Landscape Visualization by Integrating Augmented Reality and Drones with Occlusion Handling to Link Real and Virtual Worlds

Towards city digital twin realization

Naoki Kikuchi[1], Tomohiro Fukuda[2], Nobuyoshi Yabuki[3]
[1,2,3]Division of Sustainable Energy and Environmental Engineering, Graduate School of Engineering
[1]kikuchi@it.see.eng.osaka-u.ac.jp [2,3]{fukuda|yabuki}@see.eng.osaka-u.ac.jp

In the field of urban architecture and design, augmented reality (AR)-based landscape visualization is useful for building consensus among stakeholders at the design stage. An integrated AR and drone method can visualize future and past landscapes from an aerial perspective but has to address the problem of occlusion, where a 3D virtual model is displayed in front of the real-world objects. In this study, we propose an AR and drone integrated landscape visualization method to handle occlusion by linking the drone's location information in the real world and the camera in the virtual world. The method uses a 3D model of an existing building, which is part of the city model, to represent the 3D model of the design target as if the target were behind the existing building in the real world. Users can use the perspective of the drone, which flies along a set route, to examine the future landscape with high accuracy, as visualized using AR with occlusion handling.

Keywords: *Digital twin, Occlusion handling, Landscape visualization, Web-based augmented reality (web AR), Drone, Urban design*

INTRODUCTION

In the field of architectural and urban design, public involvement is expected during the planning and design stages to examine landscape changes. The design information provided to participants needs to be easy to understand. Augmented reality (AR)-based landscape visualization at the design stage effectively provides users with visual information about the before and after appearance of the designed landscape and can help to build consensus among stakeholders (Goudarznia et al. 2017). When a user wears a head-mounted display or holds a mobile device such as a smartphone, the AR perspective is limited to the user's range of action.

Recently, drone research has advanced, and drones are used for various applications. In urban planning, drones are used for environmental monitoring and urban space management, providing a perspective outside the user's range of action (Gallacher 2016). To visualize past and future landscapes from the air, an integrated AR and drone method was proposed (Wen and Kang 2014, Unal et al. 2018). However, the integrated AR and drone method has an occlusion problem. A virtual object in front of a real object wholly or partially hides the real object. In general, in AR, virtual objects are rendered last in relation to the real world. Without intervention, a virtual object is displayed in front of real objects. This discrepancy between the relative placement of a real object and a virtual object (i.e., determining which one is behind the other) is called the occlusion problem (Figure 1). Occlusion handling is a process to solve this problem.

A city model is made of 3D geospatial data that reproduce a real-world (physical space) city in a virtual world (cyberspace). City models are essential for public participation in reviewing the design of architectural and urban spaces (Ruohomäki et al. 2018). A digital twin is defined as "a digital replica of a physical asset, process, or system," where the digital twin and the physical twin continuously interact (Batty 2018). The ideal urban digital twin is a city model that reflects its combined economic, ecological, and demographic conditions and changes (Dembski et al. 2020). In the field of urban architecture and design, the use of digital twins of cities is at an early stage, and their use remains inconsistent.

This study proposes an AR and drone integrated landscape visualization method that handles occlusion by linking the camera's location information in the real world to the location of a virtual world's camera in cyberspace. First, a design target virtual 3D model and a virtual 3D model of an existing building that is part of a city model are placed in the virtual world. Next, the drone's camera and the virtual world's camera are predetermined to move along the same path simultaneously so that the real world and the virtual world are linked as digital twins. The method uses 3D models of an existing building, which are part of the city model, to represent the appearance of the design target's 3D model as if it were behind the existing building in the real world. Also, to integrate AR and a drone with high versatility, mirroring and virtual camera technologies are employed. Mirroring allows the same screen to be displayed on multiple devices. At the same time, a virtual camera treats the PC screen as a webcam video. By employing mirroring and a virtual camera, we can easily use any available drone without regard to a drone-specific software development kit (SDK). By superimposing the real scene and the occlusion-handled virtual scene, it is possible to study the future landscape with high accuracy as visualized with AR from the viewpoint of a drone flying along a set route. The drone's camera and the virtual world's camera move along the set drone route with synchronized positions.


