RESEARCH ARTICLE

Future landscape visualization using a city digital twin: integration of augmented reality and drones with implementation of 3D model-based occlusion handling

Naoki Kikuchi ⓘD, Tomohiro Fukuda ⓘD* and Nobuyoshi Yabuki ⓘD

Division of Sustainable Energy and Environmental Engineering, Graduate School of Engineering, Osaka University, Suita, Osaka 565-0824, Japan

*Corresponding author. E-mail: fukuda.tomohiro.see.eng@osaka-u.ac.jp ⓘD http://orcid.org/0000-0002-4271-4445

Abstract

The integration of augmented reality and drones allows past and future landscapes to be visualized from an aerial perspective. However, these visualizations still suffer from the occlusion problem, where the three-dimensional (3D) virtual model displayed in the real world is in front of a real-world object. Currently, city digital twins are essential for the sustainable development of cities and the development of detailed 3D models of cities. By visualizing the city digital twin, augmented reality can facilitate the participation of nonexpert citizens in the decision-making process of urban design, but research examples are limited. Here, using detailed city 3D models, we develop a digital-twin approach to outdoor augmented reality with occlusion handling for both first-person and bird's-eye views. In a verification experiment, the occlusion handling accuracy of the prototype system was measured to be about 0.8 using intersection over union. The frame rate of the entire prototype system was about 30 fps, and the delay between the controller and the augmented reality device was about 3 s. The internet-based system architecture was developed to integrate augmented reality and drone systems. Our system allows multiple stakeholders involved in building construction projects to observe aerial perspectives of those projects, both on-site and off-site via an internet browser, using augmented reality with occlusion handling.

Keywords: city digital twin; occlusion handling; landscape visualization; web-based augmented reality; drone

1. Introduction

In the field of architectural and urban design, public involvement is expected during the planning and design stages in the examining of landscape changes (Wiberg *et al*., 2019). The design information provided to participants needs to be easy to understand. Augmented reality (AR) superimposes virtual data on the real world (Milgram & Kishino, 1994; Azuma, 1997). The visualization of the landscape at the design stage using AR provides users with visual information about the result of a proposed de-

velopment and helps in building consensus among stakeholders (Goudarznia *et al*., 2017; Tomkins & Lange, 2019). Conventional outdoor landscape visualization methods using AR require users to wear a head-mounted display (HMD), such as Microsoft HoloLens (2016), or to hold a smartphone or tablet (Haynes *et al*., 2018; Tomkins & Lange, 2021). Consequently, there is a physical distance limitation between the AR device (e.g. HMD or smartphone) and the AR user, and the perspective of the AR device's camera is limited to the view from the AR user's area of action. In addition, good information graphics are needed that
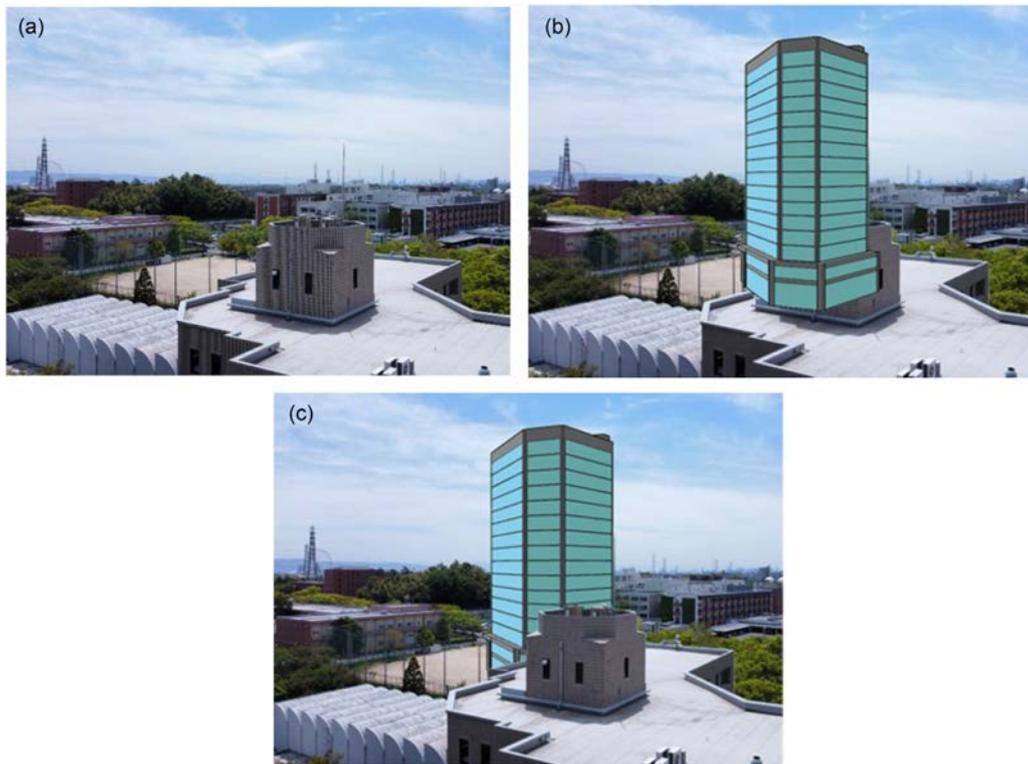
Figure 1: Occlusion problem in AR and occlusion handling effect. (a) Real world. (b) No occlusion handling (mock-up, incorrect AR). (c) Occlusion handling (mock-up, correct AR).

allow people to understand the data at both the micro and macro levels (Offenhuber & Seitinger, 2014).

The research and development of unmanned aerial vehicles, commonly called drones, has advanced greatly, and the use of drones in various applications has been proposed. For example, drone delivery services for small packages have attracted interest; Amazon (2013) is developing a drone delivery service (Hong et al., 2018), and drones are also used for both traffic and crowd monitoring (Motlagh et al., 2017; Elloumi et al., 2018). Drones provide new perspectives that are not visible from the user's range of action (Gallacher, 2017). To visualize nonexistent landscapes, both past and future, from the air, a method that integrates AR and drones has been proposed (Koch et al., 2011; Wen & Kang, 2014; Unal et al., 2018, 2020; Yan et al., 2019). This integrated AR–drone method enables the perception of augmented macro-level information by adding digital information to the aerial video acquired from the drone. However, the integrated AR–drone method has an occlusion problem (Fig. 1b). In general, virtual objects in AR are rendered last in relation to objects in the real world, and without intervention, a virtual object is displayed in front of real objects. This discrepancy between the relative placement of a real object and that of a virtual object (i.e. determining which one is behind the other) is called the occlusion problem (Fig. 1c). Occlusion handling is a method for solving this problem and is essential for clarifying the position of the virtual object (Zollmann et al., 2020). Model-based occlusion handling (Kasperi et al., 2017; Gimeno et al., 2018; Li et al., 2018; Evangelidis et al., 2021) uses a three-dimensional (3D) virtual model (occlusion model) based on the real surrounding environment, and the design target model and occlusion models are created in the virtual scene during preprocessing. The method deals with occlusion by comparing the depth of the 3D virtual model of the

design target with that of the occlusion model. In integrating AR and drones, if a city 3D model is available, then occlusion handling can provide a realistic representation of the city (Koch et al., 2011).

