Virtual Environment, Reconstruction of Vanished City: Focusing on the Serious Game *The Ruin Before*

Kyo Seok Lee[*,1], *You Seon Hahm*[1], *Sung Yup Lee*[1], *Hak Gu Kim*[1], *Hyung Gi Kim*[1], *Jin Wan Park*[1]

Graduate School of Advanced Imaging Science, Multimedia & Film, Chung-Ang University / Seoul, Korea
**Corresponding Author (kslee99@cau.ac.kr)*

Abstract: This paper introduces a serious game aimed at reconstructing war-damaged cities in a virtual setting, thereby providing an immersive experience of historical events. The game focused on the current conflict in Ukraine, reconstructing a city at a specific point in time, and allowing users to explore and interact with a meticulously detailed, historically realistic world. The principal processes in the game's development included the identification of a target city, acquisition of pertinent urban data from online sources, conversion of these data into a historically accurate style, and the design of non-playable characters (NPC) that display behaviors consistent with the era. Employing Meshy's Physically-Based Rendering (PBR) map generation technology, we accurately simulated the interplay between light and materials to authentically replicate the city's ambiance. The NPC were generated using ConvAI, a system based on ChatGPT that facilitates the incorporation of personality traits, background information, and event data, thereby animating them with contextual behaviors and narratives. The resulting game allowed users to interact with a meticulously reconstructed urban landscape and its people, providing a novel and instructive means of examining the effects of war through immersive virtual experiences.

Keywords: Virtual City Reconstruction, Generative AI, Serious Game, 3D Stylization, AI NPC

*Areas: ☐ Film, ☑ Media Art, ☐ Animation, ☐ Imaging Science(Check the box of your area)

Accepted for publication Oct. 02, 2024; published online November. 30, 2024; DOI: 10.15323/mint.2024.11.4.3.34/ISSN: 2799-8177.

1. Introduction

The devastation of cities owing to war has led to an irreversible loss of cultural heritage, historical landmarks, and community identity. As urban environments are destroyed, the stories and histories embedded within them are also destroyed. In response, this study focused on the virtual reconstruction of war-torn cities, allowing users to engage with these environments through immersive and educational experiences. By leveraging advanced technologies such as artificial intelligence (AI), 3D reconstruction techniques, and interactive nonplayer characters (NPCs), this study seeks to transform how historical education and cultural preservation are approached.

This study aimed to develop *The Ruin Before*, a serious game that meticulously recreates cities devastated by warfare. This virtual space enabled players to explore these cities as they existed during significant historical moments, offering a deeply engaging educational experience. The game not only serves as a tool for historical learning but also raises awareness of the profound human cost of war, highlighting the importance of peace and cultural preservation. The game focuses on Ukraine, which is currently embroiled in conflict, making it a highly relevant case study for demonstrating the potential of virtual city reconstruction as both an educational and empathetic tool.

The technical foundation of this study rested on extracting geospatial data and generating 3D models of cities using tools such as Google Maps and Cesium. These platforms provided high-resolution spatial data that served as the basis for recreating intricate cityscapes. To ensure that the cities are rendered accurately and reflected their historical context, Meshy's Physically-Based Rendering (PBR) technology was employed. PBR allowed the simulation of realistic interactions between light and materials, ensuring that the textures, lighting, and overall aesthetics align with the historical period being represented. The NPCs in the game were designed using ConvAI, which is a ChatGPT-based system. These NPCs exhibit behaviors, personalities, and dialogues that reflect the cultural and historical context of the period, allowing for dynamic and meaningful interactions with users.

