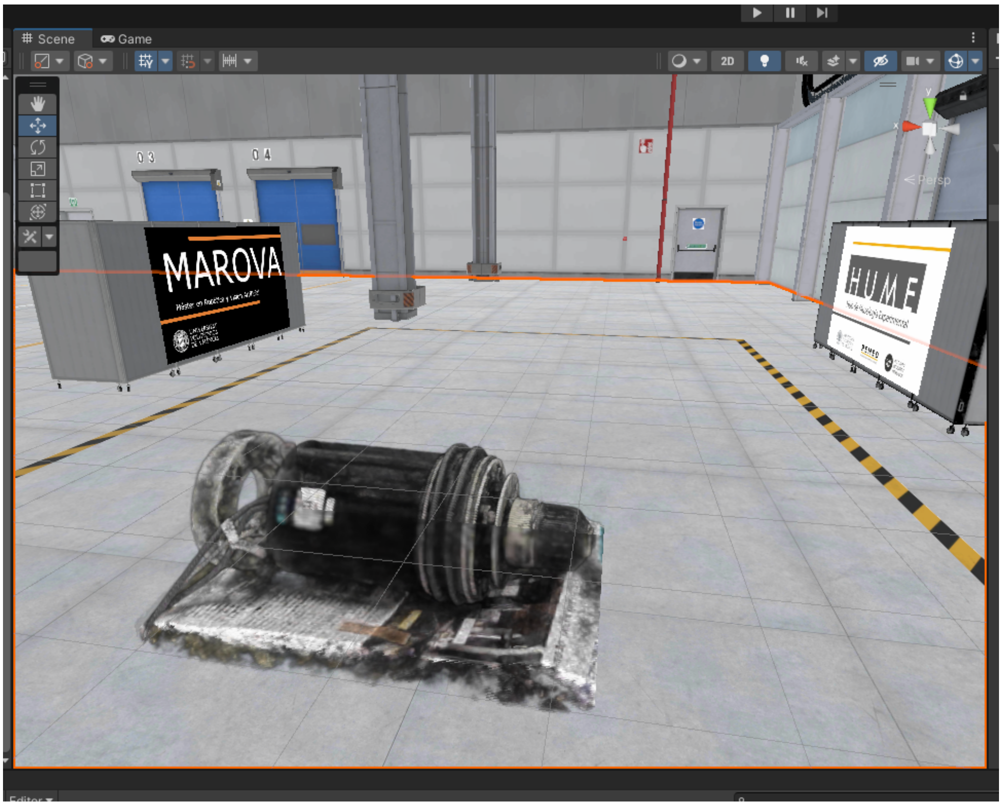


Figure 9. Representation of the NeRF model in scene.

In the final phase (S6), the focus shifts to configuring the GameObjects to become a shared resource accessible to all users within the metaverse. Specific components, as illustrated in Figure 10, are added to the GameObject. Additionally, various grab points have been strategically created on the GameObject, enabling users to interact from multiple angles.