A city 3D model is made of 3D geospatial data that reproduce a real-world (physical space) city in a virtual world (cyberspace). These models have become an essential platform for co-designing with residents in urban spaces (Ruohomäki et al., 2018). The ideal city digital twin would be a city 3D model that encompasses economic, ecological, and demographic conditions and changes (Dembski et al., 2020). A digital twin is defined as "a digital replica of a physical asset, process, or system" (Grieves, 2014). The digital twin and the physical twin continuously interact (Batty, 2018). From the perspective of architecture and urban design, the use of digital twins in cities is still in its infancy, and no methods utilizing them have been established. AR and virtual reality (VR) are gaining attention as tools for visualizing city digital twins (Mohammadi & Taylor, 2019). However, there are fewer studies in which AR has been used to visualize city digital twins (Schrotter & Hürzeler, 2020) than those in which VR has been used (Dou et al., 2019; Dembski et al., 2020; Ham et al., 2020; Herman et al., 2021; White et al., 2021). Because of the limited view available in AR while using city digital twins, it is challenging to obtain a holistic view of the city using AR, whereas this is possible with VR.

As city 3D models evolve into city digital twins, they become more detailed, larger, and require many resources for AR rendering. Therefore, a high-performance PC is required for rendering the city 3D model when using AR. This makes it difficult to implement the AR system outdoors due to the weight, power supply, and communication environment required. To achieve rapid consensus building among stakeholders, it is

necessary to design an AR system that can be easily implemented outdoors.

In this research, we use detailed city 3D models to develop a digital-twin approach to outdoor AR with occlusion handling for both first-person and bird's-eye views. The internet-based system architecture is built in an integrated AR–drone system. Multiple stakeholders involved in building construction projects can view aerial perspectives, both on-site and off-site via an internet browser, using AR videos occluded with a city 3D model. An earlier version of this paper was presented at the Education and Research in Computer Aided Architectural Design in Europe (eCAADe) 2021 conference (Kikuchi *et al.*, 2021). Important subsequent advances in our research are presented in this paper.

2. Literature Review

2.1. Integration of AR and drones

AR technology has been integrated with drone technology for AR visualization from aerial viewpoints. Liu *et al.* (2021) developed an AR visualization method for building inspections that concatenates aerial drone animations and building information model (BIM) animations of the target building while it is being navigated. They used intersection over union (IoU) to quantitatively evaluate the degree of matching between the aerial images and BIM images. The system used the drone for preprocessing, so the AR rendering and real-world photography were not synchronized. Some studies have reported methods for synchronizing AR rendering with real-world drone photography. Koch *et al.* (2011) proposed a method of visualization that integrates drones and marker-based AR. Wen and Kang (2014) proposed integrating AR and a drone to display a 3D virtual model in the real world from an aerial perspective. They designed and built the prototype from scratch, and the system's applicability to other sites was low.

A method using commercially available drones was proposed by Unal *et al.* (2018) in which GPS- and location-based AR were integrated with a drone to display 3D virtual models of cultural heritage sites from an aerial perspective. In a subsequent paper, Unal *et al.* (2020) proposed two methods for integrating drones and AR, one involving location-based positioning using GPS and accelerometers and the other involving vision-based positioning using videos from a monocular camera attached to the drone. Yan *et al.* (2019) proposed a landscape visualization method that involved AR–drone integration and that used simultaneous localization and mapping. However, the method has not yet been implemented because the drone's captured videos cannot be handled by the AR engine.

In the conventional method using commercially available drones, the integration of AR and drones was based on the software development kits (SDKs) specific to each technology, and the system integration required the use of devices that were compatible with the SDKs. In addition, the existing method of integrating AR and drones still suffered from the occlusion problem in which the 3D virtual model displayed in the real world is shown in front of the real-world object. The occlusion problem is described in detail in Section 2.2.

2.2. Occlusion handling method

Occlusion problems can mislead AR users about the positions of 3D virtual models (Tian *et al.*, 2015). To solve the occlusion problem, depth-based, foreground object-based, and model-based methods have been proposed.

In depth-based methods, both RGB-D cameras and monocular cameras have been used to obtain the depth information for the occlusion target. They have also been used to handle the occlusion by comparing the depth information obtained from the camera with the depth of the virtual object (Holynski & Kopf, 2018; Valentin *et al.*, 2018; Du *et al.*, 2020). Although the method can handle occlusion in real time, the depth information that the 3D sensing camera can obtain is limited, making it unsuitable for use in wide outdoor spaces.

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. Semantic segmentation has also been used to extract foreground objects and to handle occlusion (Roxas *et al.*, 2018; Kido *et al.*, 2021). Because the object-based foreground method analyses the foreground object and the 3D virtual model before and after the 2D image, it cannot handle positional movement of the drone camera; looking around the 3D virtual model displayed in the real world from the air is an example.

Model-based methods handle occlusions by comparing the depth of the 3D virtual model with the depth of the occlusion model (Kasperi *et al.*, 2017; Gimeno *et al.*, 2018; Li *et al.*, 2018; Evangelidis *et al.*, 2021). Although the model-based methods require the preliminary creation of an occlusion model, they provide robust occlusion handling when the occlusion model is a static physical model whose shape and position do not change irregularly with time, such as a building. In integrating AR and drones in an outdoor space, if city 3D models are available, then model-based occlusion handling can provide a realistic representation of the virtual model (Koch *et al.*, 2011).

2.3. Toward the city-digital-twin approach

Digital twin technology is rapidly developing toward being fully used for urban development (Deng *et al.*, 2021). In various cities, 3D models are being created to create the cities' digital twins. In Singapore, in the Virtual Singapore project (2017), and in Japan, in the PLATEAU project (2020), 3D models of urban spaces are being created. Dembski *et al.* (Dembski *et al.*, 2020) studied a city digital twin for urban planning in Herrenberg, Germany. A city digital twin is also being developed for Zurich, Switzerland to support decision making in urban planning (Schrotter & Hürzeler, 2020).

The research community has identified the need for advanced technological tools and applications that improve the visualization, realization, and management of city digital twins (Shahat *et al.*, 2021). AR and VR are tools for visualizing city digital twins that can facilitate nonexpert citizens' participation in the decision-making process (Mohammadi & Taylor, 2019). However, fewer studies have used AR (Schrotter & Hürzeler, 2020) than VR (Dembski *et al.*, 2020; Ham *et al.*, 2020; Schrotter & Hürzeler, 2020; Herman *et al.*, 2021; White *et al.*, 2021) to visualize city digital twins.

