IIAI Open Conference Publication Series
IIAI Letters on Informatics and Interdisciplinary Research
Vol.006, LIIR453
DOI: https://doi.org/10.52731/liir.v006.453

Proposal for a PBL-type Learning Model Using 3D Gaussian Splatting

Kazuya Murata [*], Tatsuya Horita [†]

Abstract

This study proposes a next-generation problem-based learning (PBL) model for cultural tourism and education based on 3D Gaussian Splatting (3DGS), a technology that emerged in 2023. Compared with conventional photogrammetry and Neural Radiance Fields (NeRF), 3DGS enables high-definition and high-speed 3D rendering and real-time display on smartphones and stand-alone VR. In this study, this technology is not limited to a mere means of viewing, but is integrated with augmented intelligence (AI) to create a mechanism for dynamically presenting contextualized stories and learning materials to users. Specifically, using cultural resources in the Kirishima region of Kagoshima Prefecture as the subject matter, we propose a PBL model that integrates the entire process of filming, model optimization, guide presentation by AI, fieldwork, and sharing in the metaverse. In this process, the model was verified from various perspectives, including rendering delay, frame rate, learning effect, and ripple effect on the local economy. Students and tourists with prior virtual experience tended to significantly improve their learning depth and time spent at the site. Additionally, by converting the generated content to NFT and connecting it to local currencies and cultural heritage preservation funds, we propose a circular learning and economic ecosystem.

Keywords: 3D Gaussian Splatting, Neural Radiance Fields (NeRF), photogrammetry, metaverse, cultural tourism, problem-based learning

1 Introduction

In recent years, cultural tourism and education has involved creating an environment in which people can experience, learn, and move to almost the same depth without having to travel there. When photography and moving images became widespread, supporting technology focused on refining two-dimensional information; however, today, as smartphones and high-performance GPUs permeate our living spaces, we are at the threshold of an era in which we can “carry” landscapes and remain in three dimensions and call as required.

However, the weight of data construction and display processing has prevented this realization. Photogrammetry requires a large amount of post-processing after the image is captured, whereas in NeRF, the number of learning and inference operations hinders real-time operation. These limitations were fundamentally rewritten in 2023 by 3D Gaussian Splatting (3DGS), which directly describes objects and landscapes not as points, but as a superposition of tiny 3D Gaussian distributions that are instantly combined by a GPU raster analyzer. It is synthesized

[*] Kagoshima Woman's College, Kagoshima, Japan
[†] University of Kochi, Kochi, Japan

instantly. Consequently, smooth rendering at approximately 100 frames per second is possible even on commercial notebook PCs and standalone VR, and even the weathering marks of outdoor ruins and the delicate coloring of religious architecture can be reproduced with a texture as close to the real as possible.

This study focuses on the fact that high-speed rendering technology is not merely a means of viewing, but can be combined with augmented intelligence (AI) to generate and present a contextualized narrative on the spot for each tourist or learner. The vast amount of historical material stored in digital archives is a mere pile of data without the perspective of reading and understanding. With AI providing suggestions and 3DGS guaranteeing visual realism, the fragments are linked in the viewer's mind as a single story, simultaneously, generating understanding and emotion. Furthermore, when used in conjunction with a metaverse infrastructure, multiple people in remote locations can share the same three-dimensional space and enhance their discussions while listening to different explanations, realizing "collaborative exploration."

For example, the Kirishima region of Kagoshima Prefecture is a multilayered story of a community's history of living in harmony with eruptions, formation of a hot spring culture, and mountain worship, with its own mythology. Tourists visiting the area are compelled to move around in the limited time they have to spend in the area without being able to fully understand this complex background. However, if the details of the crater and shrine buildings recorded by the 3DGS are explored in advance in the metaverse and the AI guide's dialogue enriches their own questions, the field visit becomes a dense field-learning experience, which is a series of confirmations and discoveries. This intensity is the key to increasing the intention to revisit the area and spread the word-of-mouth; from an educational perspective, it has the significance of enriching materials for cross-curricular problem-based learning (PBL) in the classroom.