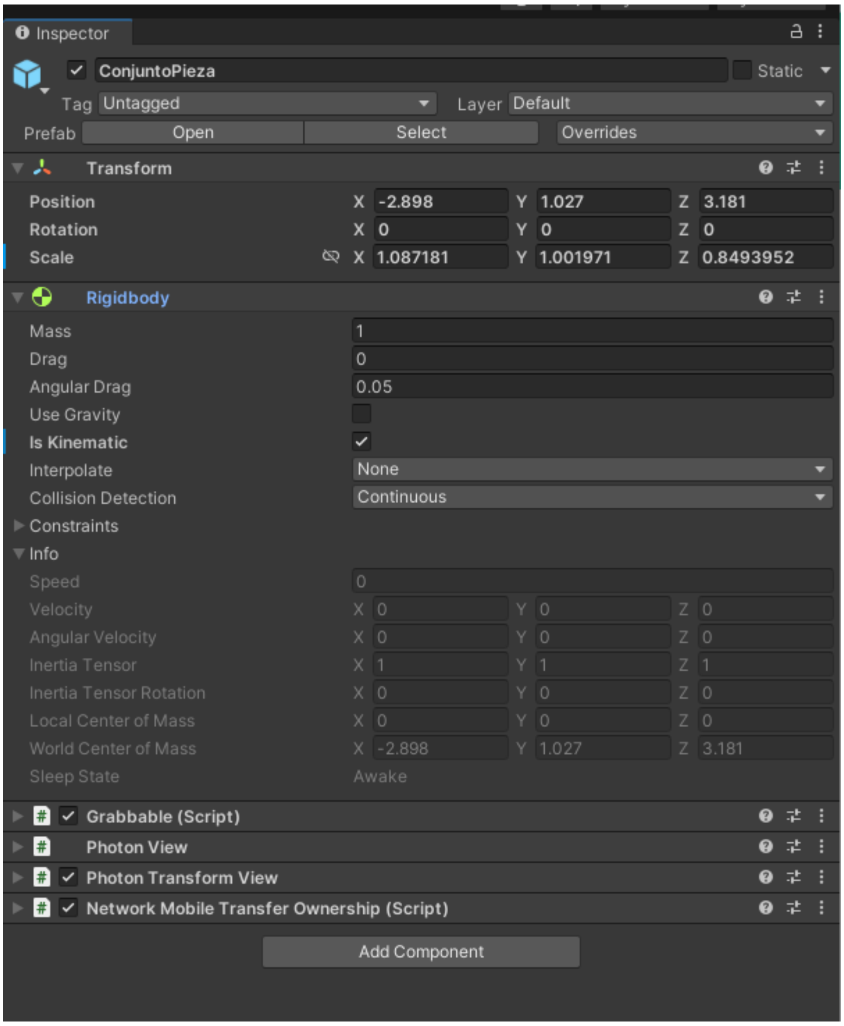


Figure 10. Inspector window with shared object components.

3. Results

3.1. Showcasing the Results through User Experience

To demonstrate the applicability of using NeRF-based 3D models for training activities, the industrial metaverse described in Section 2.2 is utilized, enabling the simultaneous connection of two users. Meta Quest 2 VR hardware devices are employed, equipped with hand tracking that allows users to interact with virtual objects naturally and keep their hands free of devices. Additionally, this environment enables voice communication, allowing operators to communicate during their training as they would in a real space.

It is worth mentioning that the industrial metaverse developed for this work offers two forms of user movement: natural translation, involving movement through the real space, and teleportation-based translation, enabling users to move even in confined real-world spaces. These two forms of movement are common in virtual spaces because, on the one hand, for true immersion in the virtual world, users should experience it similarly to how they would in the real world. On the other hand, due to real-world space limitations, teleportation is necessary to navigate through the entire virtual space.

The target objects for training can be seen in Figure 11: spiral pneumatic tube, Figure 11a; pneumatic motor for surface treatment operations with robots, Figure 11b; industrial gripper for robots, Figure 11c; vision system for studying surface defects using deflectometry, Figure 11d. These objects have been chosen for their industrial relevance and the challenge in obtaining a visually realistic model, particularly the vision system due to light reflections.

The results of the models viewed from the metaverse are depicted in Figure 12.

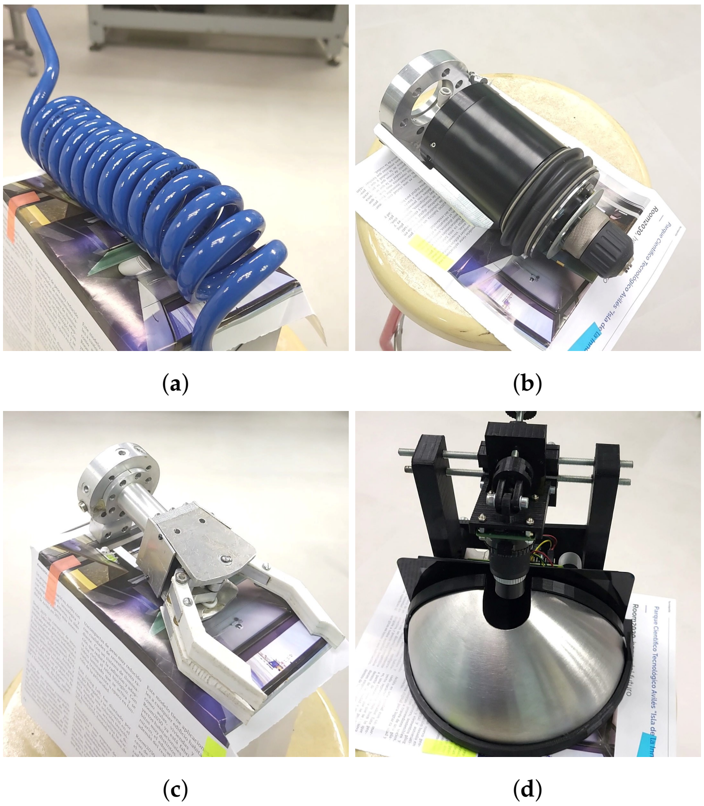


Figure 11. Real objects to be scanned and recreated using NeRFs. (**a**) Spiral pneumatic tube. (**b**) Driller. (**c**) Industrial clamp. (**d**) Vision system.