In urban spaces, residents should perceive information at both the micro and macro levels (Offenhuber & Seitinger, 2014). This study defines micro-level information as that obtained in the near-ground view and macro-level information as that obtained in the far-ground view. The BIM-based VR method (Kim *et al.*, 2021) and AR method (Fukuda *et al.*, 2019) for architectural visualization have been proposed to provide real-scale advanced micro-level information. However, due to the limitation of the physical distance between the AR device and the AR user, outdoor AR using a city model has not yet provided a holistic view of a full-scale city as VR has.
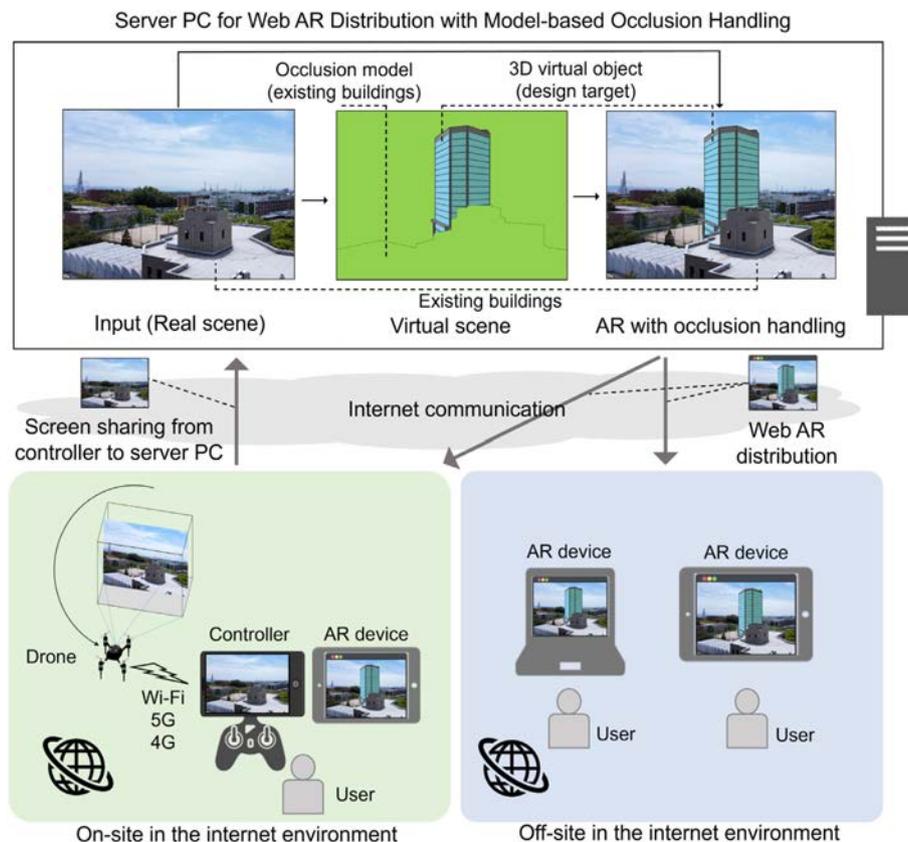
Figure 2: Conceptual diagram of the proposed method.

In the City Geography Markup Language (CityGML) (2021), which is a standardized data format for storing city 3D models, there are four levels of detail (LoD): digital terrain model (LoD0), block model (LoD1), roof model (LoD2), and exterior model (LoD3). The city digital twin envisions an LoD3 model as the ideal goal, but most models are either LoD1 or LoD2 at this stage. As the model of the city develops into a city digital twin, many LoD3 models will be added, vastly increasing the amount of data in it. As a result, the model of the city will require a large amount of resources for VR and AR rendering. One way to solve this problem is to use a high-performance PC to perform VR and AR rendering.

3. Research Contribution

In this study, we developed a digital-twin approach to landscape visualization by building an internet-based system architecture in an integrated AR–drone system, enabling us to obtain both first-person and overhead views for outdoor AR with occlusion handling using a detailed city model.

Our system uses universal technologies, such as virtual cameras and the screen-sharing functions of online meeting applications, to integrate AR and drones into the system without using SDKs. In addition, as city 3D models develop into city digital twins, the AR system configuration must cope with large amounts of data, and thus a high-performance PC is needed for AR rendering. However, problems with the weight, power supply, and communication environment of high-performance PCs make AR systems difficult to operate outdoors. Designing an outdoor AR system is necessary for achieving rapid consensus building among stakeholders. Therefore, our system uses a server PC to perform AR rendering with occlusion handling using a city 3D model and does not require a PC to be present at the construction site.

In the earlier version of this paper (Kikuchi *et al.*, 2021), we did not quantitatively evaluate the occlusion handling accuracy, and the experimental location was limited to one site. Therefore, we conducted an outdoor AR experiment to determine whether the proposed system can handle occlusion at two sites from the perspective of a drone camera that is moving in the air in an outdoor space, and IoU was used to quantitatively evaluate the occlusion handling accuracy (Liu *et al.*, 2021). The proposed method is expected to help stakeholders both understand and share plans for building, improve reviewing during the planning and design phase of development, and contribute to accelerating large-scale construction projects.

4. Methodology

We use a drone that acquires both first-person and overhead views at the building construction site, a controller that operates the drone, and a PC on a server that performs AR rendering with occlusion handling. Various AR devices, such as smartphones and PCs, are used by multiple users off-site and on-site. These devices are connected via internet communication.

The conceptual and flow diagrams for our method are shown in Figs 2 and 3, respectively.