The integration of AI-generated content and historical data presented a novel development pipeline for the creation of educational games, significantly reducing the time and cost typically associated with such studies. The

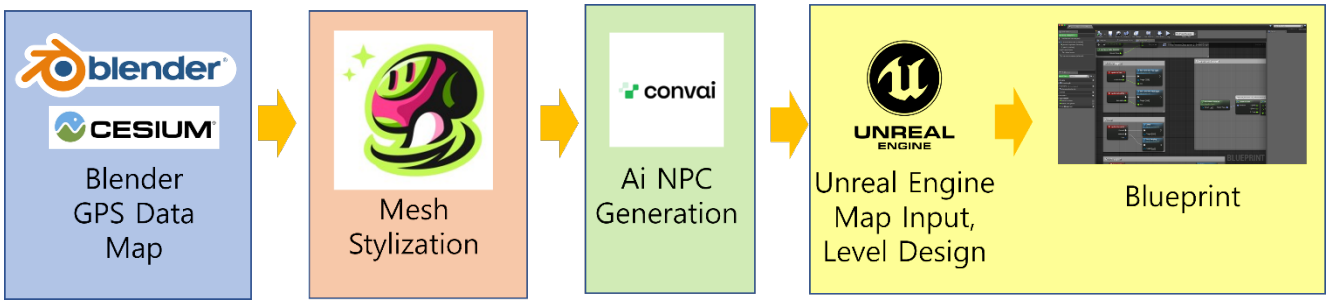

Figure 1. Overall procedure for making a virtual environment: Starting with processing geospatial data with Blender and Cesium, then mesh styling, AI-based NPC creation with ConvAI, and finally putting the maps together and designing the levels in Unreal Engine using Blueprint.

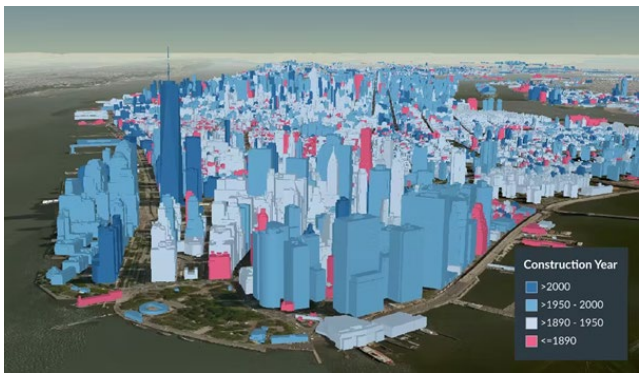

Figure 1. 3D geospatial data in each year made by Cesium

reconstructed cities in *The Ruin Before* were not simply static simulations. Instead, they offered a deeply interactive and immersive experience that encouraged users to explore history through a modern, interactive lens. Finally, this study demonstrated how virtual environments and AI can be harnessed to preserve cultural heritage, promote historical education, and contribute to both academic research and public awareness.

2. Related Works

Recent advancements in geospatial data acquisition, stylization techniques, and conversational AI have significantly enhanced the capabilities of virtual city reconstruction, interactive simulations, and gaming. LiDAR and multi-source digital map mashup schemes have improved the accuracy and efficiency of 3D data visualization, enabling the detailed reconstruction of complex environments. Stylization, driven by deep learning methods, such as Neural Style Transfer and Neural Radiance Fields (NeRFs), allows for the artistic and realistic transformation of both 2D and 3D data. In parallel, the evolution of conversational AI powered by models such as GPT and LaMDA has greatly improved NPC interactions, making virtual environments more immersive and contextually rich.

A. Geospatial data acquisition

Recent advancements in geospatial data acquisition and visualization technologies have revolutionized how 3D data is captured and utilized. These innovations are particularly crucial for applications, such as virtual city reconstruction, urban planning, and cultural heritage preservation.

LiDAR (Light Detection and Ranging) technology has become widely used owing to its high precision in capturing spatial structures. LiDAR provides high-resolution point clouds that are essential for reconstructing detailed 3D models. Kulawiak et al. [1] introduced a system for processing and disseminating LiDAR data through a 3D Web GIS framework using open standards such as LAS and 3D tiles. This system efficiently organizes and visualizes large-scale LiDAR datasets, making it suitable for urban reconstruction studies.

Multi-source digital map tile mashup schemes play a critical role in integrating various geospatial data formats into cohesive 3D models. Fan et al. [2] proposed a system utilizing CesiumJS to organize and load map tile data from different sources, ensuring compatibility between spatial references, such as WGS84 and Web Mercator. This method addresses the issue of tile deformation when merging data from multiple sources, which is crucial for the accurate reconstruction of cities.