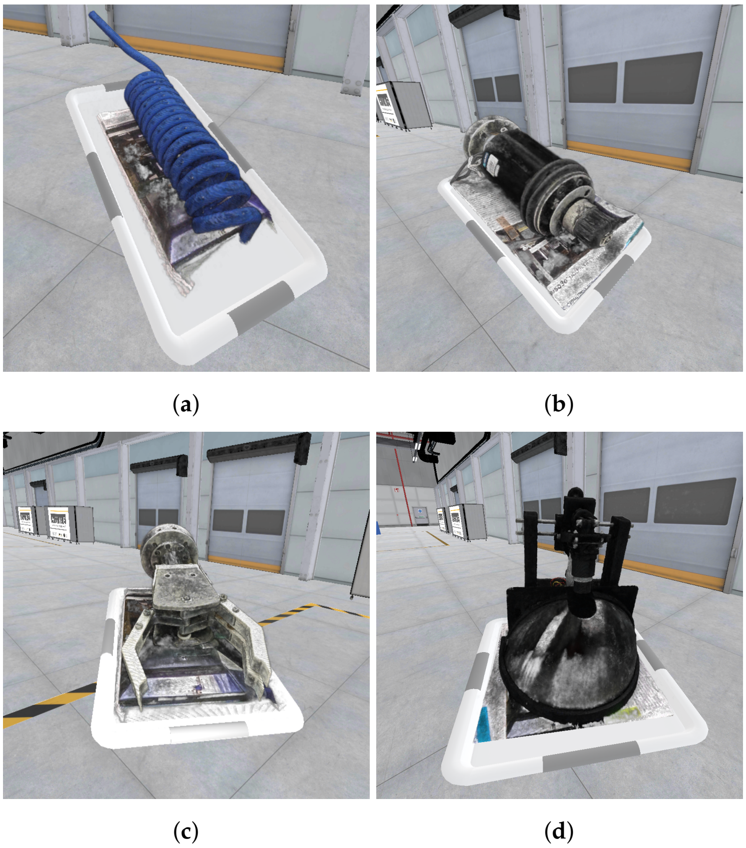


Figure 12. Results of the 3D models in NeRFs viewed from the industrial metaverse. (**a**) NeRF spring. (**b**) NeRF driller. (**c**) NeRF industrial clamp. (**d**) NeRF vision system.

To showcase the solution's results, a demonstration was conducted where two users entered the industrial metaverse developed (see https://media.upv.es/player/?id=186f82f0-613b-11ee-8548-1bcf94a5ee62) (accessed on 19 February 2024). This aimed to validate the outcomes and verify its proper functionality in real time. The demonstration provided a tangible view of the capabilities and possibilities offered by the industrial metaverse, emphasizing its real-time performance.

Figure 13 shows several captured moments from the user experience for various actions within the metaverse. In particular, Figure 13a depicts the hand gesture required for spatial navigation within the virtual environment. This not only broadens the age range of users but also grants complete freedom for natural interaction with virtual objects and avatars. In Figure 13b, one of the 6D models distributed across the industrial metaverse is showcased. Each model is positioned on a tray atop a pedestal, with the trays featuring grip points for user interaction. Figure 13c displays the available options panel within the metaverse, triggered when the user places both hands in front of their field of view. Pressing the green button activates a mobile device, allowing users to take photos to save images of the training session for later use. The red button allows users to exit the virtual space. Figure 13d,e illustrate user interaction with elements within the space: the first one shows user-to-user interaction, passing a model from one user to another, whereas the second one depicts a user comparing two 6D models, manipulating both simultaneously. Note that during the session, the users interact with each other by communicating verbally, explaining details of the actions and training provided. This can be observed at various moments in the demonstration video.

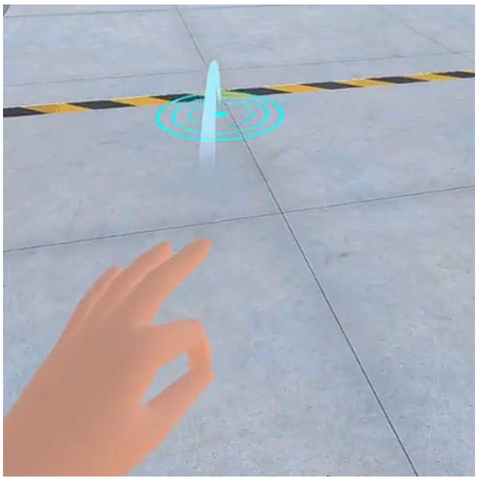
It is worth noting that, to enhance usability and ensure a satisfactory user experience when interacting with the proposed industrial models, a tray with only four grip positions has been implemented. These positions are indicated by gray handles on the tray, as illustrated in examples in Figure 12. When the user grasps the object at one of these positions on the tray, it automatically aligns so that the handgrip and the handle coincide, replicating the experience with a real handle. Additionally, the tray consistently positions itself to ensure that the model is always visible on top of it, facilitating continuous user engagement with the model (refer to Figure 13e). This design approach eliminates awkward viewing and gripping scenarios, preventing user distractions and maximizing the focus on the training content.

Qualitatively, users provided feedback on the application. Overall, it was noted that the application was user-friendly and did not necessitate prior knowledge, aside from initial instructions on navigating the environment. Socially, the realistic facial expressions of the avatars enhanced the interaction, fostering a greater sense of empathy. While interacting with objects was generally straightforward, there were occasional instances of objects being lost. This occurrence was attributed to limitations in the tracking system of the devices used.

3.2. Case Study: A Comparative Analysis of Models Based on NeRF and Models Based on Photogrammetry

In this section, 3D models obtained using standard photogrammetry techniques are compared with NeRF-based 3D models. The mesh-based models using photogrammetry were generated using RealityCapture [24] software (version 1.3), while the NeRF-based models were created using Nerfstudio [25] software (version v1.0.1). In both cases, 80 images of the objects in various positions and orientations were used as input, captured from a smartphone camera video.