4.1. Preprocessing for model-based occlusion handling

Occlusion handling is performed in the model-based method. As a preprocessing step, the occlusion model and the 3D virtual
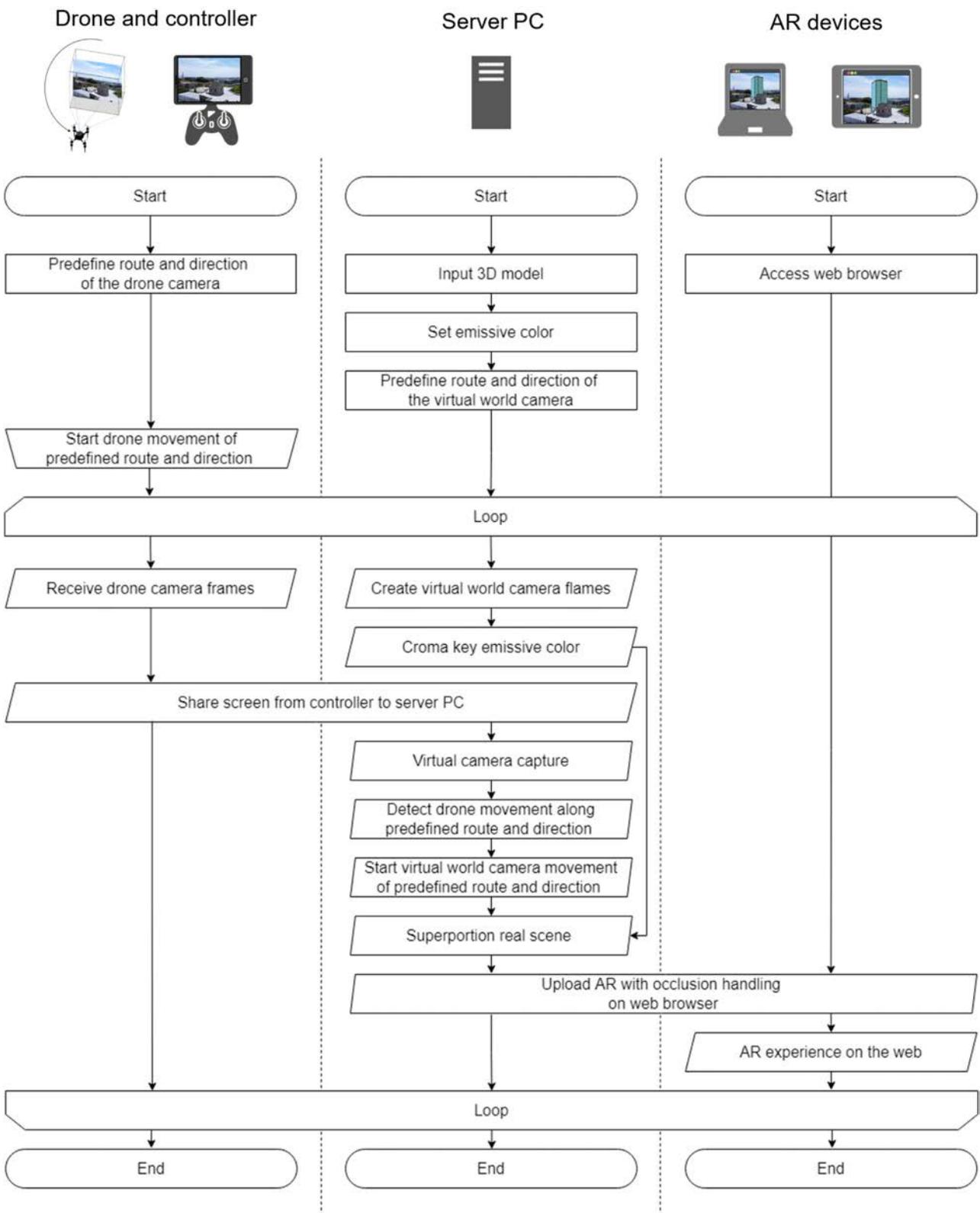
Figure 3: System flow of the proposed method.

model of the design target are created and placed in an appropriate positional relationship in the real world. To superimpose the occluded 3D virtual model of the design target on the real world during the chroma keying process, everything except the 3D model of the design target is changed to an emissive color that does not exist in the real world. Figure 4 shows how this is done. A navigational route is set up for the cameras in both worlds to synchronize the position and direction between the
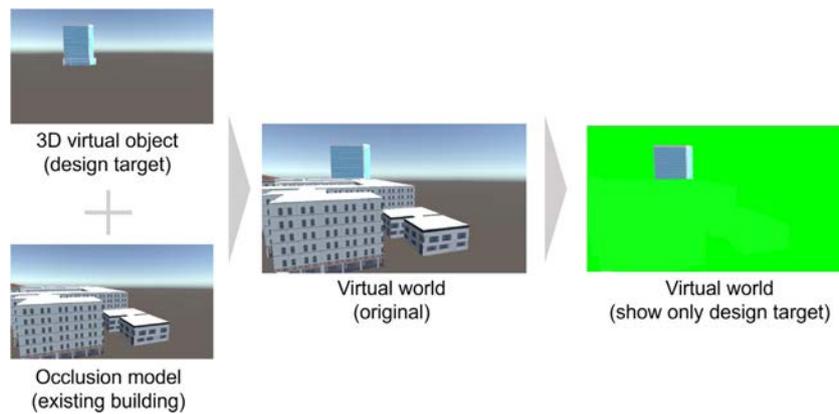
Figure 4: Example of changing the color settings for the background and occlusion models.
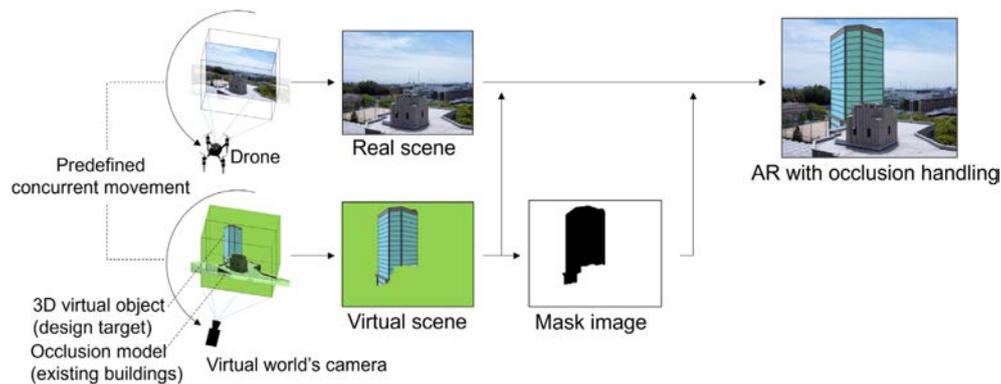

Figure 5: Conceptual diagram of AR rendering with model-based occlusion handling.

drone camera in the real world and the camera in the virtual world.

4.2 Internet-based integration of augmented reality and a drone for a city-digital-twin approach

A high-performance PC is required for the AR rendering with occlusion handling of detailed city models as occlusion models. A high-performance PC in this study is defined as a PC equipped with a GPU and CPU that enable AR rendering with occlusion handling using a detailed city digital twin. First, on-site video of a building construction project is captured in real time by a camera attached to the drone and is displayed on the controller's screen, which is a smartphone. The screen-sharing function of an online meeting application is used to communicate video over the internet between the high-performance server PC and the drone controller, and to take the video on the controller and display it on the high-performance PC's screen. The screen-sharing feature of online meeting applications shares the video on one screen with multiple devices in the internet environment. As in the earlier version of this paper (Kikuchi *et al.*, 2021), by using a virtual camera, the drone video displayed on the high-performance PC can be processed without using the drone SDK. The virtual camera treats the PC screen as a webcam video. Image processing is used to detect changes in the controller's video when the drone begins to move and to synchronize the start of movement of the drone's camera with the virtual world's camera.

4.3. AR rendering with model-based occlusion handling

Our method performs AR rendering with occlusion handling on a server PC. Figure 5 shows a conceptual diagram of the rendering with occlusion handling. First, to superimpose the 3D virtual model of the design target onto the video from the drone's viewpoint, the emissive color regions of the video acquired from the virtual world's camera are masked using chroma keying, in which mask processing based on color information is performed. Mask processing is an image processing technique that displays a specific range on an image or video while hiding the rest of the range. AR videos with occlusion handling are generated by synchronizing the start time of the drone's camera in the real world and in the virtual world and then superimposing the videos from the viewpoints of the two cameras while flying along a predetermined route.