Web-based 3D Geographic Information Systems (3D WebGISs) have evolved into platforms that enable real-time interaction and analysis of geospatial data. Auer and Zipf [3] demonstrated how browser-based 3D line-of-sight analysis powered by WebGL allows users to perform complex spatial analysis directly through the Web. This method is highly relevant for studies involving the reconstruction and simulation of urban environments.

In this study, cesium was the core technology used to reconstruct the virtual cities. Cesium is an open-source JavaScript library that enables real-time visualization of 3D geospatial data in web browsers without requiring any plugins. It is widely used across industries such as Geographic Information Systems (GIS), virtual reality (VR), urban planning, and military simulations because of its ability to efficiently visualize and interact with large geospatial datasets.

Cesium employs the 3D Tiles standard, which is optimized for handling large-scale 3D datasets by hierarchically structuring them to load only the necessary levels of detail at a given time. This makes it ideal for rendering complex environments such as entire cities in real-time. In the context of this research, Cesium allowed the seamless integration of data from multiple sources, including satellite imagery and terrain models, to create accurate and interactive 3D reconstructions.

B. Stylization

Figure 2. PBR elements predicted during the stylization of Meshy

Stylization, a technique often used in computer graphics and AI, refers to the transformation of the visual appearance of an object while preserving its structural content. This transformation can be applied to either 2D images or 3D data, enabling changes in style, such as converting an image into an artistic rendering or applying material effects to a 3D model. Stylization has become essential in various fields, including the visual arts, virtual simulations, and gaming, where both aesthetic quality and realism are desired.

The stylization of 2D images has gained prominence with the introduction of deep-learning techniques, particularly the Neural Style Transfer method developed by Gatys et al. [5]. This method separates the content and style of the two images and blends them into a new creation that maintains the original structure of one image while adopting the artistic characteristics of the other. More recent advancements such as StyleGAN [6] and AdaIN [7] have further refined this process, offering more control over style transfer and enabling a wide range of artistic effects. Although these methods have been highly successful for 2D imagery, extending them to 3D data presents new challenges because of the additional spatial dimensions involved.

In contrast to 2D images, stylizing 3D data involves transforming not only the surface appearance but also maintaining geometric consistency across multiple viewpoints. Recent studies, such as the Neural Radiance Fields (NeRF) framework [8] and its derivatives, such as StyleRF [9], have enabled zero-shot style transfer for 3D objects by learning how to apply styles to complex geometric forms. Unlike 2D stylization in which transformations occur on a flat plane, 3D stylization must account for changes in depth, lighting, and texture from different perspectives. The added complexity of 3D models requires more sophisticated algorithms to ensure that the style is consistently applied to the entire structure of an object, making it a more challenging task than 2D stylization.

In this study, we utilized Meshy, an AI-based tool specifically designed for 3D stylization and texturing. Meshy's PBR technology simulates realistic interactions between light and materials, ensuring that textures maintain visual consistency across different angles and lighting conditions. This approach contrasts with traditional 2D stylization because Meshy applies textures uniformly to the surfaces of 3D models,

maintaining the structural integrity and visual fidelity of complex environments. Using Meshy, we achieved a cohesive and immersive visual style for the virtual city environment in our project.

C. Coversational Deep Learning Models

Conversational AI has evolved significantly with the advancement of deep learning models, enabling machines to engage in natural dialogues with users. These models, which leverage natural language processing (NLP), have become integral to applications such as chatbots, virtual assistants, and interactive agents in customer service and gaming. The key developments in conversational AI include improved coherence and context awareness in interactions, allowing more meaningful and engaging user experiences.

The introduction of sequence-to-sequence (Seq2Seq) models by Sutskever et al. (2014) was a milestone in conversational AI. This model architecture uses an encoder-decoder mechanism to transform input sequences into a fixed-length vector and then decode them into output sequences, making it applicable in machine translation and dialogue generation. Despite this innovation, Seq2Seq models struggled to maintain a conversation context over multiple turns, often producing generic or repetitive responses, which limited their effectiveness in extended dialogues [10].