(a) (b) (c)

Figure 1
Occlusion problem in AR (a) Real world (b) AR with incorrect occlusion (mock-up) (c) AR with correct occlusion (mock-up)

RELATED WORK

Augmented Reality in Public Participation

In Arnstein's (1969) "Ladder of Participation," in the process of democratic empowerment, one step is providing information to the public. Specifically, AR is used as a tool to provide geospatial information to the public. In urban space, AR visualization effectively provides visual information of landscape changes before and after the design stage and can help form a consensus among stakeholders, including non-specialists such as the public. AR in urban space research (Goudarznia et al. 2017, Haynes et al. 2017) is limited to perspectives from within the AR user's area of action and does not allow users to access the aerial AR perspective.

Integration of AR and Drones

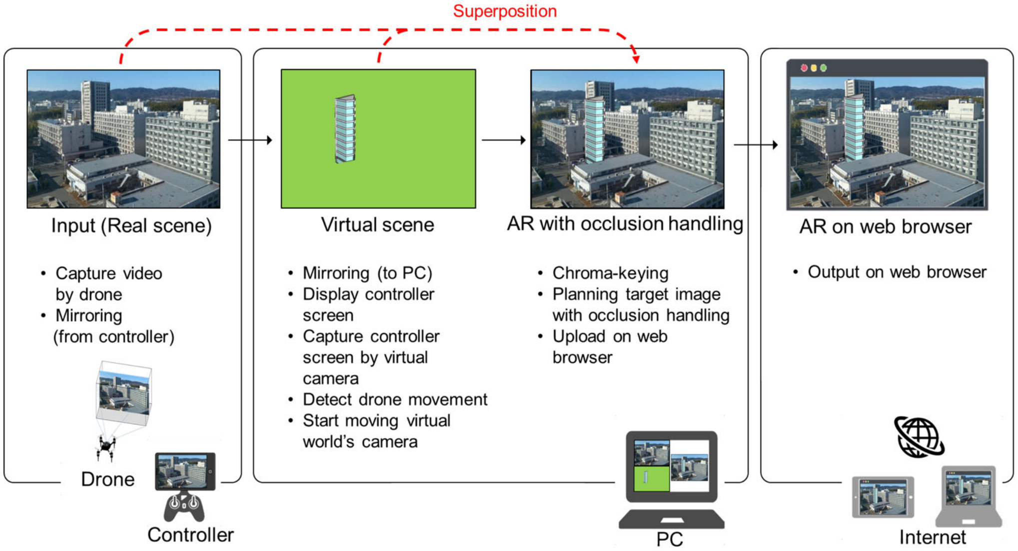
Wen and Kang (2014) integrated AR and a drone to display 3D virtual models in the real world from an aerial perspective. However, their drone was designed and built from scratch, and their method was less versatile because only specific drones could be used. Unal et al. (2018) proposed a method that integrates location-based AR and a drone to improve the accuracy of displaying 3D virtual models. Previous integration AR and drone methods had the problem occlusion: the 3D virtual model always appears in front of the real-world object. The problem was that the system configuration depended on the drone model because the SDK provided for drones has limited functionality.

Occlusion Handling Method

In the occlusion problem, AR users may misunderstand the position of the 3D virtual model when it always appears in front of the real-world object. To provide accurate information, occlusion handling is essential. To solve the occlusion problem, depth-based, foreground object-based, and model-based methods were proposed.

The depth-based method deals with occlusion by comparing the depth information of an occlusion target obtained from an RGB-D camera with the depth of a virtual object, as proposed by Du et al. (2020). The depth-based method can handle the occlusion in real time. However, the depth informa-

Figure 2
Conceptual diagram of our proposed method



tion that the 3D sensing camera can obtain is limited, making it unsuitable for use in wide outdoor spaces.