Given the challenge of using a fair quantitative metric to express the similarity between the obtained 3D models and real-world systems in the case under study, where similarity involves not only morphology but also the realism of the object including reflections, transparency, textures, and colors, the comparison will be qualitative. Models will be described and compared through images from different viewpoints as necessary.



(**a**) Time instant 1 m 37 s of the video.

(**b**) Time instant 1 m 16 s of the video.

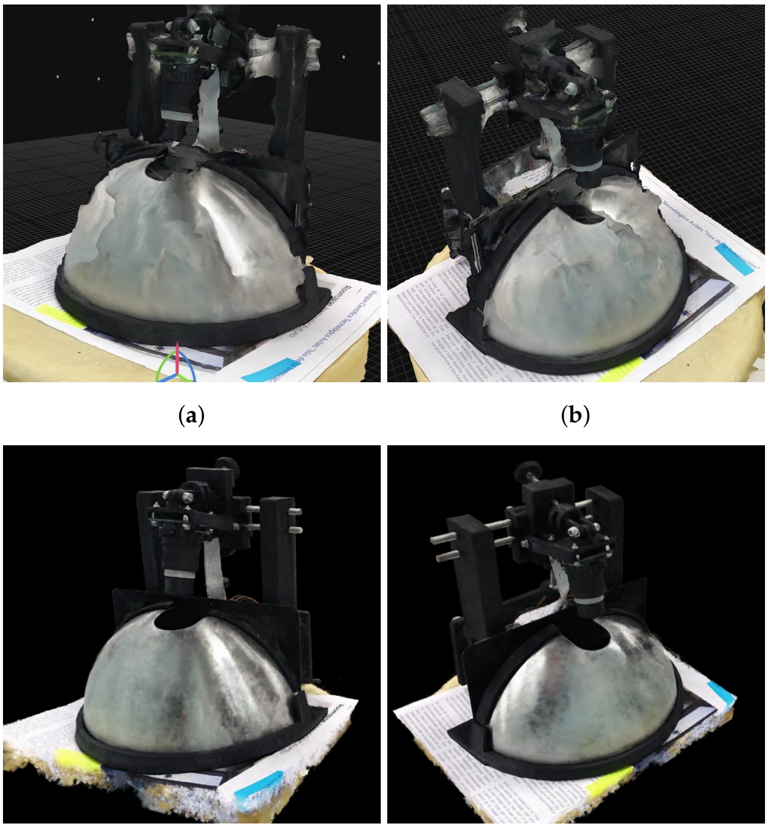
(**c**) Time instant 0 m 32 s of the video.

(**d**) Time instant 1 m 09 s of the video.

(**e**) Time instant 3 m 24 s of the video.

Figure 13. Demonstration test with users in the metaverse: (**a**) navigation mode; (**b**) industrial model; (**c**) options panel; (**d**) users' interaction; and (**e**) user interaction with different 6D models. Link to the video: users' experience video https://media.upv.es/player/?id=186f82f0-613b-11ee-8548-1bcf94a5ee62 (accessed on 19 February 2024).

Firstly, Figure 14 presents the obtained models of the vision system depicted in Figure 12d. Figure 14a,b show two different viewpoints of the mesh-based photogrammetry model obtained with RealityCapture. Figure 14c,d display two very similar viewpoints to their respective ones in Figure 14a,b of the NeRF-based model obtained with Nerfstudio. It is noticeable, first and foremost, that the NeRF model better preserves the morphology of the real target system. This is evident, for instance, in the two metallic gray support rods for the camera system at the top or in the semicircular dome, where even the mesh-based system has failed and generated holes in areas with light reflection. Additionally, regarding the visual realism of the model, it can be observed how the NeRF model resembles the real target system more closely than the mesh-based model. This is particularly clear in the dome area, where the reflection of light varies depending on the viewing angle.