The release of the Transformer model by Vaswani et al. in 2017 revolutionized the field of NLP with its self-attention mechanism, allowing models to handle long-range dependencies more efficiently. This architecture laid the foundation for more powerful conversational models such as BERT and GPT. While BERT focused on understanding language through a bidirectional context, the GPT models, especially GPT-2 and GPT-3, excelled in generating coherent, contextually appropriate text. GPT-3, with its large-scale architecture, significantly improves dialogue generation, pushing the boundaries of what conversational AI can achieve [11].

Recent advances in conversational AI have focused on improving dialogue consistency and coherence across multiple conversational turns. Models such as DialoGPT, which is based on GPT-2, are specifically fine-tuned for dialogue, enabling better maintenance of context throughout longer conversations. These models, combined with reinforcement learning techniques, have further improved the accuracy and relevance of the generated responses [12]. Google's LaMDA, which was introduced in 2021, represents a major leap forward in open-domain conversations, allowing for more natural and contextually rich interactions than earlier models [13].

In this study, we utilized ConvAI, a conversational AI platform that builds on the capabilities of advanced language models such as the GPT. ConvAI allowed us to develop intelligent NPCs for our virtual environment, enabling users to engage in meaningful and contextually aware dialogues. These AI-driven NPCs played a crucial role in enhancing the immersive experience of our game, responding to user inputs through historically and culturally relevant conversations, which enriched the overall interactivity of the reconstructed virtual city.

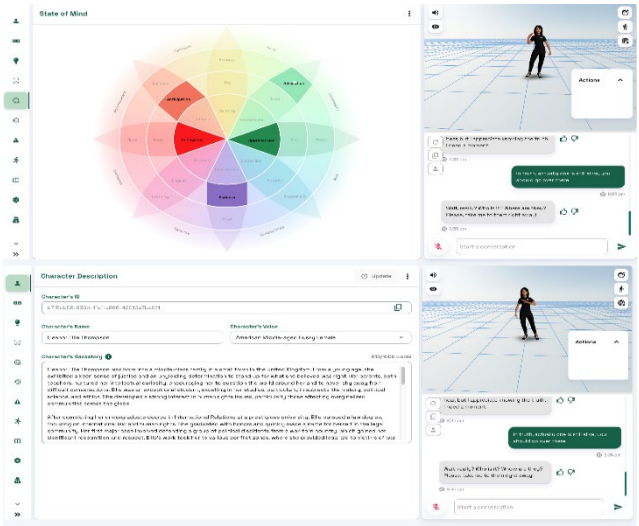

Figure 3. Creating a virtual non-player character via ConvAI. Establish the NPC's background context or psychological profile to facilitate organic dialogue.

3. Methodology

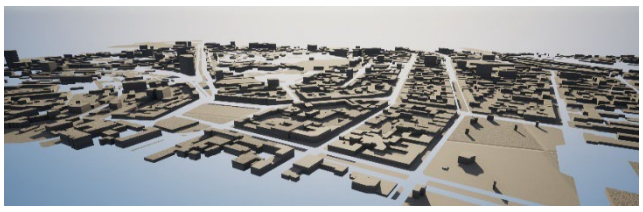

Figure 5. 3D geospatial data was taken out for the chosen city in Ukraine.

This study aimed to develop a serious game, *The Ruin Before*, that allowed users to explore virtual reconstructions of war-torn cities, enabling them to interact with historically significant environments and learn about the impacts of war. The research process was divided into four major phases: city selection, creation of NPC, 3D data extraction and stylization, and integration of these elements into a cohesive virtual environment. Each phase utilized advanced AI, 3D rendering, and simulation technologies to ensure the immersive quality of the virtual experience. A detailed explanation of each stage of the research process is provided below.