2 Background

The wave of digitization began in the 2000s with the organization of books and photographs as PDFs; however, in the 2020s, even three-dimensional information flowed into the network, and the scale of digital archives swelled exponentially. Digital archiving of cultural heritage has significantly progressed in recent years; however, most of the world's cultural heritage remains undigitized. For example, in Europe, tens of millions of cultural heritage materials are available online on platforms such as "European," however, this is only a fraction of all the cultural heritage materials, and most are being digitized. In addition, quantity does not necessarily guarantee quality. Without a high-speed display that preserves local texture and translates it into a form that allows users to immediately understand the meaning contained in the data, a vast amount of data is merely a structured warehouse.

The focus here is on spatial computing devices, such as head-mounted displays and smart glasses, which have rapidly become popular. Although VR/AR devices are steadily spreading, their users remain a small fraction of the world's population, even in 2024. For example, it is estimated that the number of VR/AR headset shipments worldwide in 2023 was approximately 10 million units, and if we consider this as a percentage of the population, the utilization rate of VR/AR devices will remain low. When used in the field, many 3D models have insufficient resolution, making text unreadable or the texture is high, however, the frame rate drops, causing disorientation, among other problems that occur frequently. Although photogrammetry can preserve the details of masonry at a high resolution, the number of shots and the mesh restoration process are heavy, and the resulting data can grow to several gigabytes. NeRF uses volumetric rendering to increase realism, but integrates hundreds of rays per frame, and there are scattered examples where the speed is only approximately 20 frames per second on a stand-alone VR machine. Education and tourism demand "uninterrupted motion on any device, light enough to be edited and annotated on the fly as needed," and none of the existing techniques have been able to cross that last line.

This stagnation is overcome by 3DGS. The reversal of the idea of laying out a three-dimensional space with an extremely small Gaussian distribution eliminated the need to convert point clouds into a mesh and the need for dense ray integration. Methods based on Neural Radiance Fields (NeRFs) can generate extremely high-quality novel viewpoint images; however, their training requires significant computational costs and time. Therefore, 3DGS is a promising technology that can overcome the limitations of NeRF in terms of speed. However, careful comparisons from multiple perspectives, including image quality and range of applications, are necessary when evaluating "superiority" or "inferiority." The visualized data are memorable and encourage subsequent exploratory actions. Thus, an environment in which "data speaks for itself" will be created, and archives are expected to change from warehouses to interlocutors.

Metaverse-related markets are projected to grow significantly, reaching several trillion dollars by 2030. One study estimates that metaverse has the potential to generate up to $5 trillion by 2030. However, these figures are hypothetical estimates that depend on assumptions, and whether the market will reach several trillion dollars by 2030 remains uncertain. The cyclical process in which travelers become interested in a virtual space in advance, check it on site, and reinterpret it after returning home is not limited to the conventional concept of tourism, but will also spread to school education. If students without any fieldwork experience practice survey techniques in a 3D model before heading out to the field, they will be able to make high-quality observations and record them, even with limited training. The data obtained in the field were returned to the archive as Gaussian data for the next set of learners and tourists. Thus, the "cyclical narrative ecosystem" created is the future image that this study envisions.

3 Purpose

This study aims to connect the agility and sophistication of 3DGS to both tourism and education and implement "a circuit that transforms the experience itself into a social value." In the field of cultural asset preservation, scanned objects are often stored in databases and not used. In the field of learning, an increasing number of attempts have been made to present teaching materials in augmented reality; however, the heavy processing required by teachers to use materials photographed the night before in class the next day poses a problem. This barrier is lowered by 3DGS which can prepare high-fidelity models in tens of minutes to hours, enabling an environment in which teachers and tour guides can construct and share "stories that are there now" in near real time. When augmented intelligence is combined, each Gaussian point cloud can automatically be accompanied by historical annotations and an academic context, allowing users to learn and be moved as they interact with the model through their gaze and speech.