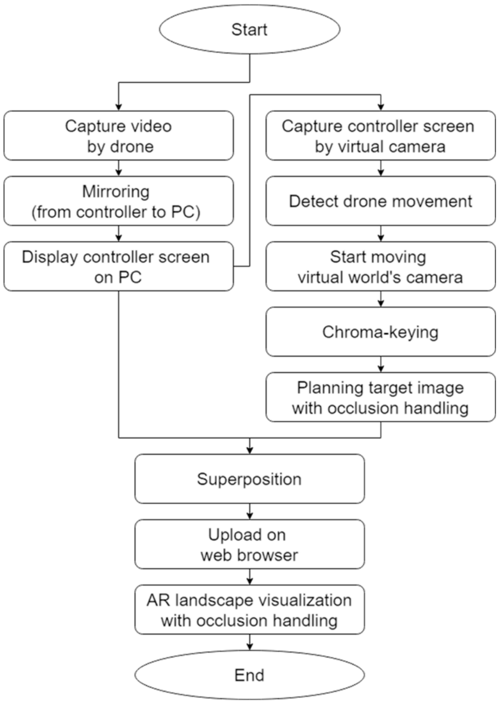


Figure 3
Flowchart of our proposed method

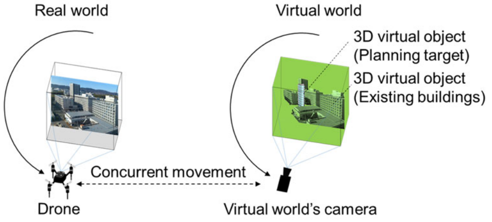


Figure 4
The concurrent movement of drone's camera and virtual world's camera

Foreground object-based methods use image processing to extract foreground objects from still images and videos and use the extracted object outline to handle the occlusions. Roxas et al. (2018) used semantic segmentation to extract the foreground objects and the occlusion handling.

In the model-based method, a 3D virtual model of the occlusion target (occlusion model) is predefined in the virtual scene. Occlusion is handled by comparing the depth of the design target 3D virtual model with the depth of the occlusion model (Kasperi et al. 2017). Model-based methods need to create an occlusion model as a pre-processing step.

Research Contribution

In our feasibility study, we handled occlusion using instance segmentation with the foreground object-based method. However, the AR frames occluded by the instance segmentation had a processing speed of only a few frames per second (fps). This paper proposes a model-based method using a 3D virtual model of an existing building that is part of a city model. The 3D virtual model of the existing building is used as an occlusion model. The occlusion problem is solved by comparing the depth from the camera in the virtual world of the 3D virtual model of the design target with the depth from the camera in the virtual world of the occlusion model. The proposed method achieved occlusion-handled AR at frame rates of tens of frames per second. We also improved the model dependency of the system configuration with the drone SDK in the integrated AR and drone method (Wen and Kang 2014, Unal et al. 2018).

METHODOLOGY

The concept and flow diagram of our method are shown in Figures 2 and 3, respectively. First, 3D virtual models of the design target and the existing building are placed in the virtual world in an appropriate relative position. The background and 3D model of the existing building are changed from RGB values to unrealistic color values and set in the model as emissive material. In order to superimpose the 3D virtual model of the design target on the video from the drone's perspective, the unrealistic colored areas on the video from the drone's perspective are masked. Mask processing is an image processing technique that displays certain areas of an image or video and hides other areas. By setting the drone's camera and the virtual world's camera to move along the same route simultaneously, the drone's location information in the real world and the camera in the virtual world are linked (Figure 4). By superimposing the video from the drone's perspective as it flies along the predetermined route with the video from the camera's perspective in the virtual world, AR with occlusion from an aerial perspective is realized. The generated AR content is uploaded to a video distribution platform using streaming software and output on the web browser of the AR user's device.

Versatile System Configuration

In traditional integration of AR and a drone (Wen and Kang 2014, Unal et al. 2018), the system configuration is limited by the drone SDK. To solve model dependency, we integrated AR and a drone without using the drone SDK but instead used mirroring and a virtual camera. Mirroring is a technology to duplicate the screen to multiple devices. A virtual camera is a technology to treat a PC screen as a webcam video. The video on the controller is displayed on the PC by mirroring. The video captured from the drone on the controller using the virtual camera can be processed. The image processing detects the change in the controller's video when the drone starts to move and synchronizes the start of movement of the drone's camera in the real world and the virtual camera in the virtual world.

AR system needs to be user-friendly to encourage public participation. The video delivery platform enables the use of AR content on web pages. Thus, users can easily experience AR content by accessing a web page using an URL or QR code without installing any specific applications.