4.4. Web AR distribution for multiple users

The AR system needs to be user friendly to encourage public participation. Web technology supports smartphones and PCs with various operating systems and specifications. It provides multiple means of access to stakeholders, including citizens, and is one of the best ways to encourage citizen participation. Streaming software delivers the AR video that is created on the server PC live to the available video delivery platform on the website (Fig. 6). By using a video delivery platform that can be used on
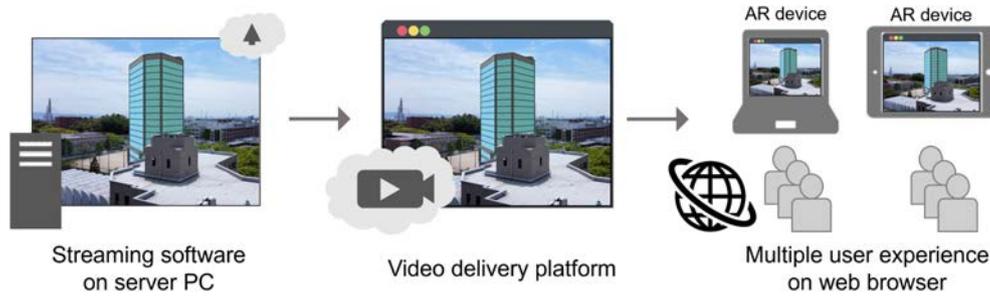
Figure 6: Multiple user AR experiences in a web browser using streaming software and a video delivery platform.

Table 1: Hardware of the prototype system.

Type	Device information	
Drone	Mavic Mini	
	Resolution	2.7K: 2720 × 1530, 30 fps
		Full HD: 1920 × 1080, 60 fps
	FOV	83°
	Horizontal FOV	68.0° (measured experimentally)
	Weight	199 g
Controller	iPhone 11	
	OS	iOS v14.4
Server PC	Self-built PC	
	OS	Windows 10 Education 64 bit
	CPU	Intel (R) Core (TM) i7-8700K CPU @ 3.70 GHz
	GPU	Geforce GTX 1080 Ti
	RAM	32 GB
AR device	iPad	
	OS	iOS v14.4

websites, users can easily watch an AR video from various sites on the internet by accessing web pages via URLs or QR codes. This can be done entirely by using their own devices, such as smartphones and PCs.

5. Experiments and Results

In this section, we describe the verification experiments that were conducted in two field tests to evaluate the occlusion handling accuracy, the communication speed, and the system latency of the prototype system that we built to test our method.

5.1. Development of the prototype system

A prototype system was constructed to verify the applicability of the method proposed in Section 4. The prototype system hardware was as follows. A self-built PC was used as the server PC to generate an AR video with occlusion handling and to upload it to the web. To obtain the first-person and overhead views in AR, a drone (Mavic Mini, DJI) weighing only 199 g that could fly in densely populated areas in Japan was used. An iPhone 11 with an installed drone control application was used as the drone controller. In addition, an iPad was used as the device for checking the uploaded AR video. Table 1 shows the hardware specifications of the prototype system.

The field of view (FOV) of the drone is listed as 83° by the manufacturer, but the horizontal FOV value is not provided. To superimpose the view from the camera in the virtual world onto the drone's camera, it is necessary to set the horizontal FOV of the virtual world's camera to the same value as that of the drone's
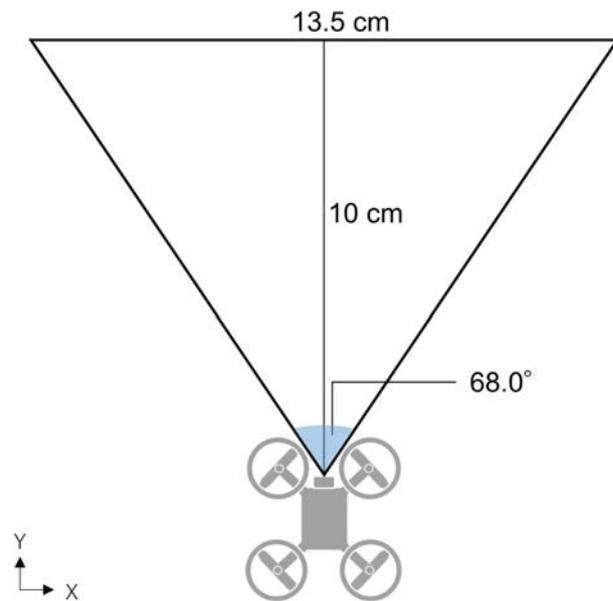

Figure 7: Experiment to calculate the horizontal FOV.

camera. We experimentally determined the horizontal FOV of the drone's camera to be 68.0° (Fig. 7).

5.1.1. Preprocessing for model-based occlusion handling
To compare the occlusion handling accuracy for LoD1 and LoD2, 3D models were created for LoD1 using the

Table 2: Details of the field test at site A.

Field test	Detail	
A	Prerecording date	16:00 JST on 20 March 2021
	Postprocessing date	16:00 JST on 24 September 2021
	Field test location	Osaka University Suita Campus, Suita, Japan

infrastructure design software InfraWorks 2019 (Autodesk) and for LoD2 using Metashape Professional (v1.6.0 build 9925 64 bit, Agisoft), which is a tool for the photogrammetry pipeline that can create 3D models by structure from motion (Westoby *et al.*, 2012). InfraWorks 2019 determines the position of the 3D model of the design target based on the base map information from the Geospatial Information Authority of Japan. In addition, the 3D modeling software SketchUp Make 2017 (Trimble Inc.) was used to create a virtual 3D model as a design target. The Unity (2020.1.7f1 64-bit) game engine was used to build the virtual world in which 3D models of existing buildings (occlusion models) were created. The navigation routes in the real and virtual worlds were chosen so that the position and direction of the drone camera in the real world and those of the camera in the virtual world were synchronized. Using the automatic drone navigation function of the drone control application DJI Fly (iOS v1.4.8, DJI), the route of the drone camera in the real world was preset. The virtual world's camera was then preset to move along this same route by programming in C# in Unity. To superimpose the 3D virtual model of the design target, which was occluded by chroma keying, onto the real world, the existing 3D model of the building and the virtual background were changed to emissive colors with RGB values [0, 255, 0], which do not exist in the real world.

5.1.2. *Internet-based integration of augmented reality and a drone for a city-digital-twin approach*

Microsoft Teams (server PC: v1.4.00.22 976 64 bit; controller: iOS v3.14.0), an online meeting application with a screen-sharing function, was used to display the same screen on different devices, such as smartphones and PCs, over the internet. Microsoft Teams was used to send video between the high-performance server PC and the drone controller in real time over the internet and to display the video from the controller on the high-performance PC. We used the streaming software OBS Studio (v27.0.1), which distributes video, such as PC screens, in real time and has a virtual camera function.

5.1.3. *AR rendering with model-based occlusion handling*

To generate an AR video with occlusion handling from the perspective of a drone flying along a preset route, the chroma keying function of OBS Studio was used.

5.1.4. *Web AR distribution for multiple users*

OBS Studio was used as the video distribution platform. To output the AR video on the web, YouTube was used as the video delivery platform. Google Chrome (v93.0.4577.82) was used as the web browser for watching YouTube.