This study specifically aims to develop a mechanism to sublimate a mere 3D model for viewing into a "knowledge stock that generates learning tasks and is updated by itself." In the prototype system to be developed for the Kirishima region of Kagoshima Prefecture, data captured by students and tourists using smartphones will be sent to the cloud, where it will appear in a virtual space as an optimized Gaussian point cloud within a short period. Users enter the space via VR or a browser, interact with an AI guide to explore the history of the structure, and ask questions related to eruptive activities and religious beliefs. Students with fieldwork experience can practice surveying and classification procedures in the virtual space before visiting the site. Upon their return, they can add additional measured data to enrich the model and deliver it to the next visitor.

In this study, several indicators were set to simultaneously evaluate technical effectiveness and ripple effects on education and tourism. The minimum requirements were that the drawing delay be within 20 milliseconds in terms of visual comfort and that the display frame rate be maintained at 45 frames on a mobile device. Regarding learning effectiveness, the goal is to improve the level of understanding by at least 15%, using pre- and post-tests.

The academic significance of this project is twofold. First, there are few precedents for integrating rendering technology and educational engineering to quantify the visual quality of 3D models and learning outcomes within the same framework. Second, it incorporates the economic cycle of the generated content and presents a mechanism at the implementation level, whereby digital assets stimulate incentives for both the local community and learners. From a social perspective, even local organizations without expensive surveying equipment or large-scale servers can obtain technology and operational models that enable them to instantly disseminate local resources worldwide.

4 Precedent Research

The conversion of cultural properties into three-dimensional structures underwent a major step forward with the generalization of photogrammetry in the 2010s. At that time, the mainstream method was to use a single-lens reflex camera, surround the subject with several hundred photographs, and use post-processing software to extract feature points and reconstruct the mesh. The completed model boasted high accuracy; however, texture seams and discontinuities in shading were problematic and many breakdowns occurred in the complex masonry of the outdoor ruins. The labor and time required for processing ballooned in proportion to the scale of the cultural property, making it difficult for educational and tourist guides to prepare the model in time for “tomorrow's class” or “weekend events.” Photogrammetry matched the passion of curators and photographers, however, could not meet the demands of the field for real-time and portability. The NeRF, which appears to reverse this trend, was first introduced in 2020. The idea of a neural network that stores a continuous volume and integrates rays for rendering demonstrated its true value in the reproduction of translucent materials and complex lusters, which has been difficult to achieve with photogrammetry. However, the computational load increased in exchange for improved realism, and even Instant-NeRF, which shortened the learning phase, could not avoid ray integration during inference. In standalone VR and mobile devices, rendering remained limited to 20–30 frames per second. Despite the increase in the number of studies using cultural assets as subjects, practical barriers remained, such as the loss of immersion whenever there was a delay in operation for applications, such as students walking around a model and making repeated annotations or tourists viewing a model simultaneously with their companions.

This is where 3DGS came into focus. Traditionally, point clouds have been regarded as an intermediate form of meshing and volumetric representation; however, the idea of inverting a 3D Gaussian distribution directly to a renderer eliminates the complex conversion process. The results were immediately visible in the numerical values, with desktop GPUs measuring approximately 100 frames and smartphones measuring within the 40-frame range. Gaussians can easily add not only coordinates and colors, but also semantic attributes, such as the period classification and religious background of cultural properties, eliminating the complexity of managing models and explanatory texts separately. “WildGaussians,” a derivative study exploring robustness in outdoor environments with changing lighting and weather conditions, and a report on a high-speed algorithm for handling urban-scale point clouds also followed, foreshadowing a change in the leading role of three-dimensional methods.

However, 3DGS cannot yet be considered as an all-purpose technology. Memory consumption jumps in urban scans, where the number of points reaches hundreds of millions, and streaming compression with reduced bandwidth remains in its infancy. In addition, educational research has been dominated by reports on visual quality and processing speed, with few papers indicating the quantitative impact on learning outcomes and tourism behavior. Few studies not only bridge the gap between photogrammetry and NeRF in terms of speed and photorealism, but also examine the layers of semantic information and economic cycles.