EXPERIMENTS AND RESULTS

To verify the applicability of the proposed method, we constructed a prototype system.

The hardware of the prototype system was as follows. We used a Galleria GCL2060RGF-T PC (Intel Core i7-10875H CPU, 32 GB RAM, NVidia GeForce RTX 2060 GPU, and Windows 10 home v20H2 OS) to generate an AR video with occlusion handling and upload it to the web. A DJI Mavic mini drone was used to realize the real-world aerial view. An iPhone 11 (iOS v14.4) with an installed drone control application was used as a drone controller. An iPad (iOS v14.4) was used as a user device to view the AR video uploaded to the web.

The software of the prototype system was as follows. DJI Fly (iOS v1.3.0) drone control software was used for the drone's automatic navigation. Reflector3 (v3.2.0) mirroring software was used to duplicate the smartphone screen onto a PC, including audio, and to transfer the video from the controller to the PC. OBS Studio (v26.1.1) streaming software enabled video streaming, and its virtual camera function was used to treat the PC screen as a webcam video. The Unity (2020.1.7f1 64-bit) game engine was used to build the virtual world. Autodesk InfraWorks 2019 infrastructure design software was used to create 3D models of existing buildings as part of the city model. The YouTube video distribution platform was used to output the created occlusion-enabled AR onto the web. The Google Chrome (v89.0.4389.82) web browser was used to access YouTube.

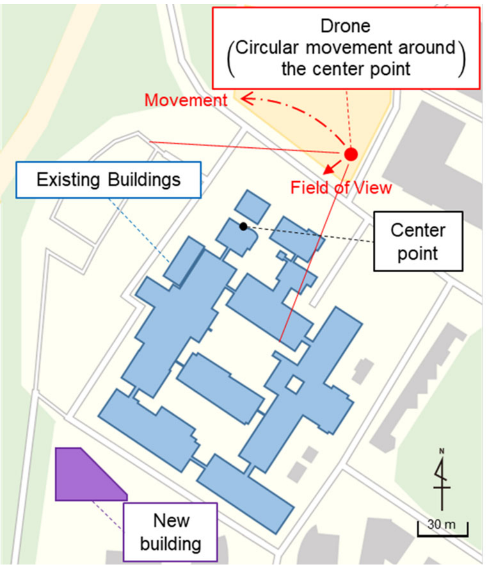


Figure 5
Arrangement of existing and planned (new) buildings

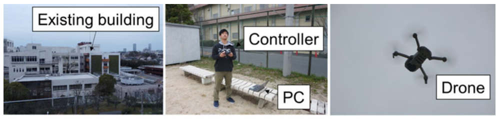


Figure 6
Photographs of experiment setting

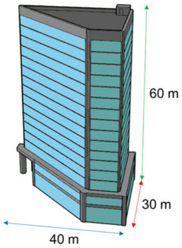


Figure 7
Dimensions of the planned building model

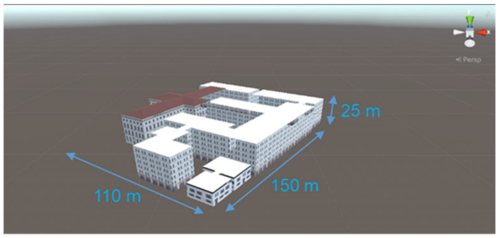


Figure 8
Occlusion model of the existing buildings

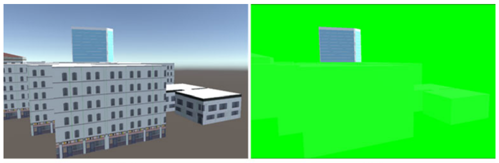


Figure 9
Changing settings of background and occlusion model:
(a) before changing RGB values and (b) after changing RGB values

Occlusion Handling Outdoors

To verify whether the proposed system can handle occlusion from the viewpoint of a drone camera flying through an expansive outdoor space, we conducted an outdoor AR experiment. Figure 5 shows the layout of the planned and existing buildings, and the camera route and direction. The drone was automatically navigated in an arc around the center point in Figure 5, facing the center at 20 ± 1 m and 40 ± 1 m altitude, and stopped after 16 s. Figure 6 shows the scene of the outdoor AR experiment. Figure 7 shows a 3D virtual model of the planned building for the landscape study. Figure 8 shows the occlusion model of the existing building, created using InfraWorks. The occlusion model created using InfraWorks was a simple box model with no detailed roof shape. The occlusion model and the background were changed to bright green ((R, G, B) = (0, 255, 0)) (Figure 9).