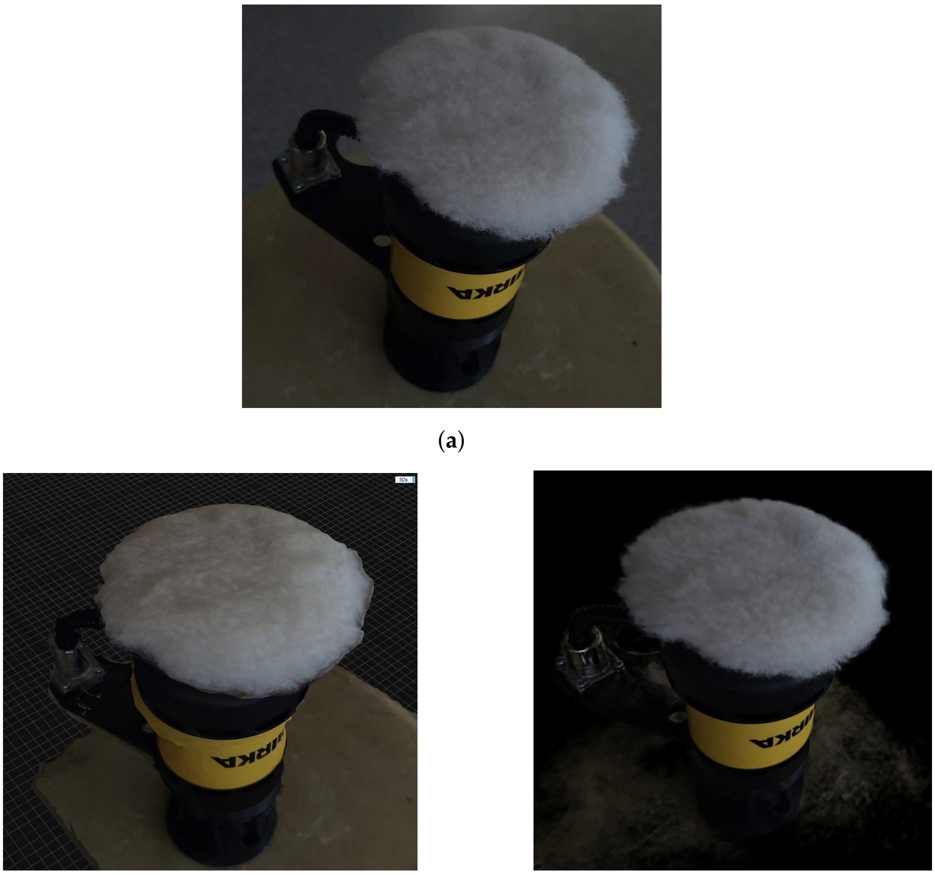


(**a**) (**b**) (**c**) (**d**)

Figure 14. Comparison between photogrammetry mesh-model and NeRF model. (**a**) Photogrammetry (view 1). (**b**) Photogrammetry (view 2). (**c**) NeRF (view 1). (**d**) NeRF (view 2).

While the morphology of the mesh-based model could be enhanced by introducing a greater number of input images captured with higher precision, the level of realism achieved would still be inferior to that of NeRF models.

To further highlight the advantages of NeRF-based modeling over conventional mesh-based techniques in industrial applications, Figure 15a showcases an electric automatic polisher commonly employed by industrial robots for surface treatment tasks, particularly in sectors such as automotive manufacturing. In the image, the tool features a non-rigid and fibrous component (depicted in white) that comes into contact with the surfaces being polished. Similar to the previous case study, the modeling process relies on 80 images captured from various angles to generate the models.



(**a**)

(**b**) (**c**)

Figure 15. Comparison between photogrammetry mesh-based model and NeRf model for the polishing tool. (**a**) Real object. (**b**) Photogrammetry mesh-based model. (**c**) NeRF model.

Figure 15c depicts the NeRF-based model obtained using Nerfstudio, illustrating detailed representations of both the rigid and non-rigid components with remarkable realism, closely resembling the actual object. In contrast, Figure 15b presents the mesh-based model derived from photogrammetry using RealityCapture. Here, noticeable mesh connections are observed in the non-rigid segment, resulting in a model that deviates from the real object compared to the NeRF-based counterpart.

3.3. Case Study: A Comparative Analysis of Models Based on NeRF and Mesh Approaches in Unity

This study conducts a comparative analysis between a model created using mesh-based approaches and the corresponding one based on NeRF in Unity. By examining their respective strengths, limitations, and application scenarios, this study aims to provide valuable insights into the optimal use cases for each method. The evaluation encompasses factors such as realism, computational efficiency, and ease of integration, shedding light on the practical considerations for developers and researchers in the field of virtual environments and 3D modeling. In Figure 16, the left side illustrates the mesh-based model obtained through LUMA AI [26], while the right side depicts the NeRF-based model.

The comparison reveals that factors such as realism, computational efficiency, and ease of integration are essential considerations for developers and researchers. As mentioned earlier, one of the advantages of the NeRF technique is its ability to capture reflections from real-world objects. Figure 16 demonstrates NeRF's capability to generate a piece with appropriate lighting, while the mesh-based model is incomplete and exhibits holes due to a lack of information for filling, attributed to the presence of light reflecting on the surface.

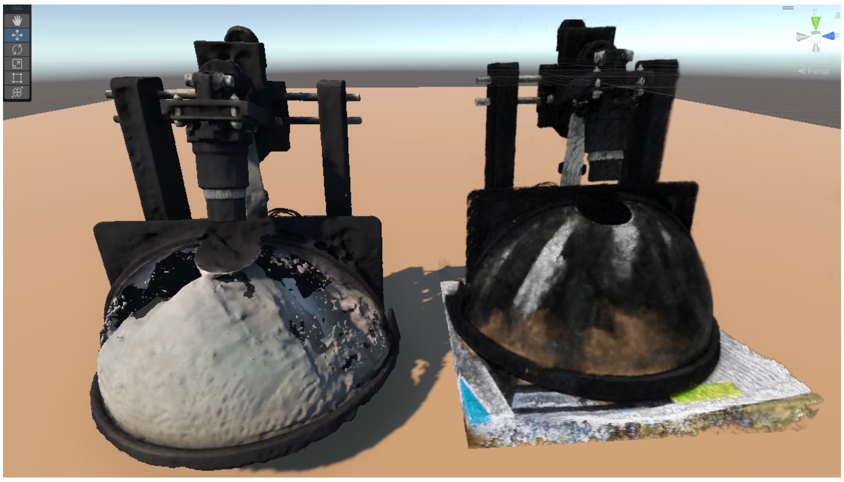


Figure 16. Comparison of mesh-based model (**left**) versus NeRF-based model in Unity (**right**).