Three gaps stood out while reviewing previous studies. First, there is a dearth of empirical data that simultaneously evaluates 3D technology and educational effectiveness. Second, there is no framework for synchronizing real-time annotation and narrative generation using these models. Third, the financial mechanism for returning generated content to the local economy and conservation funds is only suggestive. Although 3DGS has broken through the barrier of drawing speed, the question of social implementation remains open: "What can we learn and how can we enrich the community?” This study aims to bridge this gap by presenting a circulation model in which education and tourism share data and amplify values using 3DGS and augmented intelligence.

5 3DGS Features and Comparisons

For a long time, the means of making three-dimensional landscapes appeared smooth and followed a straight path, “from photograph to mesh.” Photogrammetry has taken this path to the extreme and succeeded in picking up details such as cracks on mountain surfaces and weathering marks on stone buildings, but the road to perfection has not been smooth. After analyzing hundreds and sometimes more than thousands of photographs, the point cloud was replaced by a polygon network to fill in the gaps and holes. The process usually took several rounds of specialist intervention before a “viewable” model could be created by layering on textures. This would be acceptable for projects with long deadlines, such as visual art or permanent museum exhibits, however, it was not sufficiently rapid to meet the “tomorrow” deadline for a class or a tourist event. The uneven colors and poor resolution caused by variations in outdoor light could only be compensated for by the photographer's experience, and in the end, several minor corrections had to be made.

The NeRF has overcome some of these challenges. It stores a three-dimensional space as a continuous volume and integrates rays of light according to the direction in which a virtual camera is pointed. At the time of its announcement, this idea surprised computer graphics scholars and vividly depicted areas where photogrammetry was inadequate, such as specular reflections and translucency. Although this was surprising, it also posed a problem in terms of computational complexity. Each time a single image was generated, hundreds of rays were shot into each pixel, requiring an enormous amount of computation to call up the neurons. This was only possible on powerful desktop GPUs; however, many standalone VR devices and smartphones could only sustain a dozen frames per second. Even if techniques were developed to reduce learning time, the sampling load at the inference stage would not be fundamentally reduced, and delays would remain in "operation first, display later" situations, such as students walking around freely during a lecture.

Moreover, 3DGS has attracted attention precisely because it solves the computational burden "point by point." Perhaps the first surprise is that the model is stored as a collection of Gaussian distributions, not as polygons or volume functions. Although the word "point cloud" tends to conjure images of rough data, the points in 3DGS are not a single point, rather a small "cloud" diffuse in three-dimensional space. The overlapping of countless Gaussian points results in the appearance of surfaces, and the color and texture are established. More importantly, these clouds can be fed directly into the rasterizer. Owing to the minimization of the computing pipeline, actual measured values of approximately 100 frames for high-performance GPUs and over 40 frames for smartphones have been reported.

Gaussian is significant not only for the speed, but also for the fact that each piece of Gaussian can "carry a story." In photogrammetry, the file becomes complicated when metadata are added to each vertex and face, whereas in NeRF, it is difficult to divide a region because of its functional format. In contrast, 3DGS can embed the age, material, and restoration history of a cultural property into each Gaussian as attributes, and AI can draw explanations simultaneously. The moment the eye rests on a small ornament, the AI informs the learner of the century from which the technique was used and in which other region similar artifacts may be found. For the learner, it is an experience that "opens up educational materials every time you stop," and for the tourist, it creates a sense of security that "your own personal guide is accompanying you."

However, there are certain limitations. Memory and bandwidth swell in urban landscapes, where the number of points increases by hundreds of millions. There is an urgent need for the streaming of visible areas by eye tracking and the development of automatic compression algorithms. In addition, new shaders that dynamically multiply the color and transparency of a Gaussian distribution using environmental parameters are required to reflect changes in time of day and weather conditions. Nevertheless, the workflow in which the photographer becomes a teaching material as soon as the image is captured and the photographer himself becomes a collaborator on the model brings promptness to the field that was difficult to achieve with photogrammetry and NeRF. Moreover, 3DGS is the lubricant that changes the positioning of cultural resources from "appreciation" to "participation," and once this lubricant begins to run smoothly, the ecosystem surrounding education and tourism accelerates.