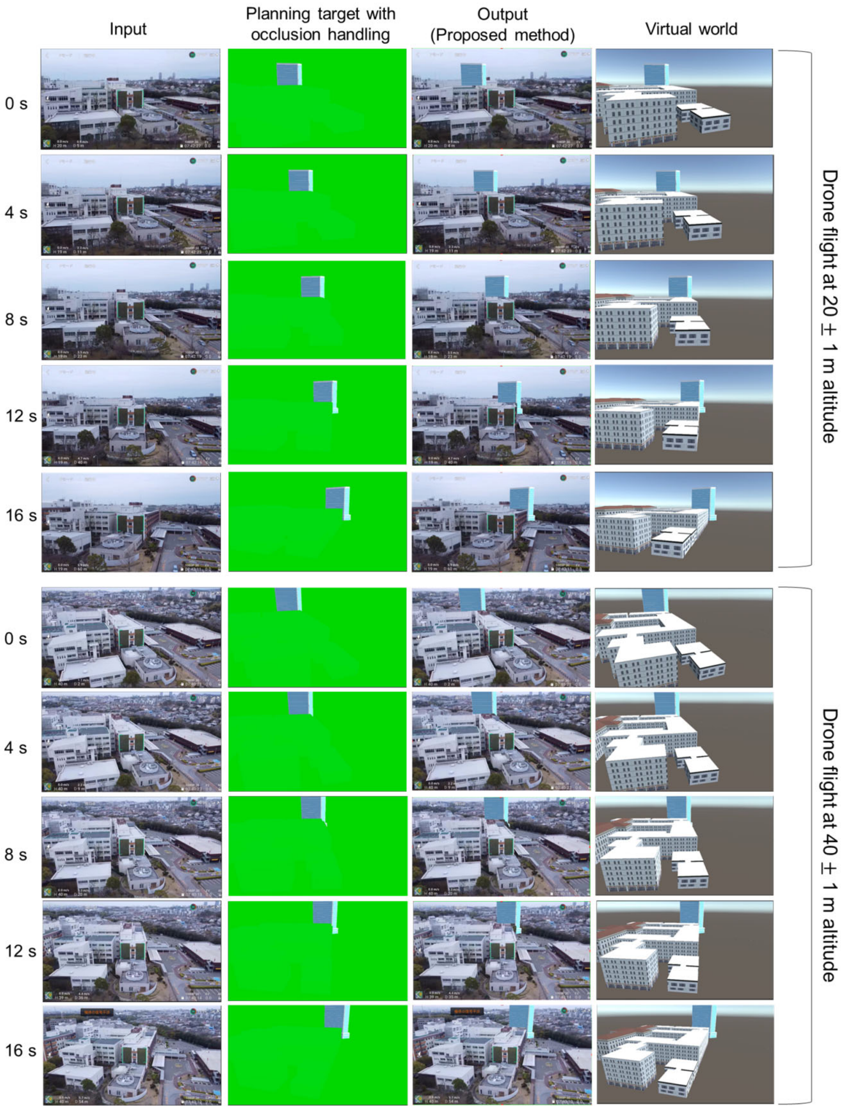
System Performance Evaluation

To evaluate the performance of the prototype system, Figure 10 shows the results every 4 s when the drone was flown at altitudes of 20 ± 1 m and 40 ± 1 m for 16 s. In this validation, the occlusion handling was possible from the drone camera's point of view, as it flew along a predetermined route in an outdoor space. The overall system processing speed and delay of the proposed system were measured. The frame rate of the video captured from the drone was about 30 fps, the frame rate of the video captured from the camera in the virtual world was average about 700 fps on Unity, and the frame rate of the video output on the web was about 30 fps. The delay between displaying the video on the controller and on the web was about 4 s.

