A Method of Crowdsourced Task Request Optimization for 3D Reconstruction in Urban Space

Shinya Akiyama†, Yuki Matsuda†, Hirohiko Suwa†, and Keiichi Yasumoto†
†*Nara Institute of Science and Technology*, Nara, Japan,
{akiyama.shinya.aq9, yukimat, h-suwa, yasumoto}@is.naist.jp

***Abstract*—This paper proposes an optimization method for scanning task requests to reduce various costs involved in creating and updating 3D models. In recent years, there has been a growing interest in research related to metaverses. Among them, the utilization of 3D models created by scanning real-world buildings has various applications. Examples include virtual travel experiences and simulations for autonomous driving. However, the creation of 3D models is time-consuming and costly, As a result, the updating of 3D models occurs on a time scale of several months to several years. However, due to the low frequency of updating 3D models, virtual travel experiences can be unsatisfying compared to real-world travel. For example, objects related to seasonal events or foliage changes in trees may not be updated, resulting in a lack of immersion in the virtual travel experience. In this study, we propose a method to induce (request) general participants interested in the target object to take photos that are useful for 3D model creation and updating. The goal of this research is to increase the frequency of creating and updating 3D models, thereby achieving a more realistic metaverse experience.**

***Index Terms*—modeling, photogrammetry, scanning task**

I. Introduction

In recent years, there has been a surge of research in the field of metaverse. Among them, a particularly noteworthy topic is the endeavor to recreate real-world cities within virtual spaces by placing 3D models of buildings and objects that are scanned from the physical environment. This holds great potential in various areas, such as virtual tourism, where users can explore virtual cities as if they were in the real world, urban planning, simulations like autonomous driving, and more. However, it is important to note that the objects required for these endeavors are not limited to static structures like buildings. In real-world travel, one can experience the differences brought by seasonal changes by visiting the same destination repeatedly. This is a common way of enjoying travel. However, due to the low frequency of updating 3D models, virtual travel experiences can be unsatisfying compared to real-world travel. For example, objects related to seasonal events or foliage changes in trees may not be updated, resulting in a lack of immersion in the virtual travel experience.

In this paper, we propose a participatory sensing system for timely 3D modeling of objects referred to as "Urban Objects" that are either installed for a limited period or require relatively frequent updates compared to annual updates. The system is designed to enable participants to complete sensing tasks using widely available devices without requiring specialized knowledge. The goal is to reduce scanning costs, improve the accuracy of 3D model creation, and increase the frequency of updates.

II. Photogrammetry: Requirements and Related Work

A. Requirements

To bring "Urban Objects" into the virtual space, it is necessary to complete the scanning of the target objects within a relatively short period while they exist in the real world. There are various methods for 3D modeling, but photogrammetry is a technique that can be used to quickly generate 3D models, including surface textures, without specific focus on the texture itself. Photogrammetry involves capturing photographs of the target object from various angles and combining them to generate a 3D object. However, one challenge of photogrammetry is that generating a seamless 3D model requires a large number of photographs taken from different angles to avoid any inconsistencies or artifacts.

To assess the required angular resolution for the target object, a dataset of 36 photographs was prepared, capturing the object from a fixed distance at 10-degree intervals, covering a complete 360-degree rotation. The ObjectCapture software, provided by Apple, was used to attempt 3D object reconstruction. The results provided insights into the photos that are useful for photogrammetry.

- Each photograph should capture both the same and different parts of the target object.
- If the above conditions are met, it was possible to generate a 3D object even from just three photographs.
- With a complete dataset covering a full 360-degree rotation of the target object, it is possible to generate a 3D object with sufficient realism.
- The parts not captured in the photograph are represented as missing in the 3D object.
- The image quality of the photographs directly affects the realism of the generated object.

B. Related Work

Structure from Motion (SfM) [1] is one of the representative methods in photogrammetry. SfM extracts feature points from numerous photographs of a target object and connects them to generate a 3D model. SfM also allows for the estimation of relative positions at which the photographs were taken with respect to the object. This has been applied in research efforts

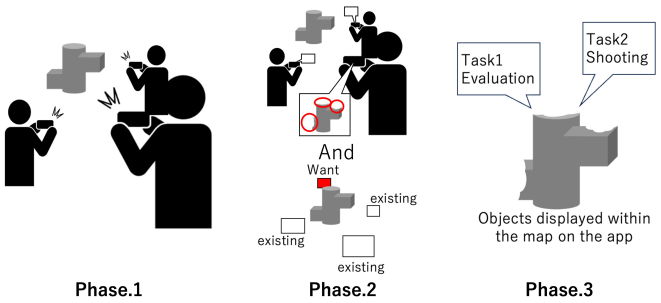

Fig. 1: Proposed Method

such as integrating sets of photographs of buildings from the internet into 3D maps to facilitate browsing and searching [2]. There are also previous examples of utilizing SfM to suggest unphotographed locations on maps and encourage players to capture photos of those places in gaming scenarios [3].