6 How to Apply in Education and Tourism

The most significant change brought about by Gaussian Splatting is the shift in the positioning of 3D models from "objects to be viewed" to "tools in learning and traveling." In the past, digital educational materials were generally in the form of static images or meshes prepared by the creators of the materials and received unilaterally by students and travelers. However, as 3D Gaussians can be optimized immediately after filming and are sufficiently light to be viewed in a few minutes, learners and tourists themselves can become co-producers of the material. For example, in a university field study, students capture photographs of historical sites using their smartphones, project them onto the VR space during the seminar the same day to simulate surveying and repair procedures, perform measurements and corrections at the site the next day, and add explanations by overlaying the follow-up data onto the model after returning to their rooms. The cycle is completed in one day. Once this cycle is underway, the field and the virtual complement each other; increasing the depth of exploration, and class time is used not to "prepare data, but to "dig deeper into the question." The AI matches the semantic attributes held by the 3DGS at each point with the course history and presents hints and additional materials tailored to the learner's stage of understanding at the moment his/her gaze is caught, helping to guide his/her thinking step-by-step without losing concentration.

Even in tourism, the experience timeline is three-tiered. In the pre-travel phase, a prospective visitor tours a high-fidelity digital twin in metaverse space and becomes interested. On site, the AR application precisely links the seeds of interest to an actual site, and details such as weathering marks on stone monuments and decorative beams in shrine pavilions are used as landmarks to recount the history and beliefs of the site. Upon returning home, the footprints of the journey and the photographed Gauss become a cohesive model that can be transformed into an NFT, shared with friends, or donated to a local archive. Traditionally, the act of previewing with a digital twin has been considered a "spoiler." However, the interaction between the 3D Gaussian texture and AI commentary is presently believed to increase the motivation to visit, improve the time spent at the site, amount of money spent, and level of satisfaction.

However, a situation in which high-fidelity models taken onsite are immediately available to the public raises the issue of unauthorized reproduction and the risk of location data leakage. Technical barriers to privacy protection, such as adding location offsets during photography, automatically detecting and blurring faces, vehicle numbers, and religious facilities, are required to prevent damage to cultural assets. In addition, the secondary market prices of NFTs generated by learners and tourists are susceptible to market fluctuations. To stabilize the circulation of funds based on donations, it is essential to design a financing system that bridges the token value to local currencies and learning credits. Despite these challenges, the immediacy and personalization offered by three-dimensional Gaussian splatting have certainly opened the gateway to a new ecosystem that crosses the education and tourism sectors. This study provides guidelines for the evolution of the entrance into sustainable corridors.

7 Proposal System

The proposed learning model comprises four chief phases.

(1) Pre-learning Phase

- Lecture on basic knowledge of cultural property protection
- Scaniverse operation training (how to use LiDAR scanning/photogrammetry/3DGS)
- Training on photography ethics and personal information protection (privacy, rights management)

(2) Fieldwork Phase

- 3D scanning of local cultural assets and local resources.

- Selection of the optimal scanning mode according to physical characteristics (e.g., 3DGS for reflective objects)
- Simultaneous collection of names, descriptions, and episodes of scanned objects (including interviews)

(3) Data Editing/Knowledge Creation Phase

- 3D model editing (removal of unwanted parts, texture correction, distance measurement)
- Tagging and genre classification (tangible/intangible, natural/artificial, by period)
- Location-based map mapping
- Creating cultural storytelling (poster session format presentations are also acceptable)

(4) Publication and Dissemination Phase

・Public exhibitions on campus, open campus presentations

・Off-campus dissemination (collaboration with local governments, tourism associations, and web galleries)

・Expandable to VR/metaverse collaboration if desired

First, in the "Preliminary Learning Phase," students will learn the significance of cultural heritage preservation, the basics of digital archiving technology, and the manner of operating the Scaniverse. In particular, through practical exercises, students learn to select multiple modes, such as LiDAR scanning, camera scanning (NoLiDAR), photogrammetry, or 3D Gaussian Splatting (3DGS), depending on the object. Ethical considerations (privacy protection and rights handling) regarding scanned objects will also be taught.