A. City Selection

The first step in this study was to select a city for virtual reconstruction. This process was guided by three primary criteria: historical significance, the availability of geospatial data, and the city's potential educational value. Given these criteria, Ukraine was chosen as the focal point of this study, focusing on the cities heavily impacted by the ongoing conflict. The selected cities provide a rich context for exploring the sociohistorical implications of war through an immersive virtual environment.

a) Historical Significance

This study aimed to reconstruct cities that had been severely impacted by war, giving users a chance to witness the aftermath of conflict and experience the gravity of historical events in a virtual space. Ukraine's cities, which have been devastated during the ongoing war, offer a perfect case study. These cities hold great historical value, and this study allowed users to engage with these significant locations in an innovative manner, thereby creating a bridge between the past and present. By selecting historically important cities, this study aimed to foster empathy and understanding of the consequences of war, ultimately promoting peace and awareness.

b) Availability of Geospatial Data

One of the key challenges in creating virtual city reconstructions was the availability of accurate geospatial data. To ensure a high level of precision in the 3D models, open-source platforms such as Google Maps and other GPS datasets were utilized. These platforms provided the necessary data to recreate a city's layout, including roads, buildings, and other key infrastructures. The accuracy of these geospatial datasets was critical for ensuring that the virtual environment mirrored real-world locations as closely as possible. Any gaps in the available data were supplemented with additional research and manual data entry to ensure completeness.

c) Educational Value

The educational value of reconstructed cities was a significant factor in the selection process. By enabling users to explore cities destroyed by war virtually, this study served as a powerful educational tool that highlights the destructive power of conflict and the importance of cultural preservation. Users are not merely passive observers. They engage with historical narratives to gain a deeper understanding of the social and emotional ramifications of war. This approach transforms the virtual experience into an interactive lesson about the importance of peace and historical awareness.

B. Creation of NPC

To enhance user experience and ensure that the virtual environment is engaging and dynamic, this study developed sophisticated NPCs. These NPCs were designed to reflect the cultural, social, and historical contexts of the selected cities, allowing users to interact with them in meaningful and informative ways.

Cultural and Historical Context: The NPCs were meticulously crafted to embody the cultural and historical characteristics of war-affected cities. Each NPC was designed with a unique background reflecting Ukraine's social and cultural environment at the time of the conflict. For example, NPCs representing rural citizens may discuss traditional agricultural practices or the challenges of sustaining a livelihood during wartime, whereas urban NPCs may reflect on the hardships of living under a siege or the destruction of their homes and communities. This level of detail in character design ensured that users could engage with historically accurate narratives that deepened their understanding of the period.

Conversational AI Integration via ConvAI: One of the most critical components of the NPC development process is the use of ConvAI, a conversational AI platform. ConvAI is powered by GPT-based natural language models that allow NPCs to engage in fluid and contextually appropriate dialogues with the users. The platform enabled the development of highly responsive NPCs capable of reacting to user inputs through historically and culturally accurate dialogue. Each NPC could engage in dynamic conversations and adapt its responses based on user actions and inquiries. This conversational capability added a layer of depth to the user experience, making NPC interactions more authentic and immersive.

Emotional and Narrative Depth: NPCs were designed with emotional depth and complex back stories, making interactions more engaging and impactful. For instance, NPCs could express grief over the loss of loved ones or describe the social and cultural impacts of the war on their daily lives. This study sought to create a more immersive and meaningful experience by allowing users to connect emotionally to these characters. This emotional connection enhanced the educational aspect of the study, helping users understand the human costs of war through direct interactions with characters.

C. 3D Data Extraction and Stylization

Figure 4. Applying stylization to results utilizing Meshy for extracted data

The accurate reconstruction of cities required detailed 3D data extraction and subsequent stylization to achieve a historically appropriate and visually immersive representation. This phase involved multiple steps, including the extraction of geospatial data, stylization to match the historical context, and the optimization of 3D models for real-time interaction.

Geospatial Data Extraction Using Cesium: Geospatial data from Google Maps and Cesium were extracted to capture the physical structure of the cities. Cesium's 3D Tiles technology was utilized to handle large datasets efficiently and render complex urban environments in real-time. The extracted geospatial data provided detailed models of buildings, roads, terrain, and infrastructure. This phase was particularly challenging because the initial 3D models were sometimes incomplete or lacked the resolution required for detailed reconstruction. In such cases, additional manual adjustments and optimizations were necessary to ensure that the models were accurate to the real-world locations.