3.4. Case Study: Real NeRF in Unreal Engine

This case study explores the visualization of a real NeRF model in the Unreal Engine. The study involves a comparison between the actual NeRF obtained from the LUMA AI plugin and a mesh-based model. The results of this comparison are depicted in Figure 17, shedding light on the performance and visual representation of a NeRF model within the Unreal Engine environment.

The superior visualization quality achieved with NeRFs in Unreal (as shown in Figure 17b) compared to its mesh approximation (depicted in Figure 17a) underscores its potential for creating realistic objects and immersive environments. This comparison distinctly showcases the effectiveness of NeRF technique in managing lighting conditions and reflective surfaces, thereby contributing to a significantly more authentic representation compared to traditional mesh-based models. The study offers valuable insights into the suitability of NeRFs for applications where accurate lighting and reflections are pivotal.

However, the success of the NeRF technique also emphasizes the importance of thoughtful consideration and resource planning when implementing it in projects with limited computational resources.

(**a**) (**b**)

Figure 17. Comparison of mesh-based model versus NeRF model in Unreal. (**a**) Mesh-based model. (**b**) NeRF model.

4. Discussion

The demonstration involving two users entering the industrial metaverse provided a real-time validation of the solution's outcomes (Section 3.1). The tangible view showcased the capabilities and possibilities offered by the industrial metaverse, emphasizing its real-time performance.

The comparison between real industrial models (Figure 11) and their counterparts obtained in the metaverse (Figure 12) confirms a high degree of realism. This affirmation underscores that the models obtained closely resemble their real-world counterparts, contributing to a heightened sense of realism within the metaverse.

As demonstrated in this study, models based on Instant NeRFs offer a more realistic visualization of objects compared to conventional techniques. This enhanced realism is particularly noticeable in objects featuring light reflections, fur, or fiber elements, as illustrated in the case studies presented in Figures 16 and 17. However, interaction with non-rigid objects such as plush toys may not provide a fully immersive experience, as manipulation does not result in deformation. Therefore, a potential avenue for further investigation is to introduce interactive capabilities where non-rigid models dynamically respond to user interaction, thereby significantly enhancing immersion.

It is worth noting that the computational cost of NeRF-based models is much higher than traditional mesh-based models. Additionally, their integration into virtual environment development applications is not straightforward, requiring the use of external plugins such as LUMA to import them into Unreal-type editors.

However, there are ways to work with NeRF approximations to reduce their computational cost. For example, one solution to reduce data computing consumption is by using the RT-NeRF technique [27]. As indicated in the study, NeRF's real-time performance on AR/VR devices is limited by uniform point sampling and the dense calculations required, despite its excellent image quality. RT-NeRF represents a significant advancement in real-time 3D rendering technology, particularly for Augmented and Virtual Reality. By effectively addressing NeRF's performance issues, RT-NeRF significantly improves speed without compromising visual quality.

Another solution could be to convert NeRF to the NVOL format [28], which stores NeRF in a compressed and optimized manner using a data structure called an octree. An octree divides the 3D space into smaller cubes, allowing for quick access to information for each point. Thus, an NVOL can display NeRF in Unreal Engine without the need to compute neural radiance on each ray. Table 1 shows the performance of NeRF and NVOL on some current Nvidia cards, using the Stable Diffusion benchmark with FP32 precision and a batch size of 16 [29]. The data are expressed in samples per second (SPS), indicating the training and inference speed of the models. As shown in the table, the use of NVOL results in a 10$\times$ acceleration in NeRF performance regardless of the card used. This is because NVOL reduces the computational load of neural radiance and better utilizes GPU memory and bandwidth.

Table 1. Comparison of the acceleration performance of NeRF and NVOL expressed in Samples Per Second (SPS) using the Stable Diffusion benchmark with FP32 precision and a batch size of 16.

Graphic Card	NeRF (SPS)	NVOL (SPS)
RTX 4090	1.2	12.0
RTX A6000	0.9	9.0
RTX 3090	0.8	8.0
RTX 3080	0.7	7.0
RTX A5000	0.6	6.0
RTX 3070	0.5	5.0

Another aspect to consider is the challenges associated with editing NeRFs. Unlike meshes, which offer explicit control over vertices and surfaces, NeRFs encode scenes in a continuous, volumetric manner, complicating tasks such as geometry and appearance editing [30].