5.2. Evaluation of occlusion handling and AR video data communication

To determine whether the proposed system is capable of occlusion handling from the perspective of a drone camera moving in the air in an outdoor space, outdoor AR experiments were con-

ducted at two sites to evaluate the accuracy of occlusion handling, communication speed, and system latency.

IoU, which is a segmentation metric, was used to evaluate the occlusion handling accuracy because it is a quantitative measure of the consistency between the occlusion target building and the occlusion model (Liu *et al.*, 2021):

$$IoU = \frac{TP}{TP + FP + FN}. \tag{1}$$

In equation (1), TP is true positive, which is the area that correctly describes parts that are buildings; FP is false positive, which is the area that describes parts that are not buildings as buildings; and FN is false negative, which is the area that describes parts that are buildings as not buildings.

The internet speed measurement service fast.com (the website; Fast.com, 2021) was used to measure the upload and download speeds of the controller (smartphone), the server PC for AR rendering with model-based occlusion handling, and the AR device. The communication speed indicates the amount of data that can be sent and received in 1 s. The communication speed of the prototype was measured 10 times, and the average value was obtained.

Latency is the time it takes to transmit data between devices. The controller–server PC latency was determined by the time difference between the transmission of video from the controller and the display of that video on the server PC. The latency between the server PC and the AR device was determined by the time difference between when the video for the AR device was sent from the server PC and when it was actually displayed on the device. The latency of the prototype was measured 10 times, and the average value was obtained.

5.2.1. *Field test at site A*

At the first site, a 3D model of a new building was superimposed on the parking lot of Osaka University Dental Hospital, Osaka University Suita Campus in Suita, Japan, from the perspective of the ground. Table 2 shows the details of the experiment.

The 3D model of the design target that was created in Sketch Up Make 2017 is shown in Fig. 8. This study focuses on model-based occlusion handling using a detailed city model, so we use a 3D model of only a simplified design target.

The occlusion target was the existing building between the AR camera and the 3D model of the design target. The occlusion model of the existing building created in InfraWorks is shown in Fig. 9.

The 3D model featuring both the design target building and the occlusion model in the virtual world is shown in Fig. 10.

Figure 11 shows the layout of the design target and the existing buildings as well as the route and direction of the camera. The drone was automatically navigated in an arc around the center point in Fig. 11, facing the center at 40 ± 1 m altitude, and stopped after 16 s.

Figure 12 shows the current state of the site.

Figure 13 shows snapshots of the AR video with occlusion handling.

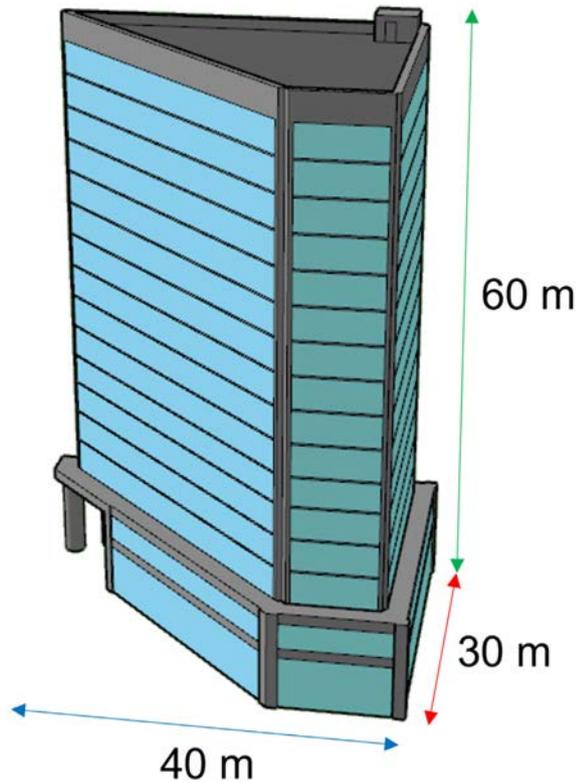Figure 14 provides a more detailed example of the results.

Figure 8: The 3D model of the design target that was created in Sketch Up.

and a latency (controller–server PC) of 0.2 s, we concluded that the effect of using the prerecorded video was small. The drone video was prerecorded at 16:00 Japan Standard Time (JST) on 20 March 2021, and the weather was cloudy. Postprocessing of the drone videos was conducted at 16:00 JST on 24 September 2021.

Figure 16 shows the occlusion handling accuracy over time.

The specifications of the prototype system used in the field test at site A are shown in Table 3, the verification results of the occlusion handling accuracy are shown in Table 4, the communication speed measurements are shown in Table 5, and the latency measurements are shown in Table 6.

The results demonstrated that AR with occlusion handling is feasible using an occlusion model (existing building), which is part of a city 3D model, from the viewpoint of a drone camera moving along a preset route in the air in an open outdoor space. The occlusion handling accuracy, processing speed, communication speed, and latency of the proposed system were also determined. The average IoU value, indicating the occlusion handling accuracy, was 0.820. We also confirmed that the occlusion handling accuracy decreased with time. The frame rate of the video acquired from the drone was about 30 fps, the frame rate of the video acquired from the camera in the virtual world was about 3000 fps, and the frame rate of the video output on the web was about 30 fps. The number of polygons drawn in the scene was 2.9×10^3, the number of vertices was 4.4×10^3, and the amount of data was 7.06 MB. The latency between the controller smartphone and the server PC was about 0.23 s, and the latency between the server PC and the AR device was about 2.814 s. The video was output on an iPad, which is a common device that users bring with them.

Figure 15 shows how the occlusion handling accuracy is visualized by superimposing part of the occlusion model (transparent shading) onto the drone video. Figure 15 was created with Adobe Photoshop CS4 (v11.0) using a specific range of the virtual world and real-world buildings. The experiment was performed twice to obtain the results shown in Figs 13 and 15. Using live drone video, which the device was designed for, the length and content of the live video were recorded each time the experiment was conducted. Therefore, we prerecorded the drone video and used the prerecorded video as the input video to perform the processing twice. Because the live drone video and the prerecorded video had a real scene frame rate of about 30 fps

5.2.2. Field test at site B

At the second site, a 3D model of a new building was superimposed on the parking lot of the Osaka University Dental Hospital from the perspective of the Osaka University Suita Campus athletic field. Table 7 shows the details of the experiment. The 3D model of the design target was the same as in the first validation (Fig. 8). The occlusion target was a group of existing buildings between the AR camera and the 3D model of the design target. The occlusion model of the existing buildings was created from 369 photos taken by the drone by using structure from motion in Agisoft Metashape (Fig. 17).
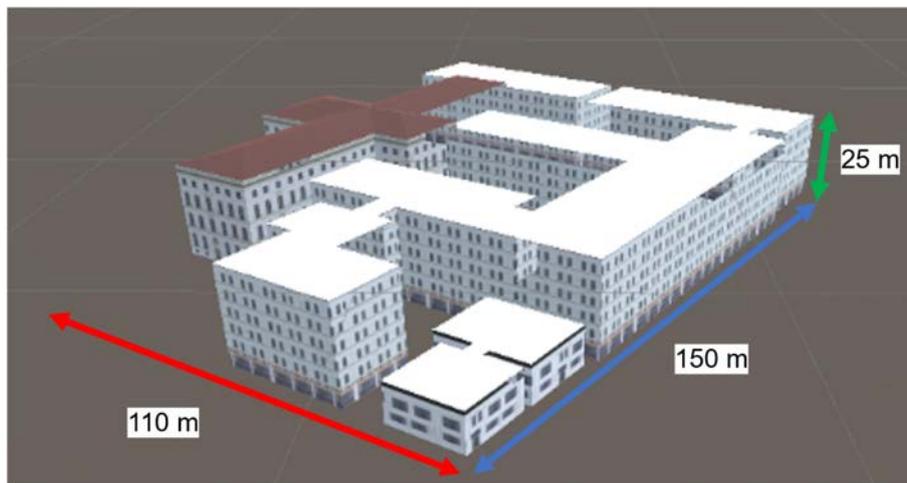