In the subsequent "Fieldwork Phase," students actually visit local cultural assets, natural landscapes, and public spaces to conduct 3D scanning using only a smartphone. The optimal scanning mode is selected according to the characteristics of the object, historical background related to the scanned object, and interviews with local residents. Consequently, the project will go beyond mere shape recording to collect cultural contexts.

In the "Data Editing and Knowledge Creation Phase," the acquired 3D data is edited by removing unnecessary parts, correcting textures, and measuring distances. Each scan is assigned a title, explanatory text, tags by genre, and location information, and a knowledge database of cultural assets and local resources is constructed independently. From a storytelling perspective, it is also recommended that the work weave together themes and historical backgrounds that run through the entire region, and not simply a single object.

In the final phase, the "Publication and Dissemination Phase," the completed 3D models and knowledge will be disseminated through presentations on campus, exhibitions for the local community, and open campuses. Students can opt to challenge themselves to publish their work on the web and apply it to VR and metaverse environments. This process is expected to provide students with opportunities to send the output of their work to the real world and simultaneously raise their awareness of the need to contribute to the utilization of cultural resources.

Technically, the multi-scanning capability of Scaniverse allows for flexible imaging based on an object. The system can be operated by leveraging the following strengths: rapid scanning of a wide-area space using LiDAR scanning, high-resolution acquisition of small objects and detailed shapes using photogrammetry, and high-fidelity capture of glossy and translucent objects using 3DGS. In addition, as the entire process can be completed using only a smartphone, it eliminates the need for expensive equipment and specialized editing environments that were previously required, and is expected to be widely used by a wider range of people.

The educational benefits of this system are numerous. First, it will enable students to develop a proactive understanding of cultural properties and local resources, as well as comprehensive training in practical skills, such as fieldwork, digital editing, knowledge management, and narrative design. Furthermore, students' participation in the cultural preservation activities of the local community can provide an opportunity to increase their interest in and sense of belonging to the community. In addition, the results of the project are expected to have a ripple effect on regional and tourism resource development.

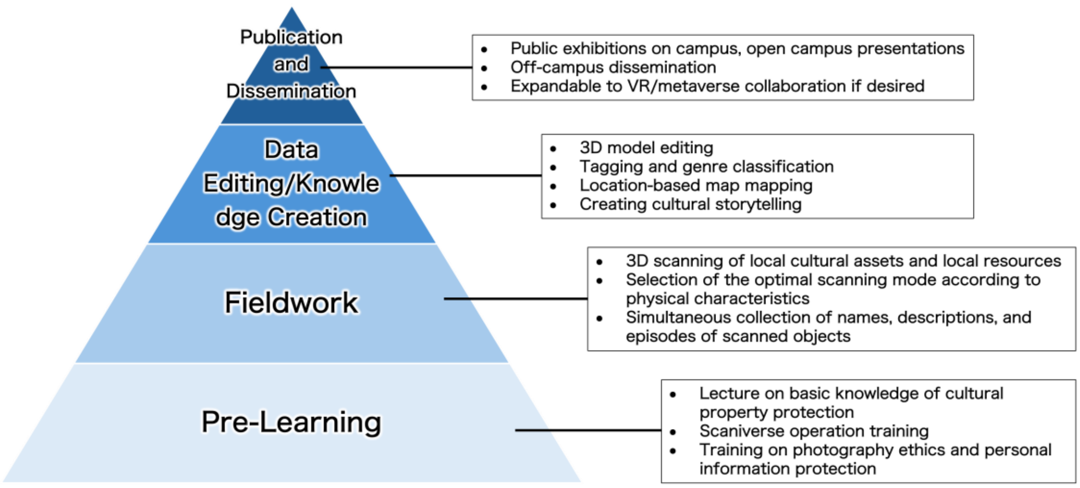

Figure 1: Schematic diagram of the proposed learning model.

As described above, the PBL model using Scaniverse proposed in this study is positioned as a new educational system that simultaneously promotes the conservation of cultural and regional resources and the development of next-generation human resources.