A key challenge in NeRF editing lies in the absence of an object-centric decomposition, making it arduous to isolate and modify specific scene elements without comprehensive retraining or elaborate manipulation strategies [31]. Moreover, translating 2D edits into the 3D domain is non-trivial, necessitating advanced techniques to ensure accurate and localized modifications without unintended global repercussions [32].

Further complicating NeRF editing are the difficulties in establishing dense correspondences for tasks like texture transfer, a challenge attributed to NeRF's implicit nature and the lack of explicit geometric features [33]. Additionally, achieving view-consistent, artifact-free color editing demands innovative solutions to modulate color across all viewpoints, a simpler task in mesh-based models where texture maps can be directly altered [34].

Lastly, the computational demands for real-time editing and rendering of NeRFs significantly surpass those of traditional meshes. This necessitates research into efficient manipulation and rendering techniques to make NeRF-based editing feasible for practical applications [35].

5. Conclusions

This research marks a significant stride in reshaping industries through virtual reality and the creation of metaverse spaces. By enabling users in the industrial metaverse to collaboratively explore and study 6D models, this solution opens new avenues for immersive experiences and joint decision-making. Rigorous testing with diverse programs has been instrumental in evaluating the strengths and limitations of each method.

The findings highlight the substantial potential of this approach to drive the adoption of emerging technologies in industrial contexts. The development of a realistic industrial metaverse not only promises fresh opportunities in design, training, and collaboration, but also establishes a sturdy foundation for advancing the representation and visualization of industrial environments. This, in turn, serves as a crucial step towards industry transformation through the fusion of virtual reality and metaverse technologies.

This article signifies a pivotal move towards industry transformation, leveraging the NeRF technique for highly realistic 6D model generation. The incorporation of these models into the industrial metaverse enhances the capture of virtual object appearance and lighting, surpassing traditional meshing methods. Beyond improving simulation, design, and decision-making, this approach holds the promise of innovative applications, ensuring businesses stay at the forefront of technology and competitiveness in an increasingly digital and collaborative business landscape.

Author Contributions: Conceptualization, A.M.; Methodology, J.E.S.; Software, L.F.; Validation, A.M.-T.; Formal analysis, L.F.; Investigation, A.A.; Data curation, A.M.-T.; Writing—original draft, L.F. and A.A.; Writing—review & editing, J.E.S. and L.G.; Supervision, A.M. and A.M.-T.; Funding acquisition, A.M. and L.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been funded by the Spanish Government (Grant PID2020-117421RB-C21 funded by MCIN/AEI/10.13039/501100011033).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mostajeran, F.; Fischer, M.; Steinicke, F.; Kühn, S. Effects of exposure to immersive computer-generated virtual nature and control environments on affect and cognition. *Sci. Rep.* **2023**, *13*, 220. [CrossRef] [PubMed]
2. Umair, M.; Sharafat, A.; Lee, D.E.; Seo, J. Impact of Virtual Reality-Based Design Review System on User's Performance and Cognitive Behavior for Building Design Review Tasks. *Appl. Sci.* **2022**, *12*, 7249. [CrossRef]
3. Solanes, J.E.; Montava-Jordà, S.; Golf-Laville, E.; Colomer-Romero, V.; Gracia, L.; Muñoz, A. Enhancing STEM Education through Interactive Metaverses: A Case Study and Methodological Framework. *Appl. Sci.* **2023**, *13*, 10785. [CrossRef]