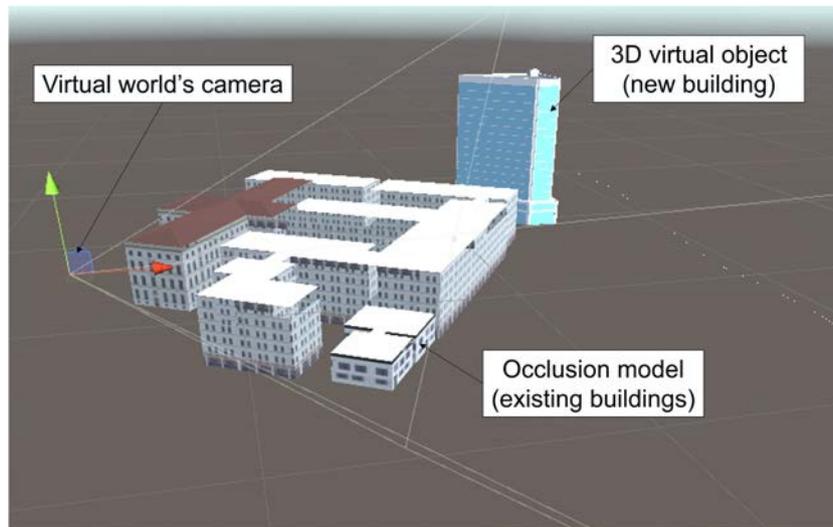**Figure 9:** LoD1 occlusion model created in InfraWorks.

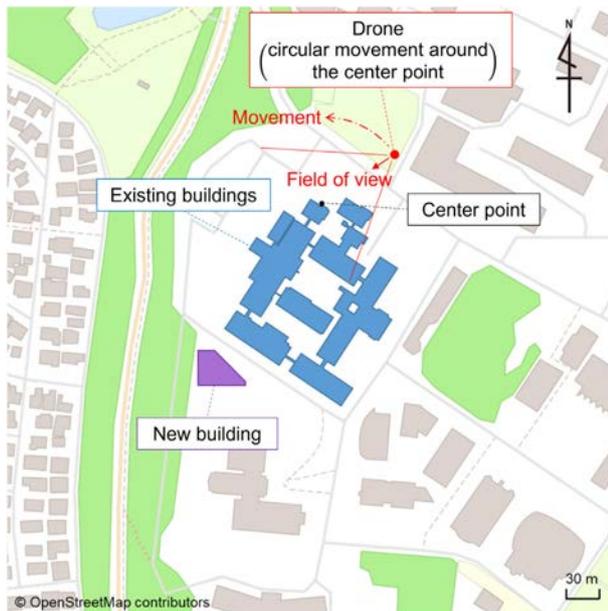Figure 10: Virtual world for the site A field test.

Figure 11: Map of the existing and design target (new) buildings for the site A field test.

Figure 12: Photograph of the current environment for the site A field test (20 ± 1 m altitude).

Figure 18 shows the placement of both the target 3D model and the occlusion model in the virtual world.

Figure 19 shows the layout of the design target and the existing buildings as well as the route and direction of the camera. The drone was automatically navigated in an arc around the center point in Fig. 19, facing the center at 60 ± 1 m altitude, and stopped after 16 s.

Figure 20 shows the scene during the outdoor AR experiment.

Figure 21 shows snapshots from the AR video with occlusion handling.

Figure 22 shows an example of the results in greater detail.

Figure 23 shows how the occlusion handling accuracy was visualized by superimposing the occlusion model (transparent shading) onto the drone video. Figure 23 was created with Adobe Photoshop CS4 (v11.0) using a specific range of the virtual world and real-world buildings. To obtain the results shown in Figs 21

and 23, the experiment was performed twice. Therefore, for the same reason as in Section 4.2.1, the drone video was prerecorded, and the processing was performed twice using the prerecorded video as the input. The real scene frame rate and latency (controller–server PC) for both the live drone video and the prerecorded video were about 30 fps and about 0.2 s, respectively; thus, the effect of using the prerecorded video was small. The drone video was prerecorded at 06:00 JST on 30 August 2021, and the weather was clear. Postprocessing of the drone videos was done at 07:00 JST on 2 October 2021.

Figure 24 shows the occlusion handling accuracy over time.

The specifications of the prototype used in the field test at site B are shown in Table 8, the verification results of the occlusion handling accuracy are shown in Table 9, the measurement results of the communication speed are shown in Table 10, and the measurement results of the latency are shown in Table 11.

As in the first experiment, we demonstrated that AR with occlusion handling could be performed in the model-based method using the occlusion model (existing buildings), which was part of the city 3D model. The occlusion handling accuracy in the second validation was 0.786, which was lower than the
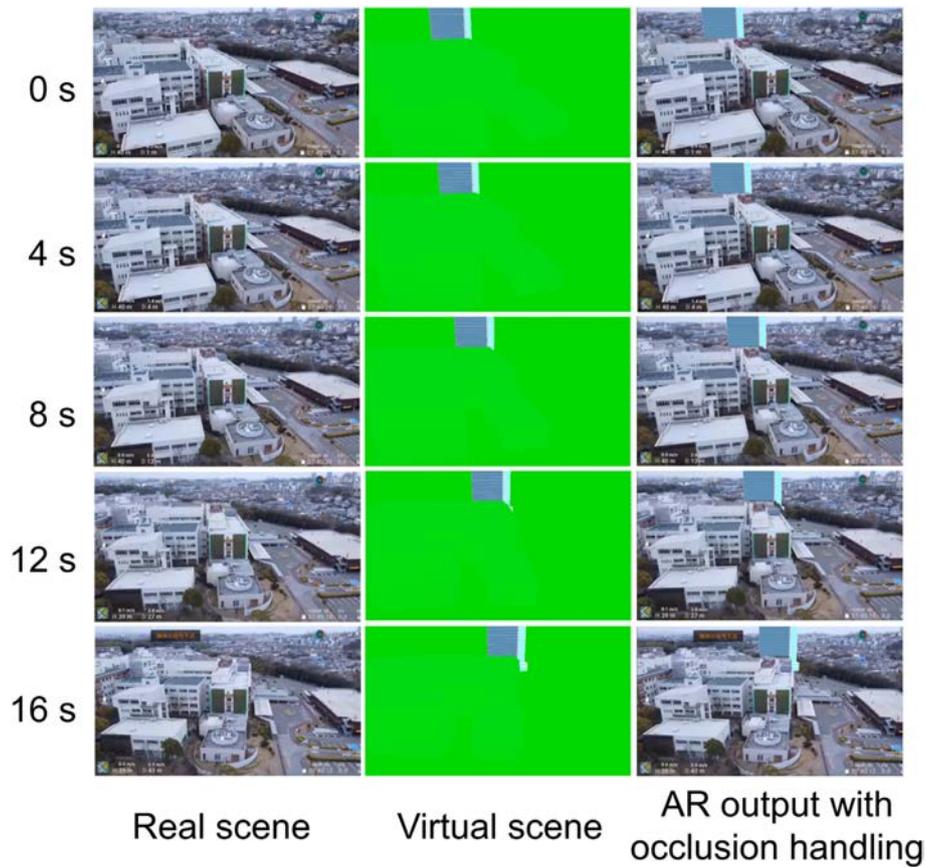
Figure 13: Snapshots from the site A field test (40 ± 1 m altitude).