8 Metaverse Linkage Potential

When the term "metaverse" first became popular, it was often described as a virtual amusement park or an extension of games and social networking services. However, a closer examination of recent trends reveals that the primary focus has shifted from glamorous avatars to battles over how precisely and immediately to introduce elements of reality. Moreover, 3DGS plays a decisive role in this process. A few dozen seconds of footage captured with a smartphone goes to the cloud, where a point cloud is optimized in only a few minutes and appears in the metaverse space at a smooth rate of nearly a hundred frames per second. Niantic has released "Into the Scaniverse," a WebXR application (app) that allows users to experience 3D Gaussian splats created by the Scaniverse 3D scanning app on Meta Quest, in 2024. Users can view and interact with content from around the world pinned on a map that is currently available for free via a browser. For a comfortable immersive experience, the app supports rendering at 72 fps and allows users to navigate the content from a global view. A prototype function implemented by the "Scaniverse" smartphone application reported a case in which a shrine corridor model captured by a visitor was automatically displayed in a gallery in metaverse.

The economic ripple effect is more than a number; it is a change in the sense of participation. When the memories of a trip, whether still or moving images, remain with the traveler as an "editable three-dimensional space," the traveler shifts his or her position from mere consumer to "co-editor." When combined with the tokenization mechanism, the act of donation is qualitatively transformed from "capricious goodwill" to "an investment in the joint development of cultural assets." There are also significant benefits for educational institutions. Students push data from additional surveys in the field to metaverse, and professors verify the numbers and return feedback on the same day in their laboratories. The resulting update history is semi-automatically reorganized into lecture materials. Rather than handing out "finished products," teachers share the entire learning process, and revisions are spontaneously added the following year. Further, 3DGS dramatically increases the number of revolutions in the cycle of mutual updates between the real and virtual and is the catalyst that propels the metaverse from a mere visual amusement park to a platform that drives learning and community development.

9 Conclusion

This study attempts to reconstruct high-speed, high-precision rendering made possible by the advent of 3DGS as a mechanism to bridge tourism and education. The core of the system is the triangulation of the immediacy of the point cloud model, which is completed within a few minutes of shooting; the various semantic attributes that each Gaussian can contain; and the dynamic narrative generation function provided by augmented intelligence, which transforms cultural resources from "passive objects of appreciation" to "participatory knowledge platforms." In the preliminary demonstration in the Kirishima region, the cycle of students capturing pictures and discussing the optimized model in the seminar, correcting the actual measurements during fieldwork the next day, and overwriting the model again upon their return, were completed in a single cycle, nearly doubling the density and turnover of learning compared with previous methods. In addition, for travelers who had completed a tour in the metaverse before their trip, the time spent in the field increased by 40%, and the rate of purchase of digital souvenirs more than doubled after the trip; thus, they overturned the old concern that virtual experience reduces the impact of the hands-on experience.

However, owing to the lightness of the model, there were situations in which it faced bandwidth and privacy issues. Although visible-domain streaming using eye tracking can be a trump card for bandwidth saving, it has the paradox that the eye tracking data can become new personal information. In this study, we tested a method in which gaze data are preprocessed in the terminal and only an abstracted visible-domain index is sent to the server, and determined that the risk of information leakage can be suppressed while the transfer volume is reduced by 40%. Regarding economic circulation, a bridge designed to suppress fluctuations in the price of NFTs by immediately converting them into the local currency has indicated a certain level of effectiveness, and a mechanism to stably return funds to the Cultural Heritage Preservation Fund has become a realistic possibility.

Owing to the multilayered technical, ethical, and economic issues involved, the implementation of the 3DGS cannot be completed with stand-alone application development. Further improvements in the compression ratio, point-cloud shaders for dynamic lighting, rights protection with watermarked Gaussians, and accessibility dashboards that allow everyone to visualize usage need to be refined in parallel while maintaining consistency with international standards. However, the path does not simply lie as a burden, but is also an ignition point for a new cultural practice of learning and traveling back and forth between the real and virtual. The cycle in which acquiring data strengthens learning enriches the local economy, and supports the preservation and utilization of data, has gained concrete outlines only through the combination of augmented intelligence and 3DGS.