4. Dwivedi, Y.K.; Hughes, L.; Baabdullah, A.M.; Ribeiro-Navarrete, S.; Giannakis, M.; Al-Debei, M.M.; Dennehy, D.; Metri, B.; Buhalis, D.; Cheung, C.M.; et al. Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *Int. J. Inf. Manag.* **2022**, *66*, 102542. [CrossRef]
5. Bhattacharya, P.; Saraswat, D.; Savaliya, D.; Sanghavi, S.; Verma, A.; Sakariya, V.; Tanwar, S.; Sharma, R.; Raboaca, M.S.; Manea, D.L. Towards future internet: The metaverse perspective for diverse industrial applications. *Mathematics* **2023**, *11*, 941. [CrossRef]
6. Liu, S.; Xie, J.; Wang, X. QoE enhancement of the industrial metaverse based on Mixed Reality application optimization. *Displays* **2023**, *79*, 102463. [CrossRef]
7. Kontkanen, I. Using 3D techniques in video game concept art: Creating environmental concept art for a video game. BSc Thesis. Tampere University of Applied Sciences (TAMK). May 2021. Available online: https://www.theseus.fi/bitstream/handle/10024/501432/Iida_Kontkanen.pdf (accessed on 19 February 2024).
8. Ortiz, J.S.; Guevara, B.S.; Espinosa, E.G.; Santana, J.; Tamayo, L.R.; Andaluz, V.H. 3D Virtual Content for Education Applications. In Proceedings of the 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), Seville, Spain, 24–27 June 2020; pp. 1–5. [CrossRef]
9. Alhonkoski, M.; Salminen, L.; Pakarinen, A.; Veermans, M. 3D technology to support teaching and learning in health care education—A scoping review. *Int. J. Educ. Res.* **2021**, *105*, 101699. [CrossRef]
10. Montusiewicz, J.; Barszcz, M.; Korga, S. Preparation of 3D Models of Cultural Heritage Objects to Be Recognised by Touch by the Blind-Case Studies. *Appl. Sci.* **2022**, *12*, 11910. [CrossRef]
11. Remondino, F.; El-Hakim, S. Image-based 3D modelling: A review. *Photogramm. Rec.* **2006**, *21*, 269–291. [CrossRef]
12. Mildenhall, B.; Srinivasan, P.P.; Tancik, M.; Barron, J.T.; Ramamoorthi, R.; Ng, R. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *Commun. ACM* **2021**, *65*, 99–106. [CrossRef]
13. Pumarola, A.; Corona, E.; Pons-Moll, G.; Moreno-Noguer, F. D-nerf: Neural radiance fields for dynamic scenes. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 10318–10327.
14. Xu, Q.; Xu, Z.; Philip, J.; Bi, S.; Shu, Z.; Sunkavalli, K.; Neumann, U. Point-nerf: Point-based neural radiance fields. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 5438–5448.
15. Unity. Game Development Platform. Available online: https://unity.com/ (accessed on 19 February 2024).
16. UnrealEngine. Available online: https://www.unrealengine.com (accessed on 19 February 2024).
17. Quest, M. Oculus SDK Download. Available online: https://developer.oculus.com/downloads/package/oculus-platform-sdk (accessed on 19 February 2024).
18. Engine, P. Photon Unity Networking 2 Download. Available online: https://www.photonengine.com/pun (accessed on 19 February 2024).
19. Engine, P. Photon Voice Download. Available online: https://www.photonengine.com/voice (accessed on 19 February 2024).
20. NVlabs. Instant Neural Graphics Primitives. Available online: https://github.com/NVlabs/instant-ngp (accessed on 19 February 2024).
21. Colmap. Available online: https://colmap.github.io/install.html (accessed on 19 February 2024).
22. ImageMagick Studio LLC. Available online: https://imagemagick.org (accessed on 19 February 2024).
23. kwea123. volumeShad2.shader. Software Resource. 2021. Available online: https://github.com/MistyNight/NerfUnity/commit/918ecd1bcb00b4df6cb1fa48678505ce95ce9bec (accessed on 19 February 2024).
24. RealityCapture—3D Models from Photos and/or Laser Scans. Available online: https://www.capturingreality.com/ (accessed on 22 February 2024).
25. Tancik, M.; Weber, E.; Ng, E.; Li, R.; Yi, B.; Wang, T.; Kristoffersen, A.; Austin, J.; Salahi, K.; Ahuja, A.; et al. Nerfstudio: A Modular Framework for Neural Radiance Field Development. In Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Proceedings, Los Angeles CA USA, 6–10 August 2023. [CrossRef]
26. Luma AI. Available online: https://lumalabs.ai/ (accessed on 22 February 2024).
27. Li, C.; Li, S.; Zhao, Y.; Zhu, W.; Lin, Y. RT-NeRF: Real-Time On-Device Neural Radiance Fields Towards Immersive AR/VR Rendering. In Proceedings of the 41st IEEE/ACM International Conference on Computer-Aided Design, San Diego, CA, USA, 30 October–3 November 2022; pp. 1–9.
28. Volinga. Available online: https://volinga.ai/ (accessed on 22 February 2024).
29. GPU Benchmarks for Deep Learning | Lambda. Available online: https://lambdalabs.com/gpu-benchmarks (accessed on 22 February 2024).
30. Liu, S.; Zhang, X.; Zhang, Z.; Zhang, R.; Zhu, J.Y.; Russell, B.C. Editing Conditional Radiance Fields. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, 11–17 October 2021; pp. 5753–5763. [CrossRef]
31. Kobayashi, S.; Matsumoto, E.; Sitzmann, V. Decomposing NeRF for Editing via Feature Field Distillation. *arXiv* **2022**, arXiv:2205.15585. [CrossRef]
32. Mirzaei, A.; Aumentado-Armstrong, T.; Derpanis, K.; Kelly, J.; Brubaker, M.A.; Gilitschenski, I.; Levinshtein, A. SPIn-NeRF: Multiview Segmentation and Perceptual Inpainting with Neural Radiance Fields. *arXiv* **2022**, arXiv:2211.12254. [CrossRef]

33. Lan, Y.; Loy, C.C.; Dai, B. Correspondence Distillation from NeRF-based GAN. *arXiv* **2022**, arXiv:2212.09735. [CrossRef]
34. Wu, Q.; Tan, J.; Xu, K. PaletteNeRF: Palette-based Color Editing for NeRFs. *arXiv* **2022**, arXiv:2212.12871. [CrossRef]
35. Hedman, P.; Srinivasan, P.P.; Mildenhall, B.; Barron, J.; Debevec, P. Baking Neural Radiance Fields for Real-Time View Synthesis. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, 11–17 October 2021; pp. 5855–5864. [CrossRef]