Challenges in 3D Model Refinement: The Raw data extracted from Cesium and Google Maps often required further refinement. For example, the initially generated models were not always capable of reflecting the fine architectural details required to authentically represent historical cities. Some buildings appeared as overly simplified structures, which diminished the realism of the environment. To resolve these issues, the team used detailed reference materials, including photographs, maps, and historical documents, to manually enhance 3D models and add the necessary architectural details.

Stylization with Meshy's PBR Technology: Once 3D models were extracted and refined, Meshy, an AI-based stylization tool, was used to apply appropriate textures and lighting to the virtual environment. Meshy's PBR technology simulated realistic interactions between light and materials, ensuring that the textures of buildings, streets, and landscapes were visually consistent from all angles and under different lighting conditions. This process was particularly important to ensure that the virtual city retained a sense of historical accuracy and visual realism. For

Figure 5. Place styled 3D objects within a virtual environment using Blender.

instance, buildings damaged by war were given textures that reflected the destruction—crumbling walls, debris, and charred façades — whereas lighting effects captured the atmosphere of the city during and after the conflict.

D. Integration of Virtual Environments and NPC

The final phase involved the integration of stylized 3D models with interactive NPCs to create a cohesive and immersive virtual environment. This phase was crucial to ensure that the user experience was seamless and engaging.

The virtual environment was constructed using Unreal Engine, which allowed for real-time rendering of the reconstructed cities and the seamless integration of NPCs. Unreal Engine provided the computational power necessary to handle large-scale urban environments while maintaining smooth performance and high visual fidelity. The advanced rendering capabilities of the engine allowed realistic lighting, shadows, and reflections, further enhancing the immersive quality of the environment. Users were able to navigate the city freely, interact with the NPCs, and explore various historical locations in detail.

NPCs powered by ConvAI were embedded in the virtual environment to dynamically interact with users. These characters acted as guides, storytellers, and even participants in historical events, offering rich contextual information about the city's history, culture, and the impact of the war. The AI-driven dialogue system ensured that NPCs could respond to user input in real-time by adapting their dialogue and behavior based on the user's actions. This interactivity allows for a personalized user experience, with NPCs reacting differently depending on the user's choices and creating multiple narrative pathways and outcomes.

The seamless integration of the 3D environment with interactive NPCs resulted in an immersive user experience that was both educational and emotionally resonant. Users could explore the reconstructed city at their own pace, engage in meaningful conversations with NPCs, and participate in historical events that unfolded in a virtual space. The combination of historically accurate visuals, detailed NPC interactions, and real-time exploration provided users with a powerful tool for learning about the impact of war and the importance of cultural preservation.

4. Result

The game developed in this study aimed not only to provide an immersive experience but also to evoke deep emotional engagement and introspection in players. Interactions with NPCs, especially the war orphan Sofia, were designed to explore the psychological effects of war and how individuals cope with trauma. By crafting these interactions, the game emotionally connects the players to the storyline and enhances its educational value by fostering empathy and reflection.

In the game, the players encounter a war orphan in a city ravaged by bombings. Sofia pleads with the player to help her find her missing parents. As the players navigate the desolate landscape, they again encounter the same orphan as an adult, still searching for answers. This storyline illustrates the long-term psychological effects of war, such

Figure 6. Interact with post-war situations including stylized objects and non-playable characters in immersive virtual environments.

as trauma, grief, and enduring hope for closure, while prompting the player to consider the human toll of conflict.

As the players move through the city, they encounter a variety of objects, such as abandoned buildings, shattered windows, burnt-out vehicles, and personal items, such as toys and family photographs scattered across the landscape. Each of these objects is deliberately placed to convey a sense of loss or destruction, serving as emotional markers that guide the players through the war-torn environment. For example, a child's toy found amidst the rubble near the city center directly connects to the narrative of the orphaned child, emphasizing the theme of lost innocence and highlighting the personal cost of the conflict.