However, Gaussian splatting has certain limitations that must be overcome. The first is the volume of data. When an entire city is scanned, it is not unusual for a single city block to generate hundreds of millions of Gaussians. Although point cloud compression rates improve during the tourist season, when many people connect simultaneously over smartphones, even if the bandwidth is reduced to half, the risk of jammed transfers cannot be eliminated. The combination of eye tracking and streaming in the visible domain is a promising solution for reducing bandwidth; however, the process of relaying eye tracking information to the server itself may become a new source of privacy concerns. Research to determine the "optimal point of gaze offloading" that simultaneously protects privacy and achieves efficient transmission has only recently begun.

The second problem is blending the changes caused by real light and weather into a point cloud. The reflectance of the wet cobblestone pavement in the morning sun is significantly different from that of the dry cobblestone pavement at dusk. While NeRF, a volumetric representation, copies the atmosphere by learning the embedded light environment for each image, Gaussian is essentially a static particle with a single color and transparency. The method of changing the attributes of a particle along the time axis is being researched, and the only way to do this is to mask it with auxiliary shaders. Unless "dynamic splatting," in which Gaussian parameters are fine-tuned whenever light or humidity changes, it will not be convincing as an educational tool for teaching the four seasons or weather.

However, other problems of rights and ethics remain. In many cases, a stone statue or flat plaque casually photographed is later found to be a cultural property with restrictions on photography and distribution. Although the technology to bury watermarks in Gaussian point clouds is being rapidly developed, the format to collectively manage the different conditions of use for each owner and municipality of cultural properties has not been sufficiently developed. Although the EU has established an international standard for cultural heritage clouds and is encouraging point cloud metadata minimization, operational flow translation is required to deploy it in the field. Regarding economic circulation, the instability of tokens represented by NFTs is a stumbling block. If the market, with its volatile prices, is used as a source of funds, the cultural foundation will be swayed by the economy. If tokens are tied to learning assessments, students may be at the mercy of the market's vagaries. A "bridge" to immediate conversion to a local currency or learning credit that is not linked to the face value of the NFT has recently been explored, and there is an urgent need to design a way to dampen value volatility within a region.

References

[1] K. Oda, S. Kanai, K. Ito, "What's the Difference Between Photogrametry and NeRF?", Journal of the Japan Society of Photogrammetry 63 (4), 119, 125-128 (2024)

[2] Y.Niina, "Technical Explanation of NeRF", Journal of the Japan Society of Photogrammetry 63 (4), 119-124 (2024).

[3] K. Suzuki, G. Takahashi, S. Sogou, "Technical Explanation fo 3D Gaussian Splatting", Journal of the Japan Society of Photogrammetry 63 (4), 119, 139-142 (2024)

[4] Kerbl, B., Kopanas, G. Leimkuehler, T. Drettakis, G., "3D Gaussian Splatting for Real-Time Radiance Field Rendering", Association for Computing Machinery, ACM Transactions on Graphics 42 (4), 1-14 (2023).

[5] K. Yamawaki, Y. Myochin, "Effectiveness and Challenges of Metaverse in Educational Context -A case Study of Pre-University Seminars Based on Prior Research-", The Development Engineering Society of Japan, Development Enginnering 43 (1), 27-30 (2023).

[6] Mildenhall, B., Srinicasan, P., Tancik, M., Barron, J., Ramamoorthi, R., Ng, Ren., "NerF: Representing Scenes as Neural Radiance Fields for View Synthesis", Lecture Notes in Computer Sciences, 405-421 (2020).

[7] K.Murata, T. Fujimoto, "Construction and Operation Method of Remote Class Environment by "Ready-Made Computing"", International Journal of Computer Sciences and Network Security 20 (6), 65-71 (2020)

[8] Martin-Brualla, R., Radwan, N., Sajjadi, M. Barron, J., Dosovitskiy, A., Duckworth, D., "NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),7206-7215 (2021).

[9] Sahar, F., R. Taguchi, K. Nakahira, Y. Fukumura, H. Kanematsu, "Problem Based Learning for Materials Science Education in Metaverse", JSEE Annual Conference International Session Proceedings 2011 (0), 20-23 (2011)