III. Proposed Method

In this paper, we propose a task assignment approach that encourages event participants and tourists to naturally capture the photos required for photogrammetry. While this approach shares similarities with the study conducted by Tuite et al. [3], the key difference lies in the photographers themselves. Instead of targeting specific players in a game, our approach aims to involve general participants such as event attendees and tourists who encounter "Urban Objects" in the city. The objective is to encourage a wide range of participants to capture photos.

This method of requesting photography tasks is intended to be integrated into a 3D object generation service. This 3D object generation service possesses the characteristics of low-cost 3D generation and frequent object updates, making it widely applicable to the conventional needs of 3D object generation. For instance, it can be utilized in contexts such as content creation for movies, games, and virtual tours, simulations for urban planning, or the creation of miniatures using 3D printers. Our proposed method consists of three phases, but it is premised on the initiation of requests for 3D object generation from individuals who require 3D objects with the aforementioned needs. Individuals who wish to create 3D objects are free to determine their budget for object creation, and our envisioned service determines the amount of compensation based on their budget and the contributions of participants. This compensation can be in the form of cash, souvenir coupons, in-game currency, or items that photographers find valuable—anything that holds value for photographers. This mechanism itself is akin to crowdfunding.

The overview of the proposed method is illustrated in Fig. 1.

In the first phase, event participants or tourists capture photos of the target objects in a spontaneous manner and upload them to the system. These photos are taken with a sense of taking commemorative photos and are considered sufficient. Regarding the uploaded photos, a uniform reward will be given to the uploader for each uploaded photo.

Next, using Structure-from-Motion (SfM), the collected set of photos is used to generate 3D models of the captured objects, and the locations where the photos were taken are determined.

The second phase is the evaluation phase. Participants who have uploaded photos are sent links to the 3D models and are asked to identify any missing or unnatural parts of the models. Participants who provide such feedback are rewarded with a modest compensation. Based on the feedback, photos that are needed for capturing the missing parts (e.g., higher-resolution photos or photos from specific angles) are determined. Additionally, using the calculated shooting locations from the set of photos obtained in the first phase, angles or points where the objects have not been captured are identified and matched with the photos determined based on the feedback.

The third phase is the task assignment phase. In this phase, individuals present in the vicinity of the object are assigned the task of capturing the required photos identified in the second phase. This is done by placing the 3D model created in the first phase on a map that can be viewed on a smartphone. Participants can select the model as an icon and view the "capture task" associated with it. The capture task includes information such as the duration of the task, desired weather conditions, image quality, and specific shooting angles. Participants capture photos according to the task requirements and upload them. The system evaluates the task completion rate of the uploaded photos using mathematical criteria and rewards the photographers based on their achievements. The rewards offered for task completion are set relatively higher than those obtained in the first and second phases.

Afterwards, the second phase and the third phase are essentially repeated until one of the following conditions is met: the number of identified issues in the second phase falls below a certain threshold, the client's budget is depleted, the client is satisfied with the quality of the 3D object and chooses to conclude the request, or the removal of the object is reported within the second phase.

IV. Conclusion

In this paper, we propose a method for increasing the update frequency of 3D models of objects at a low cost by widely outsourcing scanning tasks to general users. In future research, we plan to validate the effectiveness of this proposed method within a 3D object generation service, focusing on streamlining and cost reduction in 3D model updates.

References

[1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2003.

[2] N. Snavely, S. M. Seitz, and R. Szeliski, "Photo Tourism: Exploring Photo Collections in 3D," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 835–846, 2006. [Online]. Available: https://doi.org/10.1145/1141911.1141964

[3] K. Tuite, N. Snavely, D.-y. Hsiao, N. Tabing, and Z. Popovic, "PhotoCity: Training Experts at Large-Scale Image Acquisition through a Competitive Game," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11, 2011, pp. 1383–1392. [Online]. Available: https://doi.org/10.1145/1978942.1979146