NPCs are also carefully positioned to enhance the emotional impact. An orphan child is encountered in a significant location in the heart of the bombed city. The setting immediately evokes the gravity of the situation, and the interaction with the child adds emotional depth, anchoring the player to the narrative. Rather than being scattered randomly, NPCs are strategically placed to maximize their emotional resonance, particularly in locations holding historical or personal significance. For instance, encounters near monuments or homes destroyed by war make NPCs feel like part of the story, providing a deeper connection between the player and the environment.

Environmental storytelling plays a critical role in player immersion. The bombed-out cityscape and derelict structures are not just background elements, but integral parts of the narrative. As the player moves through the ruined city, visual cues, such as graffiti and destroyed landmarks, serve as reminders of the chaos and fear experienced by the city's residents. These elements subtly reinforce the war orphan's journey and provide the players with a deeper understanding of the city's history, immersing them in a world where every detail helps tell the story.

ConvAI allows NPCs to react dynamically to player inputs, making each interaction feel personal and grounded in the environment. For example, a war orphan's voice reflects her surroundings, carrying a tone of despair that mirrors the destruction around her. The player's choices in conversation influence the tone and direction of each interaction, strengthening the emotional bond between the player and NPC. The deliberate placement of NPCs in emotionally charged locations enhances the psychological impact of a storyline. The reunion of the orphan as an adult, set in a desolate part of the city, visually portrays the passage of time and the enduring effects of the conflict. This transformation from a hopeful child to a disillusioned adult reinforces the narrative's emotional gravity.

5. Conclusion

This study successfully developed *The Ruin Before*, which is a serious game that reconstructs war-torn cities in a virtual environment, allowing users to explore and interact with historically significant locations. This study aimed to create an immersive educational tool that highlights the impact of war and the importance of cultural preservation. By combining geospatial data acquisition, AI-based NPC creation, and 3D rendering technologies, we provided users with a deeply engaging and educational experience.

The reconstruction of Ukrainian cities was based on accurate geospatial data extracted using the Cesium, which enabled the creation of realistic 3D environments.

Challenges related to data resolution and model refinement were overcome through manual optimization, ensuring a faithful representation of cities. Meshy's PBR technology was used for stylization to provide visually consistent and historically accurate textures.

The NPCs developed using the ConvAI platform were a key feature of this study. These AI-driven characters engaged users in meaningful conversations and provided cultural and historical contexts. Their ability to dynamically react to user inputs enhanced the immersive qualities of a virtual environment.

The Ruin Before demonstrated the potential of using advanced AI and 3D technologies to create interactive educational tools. This study offers a unique way to explore history and experience the emotional and cultural significance of cities affected by war. This study opens new possibilities for virtual learning environments and promotes a deeper understanding of the impact of war, emphasizing the importance of peace and cultural heritage preservation.

6. Future work

Several challenges were encountered during the study, which necessitate further research to improve the performance and realism of the virtual environments and NPCs.

Consistent 3D Rendering and Stylization for Large-Scale City Data: One of the key technical challenges encountered in this study was the inability to handle large-scale city data for 3D rendering and stylization. To address this issue, the city models were split into smaller parts, allowing for manageable processing. However, this method introduced inconsistencies in the rendering and stylization across the divided sections. Future work should focus on developing techniques that enable seamless and consistent 3D rendering and stylization in large urban environments. Research on optimized tiling systems or AI-driven methods that can manage the visual integrity of large datasets without compromising performance is crucial.

Improving NPC Voice and Movement Realism: Another issue encountered in this study was the lack of alignment between the NPC's voice and its appearance or age, as well as unnatural body movements. The NPCs generated using ConvAI sometimes exhibited voices that did not match their visual representations or age groups, and their physical animations lacked fluidity. Future research should explore methods for generating more contextually appropriate voices for NPCs using advanced voice synthesis technologies to ensure that the voices of NPCs are realistic and aligned with their visual traits. In addition, the development of more natural and responsive animations for NPCs is necessary to create fluid and realistic interactions between users and characters in a virtual environment.

References

[1] M. Kulawiak, M. Kulawiak, and Z. Lubniewski, "Integration, processing and dissemination of LiDAR data in a 3D web-GIS," ISPRS Int. J. Geo-Inf., vol. 8, no. 3, pp. 144, Mar. 2019.