DISCUSSION

In the prototype system, the occlusion-handled AR imagery could be visualized in a web browser from the viewpoint of a drone that automatically navigated along a predefined route. The overall prototype system frame rate was about 30 fps, and the delay was about 4 s. User devices are commonly used for transmitting public information. The frame rate was 30 fps, which is the frame rate of widely used devices, and the processing speed was sufficient. The latency was about 4 s, a difference that is noticeable and needs to be improved. The results in Figure 10 show that the occlusion handling is not accurate around the boundary between the existing building and the 3D model of the planned building. The accuracy of the occlusion handling needs to be improved. The occlusion handling's low accuracy is thought to be caused by two main factors: the occlusion model's accuracy and the accuracy of the AR alignment.

Figure 10
Results of occlusion handling



CONCLUSION

The conclusions of this study are as follows:

- To realize an AR landscape visualization method with occlusion handling from the air, we integrated AR and drone technologies and handled the occlusion based on a model-based method using 3D virtual models of existing buildings that are part of a city model.
- To improve the system configuration limitation, we integrated the AR and drone techniques without using the drone SDK by using mirroring and virtual camera techniques.
- The prototype system frame rate was confirmed at about 30 fps, and the delay was about 4 s.

In future work, to improve the accuracy of occlusion handling, it is necessary to develop a method to correct the position information of the drone camera and the camera in the virtual world. We will develop an AR method that includes the digital twin concept, which continuously links real-world location information and the virtual world via a drone, and the location information of the drone camera in the real world and the camera in the virtual world.

ACKNOWLEDGMENT

This research was partly supported by the Japan Society for the Promotion of Science, KAKENHI Grant Number JP19K12681. All images were created by the authors from 2020 to 2021.

REFERENCES

Arnstein, S.R. 1969, 'A Ladder of Citizen Participation', *Journal of American Institute of Planners*, 35, pp. 216-224

Batty, M. 2018, 'Digital twins', *Environment and Planning B: Urban Analytics and City Science*, 45(5), pp. 817-820

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M. and Yamu, C. 2020, 'Urban Digital Twins for Smart Cities and Citizens: The Case Study of Herrenberg, Germany', *Sustainability*, 12(6), pp. 1-17

Du, R., Turner, E., Dzitsiuk, M., Prasso, L., Duarte, I., Dourgarian, J., Afonso, J., Pascoal, J., Gladstone, J., Cruces, N., Izadi, S., Kowdle, A., Tsotsos, K. and Kim, D. 2020 'DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality', *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, Virtual Event, USA, p. 829–843

Gallacher, D. 2017, 'Drone Applications for Environmental Management in Urban Spaces: A Review', *International Journal of Sustainable Land Use and Urban Planning*, 3(4), pp. 1-14

Goudarznia, T., Pietsch, M. and Krug, R. 2017, 'Testing the Effectiveness of Augmented Reality in the Public Participation Process: A Case Study in the City of Bernburg', *Journal of Digital Landscape Architecture*, 2, pp. 244-251

Haynes, P., Hehl-Lange, S. and Lange, E. 2018, 'Mobile Augmented Reality for Flood Visualisation', *Environmental Modelling & Software*, 109, pp. 380-389

Kasperi, J., Edwardsson, M.P. and Romero, M. 2017 'Occlusion in outdoor augmented reality using geospatial building data', *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, NY, USA, p. Article 30, 1–10

Roxas, M., Hori, T., Fukiage, T., Okamoto, Y. and Oishi, T. 2018 'Occlusion handling using semantic segmentation and visibility-based rendering for mixed reality', *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, 20, pp. 1-8

Ruohomäki, T., Airaksinen, E., Huuska, P., Kesäniemi, O., Martikka, M. and Suomisto, J. 2018 'Smart City Platform Enabling Digital Twin', *2018 International Conference on Intelligent Systems*, Funchal, Portugal, pp. 155-161

Unal, M., Bostanci, E., Sertalp, E., Guzel, M.S. and Kanwal, N. 2018 'Geo-location Based Augmented Reality Application for Cultural Heritage Using Drones', *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies*, Ankara, Turkey, pp. 1-4

Wen, M.C. and Kang, S.C. 2014, *Augmented Reality and Unmanned Aerial Vehicle Assist in Construction Management*, Computing in Civil and Building Engineering (2014) (pp. 1570–1577)