[2] M. Auer and A. Zipf, "3D WebGIS: From visualization to analysis. An efficient browser-based 3D line-of-sight analysis," ISPRS Int. J. Geo-Inf., vol. 7, no. 7, pp. 279, Jul. 2018.

[3] J. F. Fan, T. Y. Hu, H. X. He, et al., "Multi-source digital map tile data mashup scheme design based on Cesium," Natl. Remote Sens. Bull., vol. 23, no. 4, pp. 695-705, 2019.

[4] P. Cozzi and K. Ring, 3D Engine Design for Virtual Globes, AK Peters/CRC Press, 2011

[5] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2414-2423, 2016.

[6] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4401-4410, 2019.

[7] X. Huang and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," in Proc. IEEE International Conference on Computer Vision (ICCV), pp. 1501-1510, 2017.

[8] B. Mildenhall, P. P. Srinivasan, M. Tancik, et al., "NeRF: Representing scenes as neural radiance fields for view synthesis," in Proc. European Conference on Computer Vision (ECCV), pp. 405-421, 2020.

[9] M. Zhang, M. Li, J. Zhang, et al., "StyleRF: Zero-shot style transfer for neural radiance fields," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7818-7827, 2023.

[10] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," Advances in Neural Information Processing Systems (NIPS), pp. 3104-3112, 2014.

[11] A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention is all you need," Advances in Neural Information Processing Systems (NeurIPS), pp. 6000-6010, 2017.

[12] A. Radford, J. Wu, R. Child, et al., "Language models are unsupervised multitask learners," OpenAI Technical Report, 2019.

[13] Y. Li, H. Su, X. Shen, et al., "DialoGPT: Large-scale generative pre-training for conversational response generation," arXiv preprint arXiv:1911.00536, 2019.

[14] Q. V. Le, A. Chowdhery, Z. Devlin, et al., "LaMDA: Language models for dialog applications," Google AI Research Paper, 2021.

Acknowledgement

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Graduate School of Metaverse Convergence support program(IITP-2024-RS-2024-00418847) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation).

Biographies

Kyo Seok Lee: He (14.06.1999) received a B.S. degree in Medical IT from Eulji University, Seongnam, South Korea in 2023. He is currently pursuing an M.S. degree at the Graduate School of Advanced Imaging Science, Multimedia & Film (GSAIM), Chung-Ang University, Seoul, South Korea. His research interests include deep learning, and adversarial attack.

You Seon HAHM:
2022-2024; Chung-Ang University GSAIM Department of Technology Art
2024-present; Chung-Ang University Department of Immersive Media/Animation Contents
2015-2019: Seokyeong University. College of Arts Department of Film and Media (B.F.A)
Research interest; Unreal Engine, 3D Animation, Digital Media, Film Directing

Sung Yup Lee:
2022-2024; Chung-Ang University GSAIM Department of Technology Art
2017-2022: Seokyeong University. College of Arts Department of Film and Media (B.F.A)
Research interest: Generative AI, Stable Diffusion, Digital Media, AI Art

Hak Gu Kim: He (31.03.1988) received B.S. and M.S. degrees from Inha University, Incheon, South Korea and a Ph.D. degree from the Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea. He was a Postdoctoral Researcher at École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He is currently an Assistant Professor at the Graduate School of Advanced Imaging Science, Multimedia & Film (GSAIM), Chung-Ang University, Seoul, South Korea. He has been conducting research on a wide spectrum of artificial intelligence and image processing topics. His research interests include immersive content analysis/representations, multimodal learning, and the vulnerability of neural networks.

Hyung Gi KIM:
2004-present; Professor, Chung-Ang University GSAIM
Department of Technology Art
Research interest; Media art, AI Art, Projection Mapping, LED Video
Sculpture, Media Facade

Jin Wan PARK:
1995; Chung-Ang University, Computer Engineering
1999; Pratt Institue, CGIM
2003-present; GSAIM, Chung-Ang University, Professor