Figure 14: Example of the results from the site A field test (40 ± 1 m altitude).

average IoU value of 0.820. As in the first verification, the occlusion handling accuracy decreased with time. We measured the processing speed and latency for the entirety of the proposed system. The frame rate of the video acquired from the drone (real scene) was about 30 fps, the frame rate of the video acquired from the camera in the virtual world (virtual scene) was about 200 fps, and the frame rate of the video output on the web (AR output with occlusion handling) was about 30 fps.
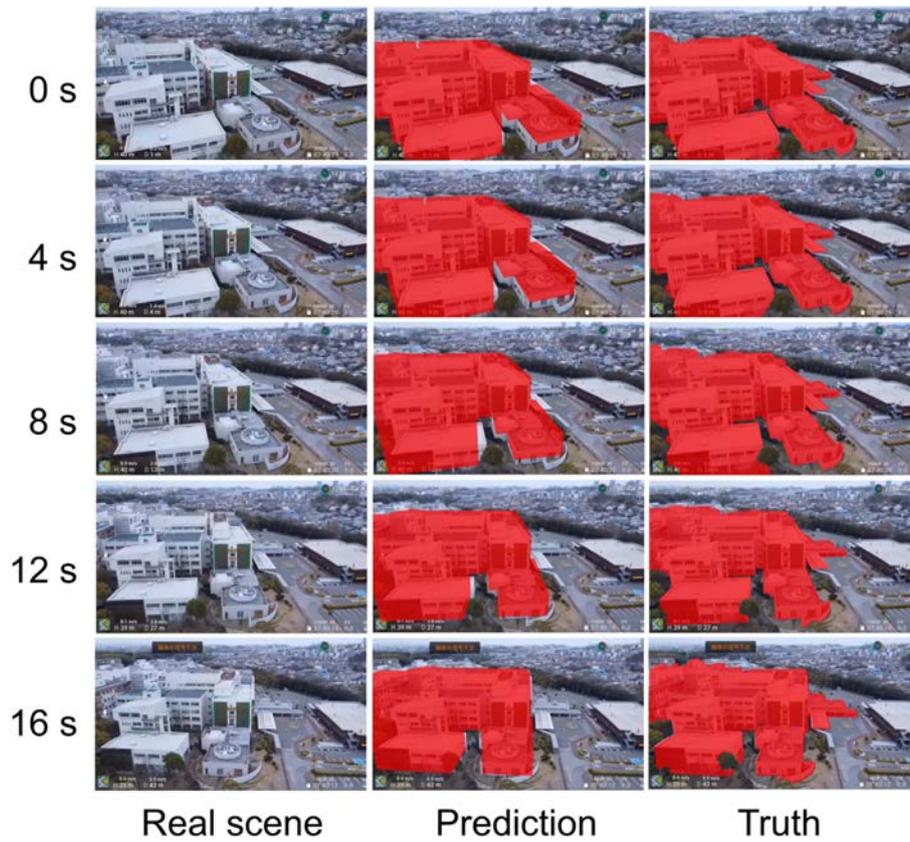
Figure 15: Occlusion handling accuracy in the site A field test (40 ± 1 m altitude).
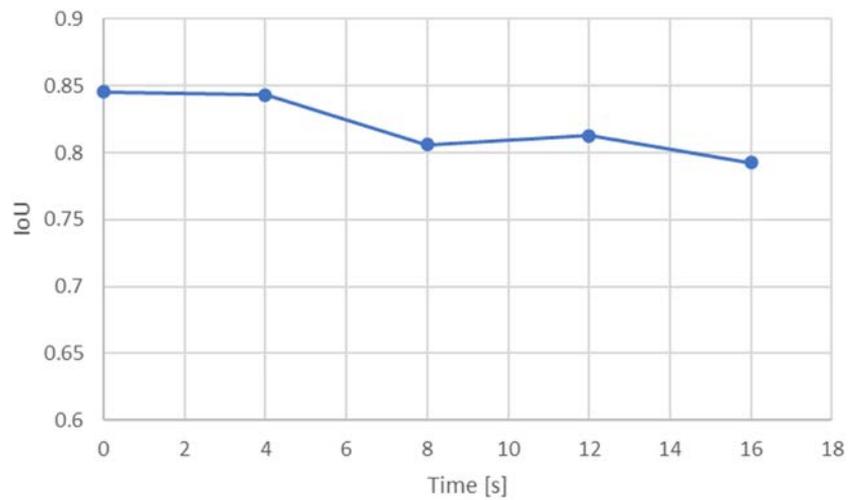

Figure 16: Occlusion handling accuracy as a function of time in the field test at site A.

Table 3: Specifications of the prototype used in the site A field test.

3D model	Number of polygons	2.9×10^3
	Number of vertices	4.4×10^3
Total system	Data amount (MB)	7.06
	Output frame rate (fps)	30

Table 4: Accuracy of the occlusion handling in the site A field test.

IoU targets	IoU
Occlusion target–Occlusion model	0.820

Table 5: Internet speed in the site A field test.

Device	Internet speed (Mbps)	
	Download	Upload
Controller internet communication speed	98.7	24.9
Server PC internet communication speed	758	882
AR device communication speed	81.4	16.8

Table 6: Latency in the site A field test.

Communication between devices	Latency (s)
Controller–server PC	0.230
Server PC–AR device	2.814

The number of polygons drawn in the scene was 2.09×10^7, the number of vertices was 1.38×10^7, and the amount of data was 8.46×10^2 MB, but the frame rate was about 30 fps, the same as the experiment at site A. The latency between the smartphone attached to the controller and the server PC was 0.248 s, and the latency between the server PC and the AR device was 3.223 s. As in the first experiment, the video output was sent to an iPad.

6. Discussion

This section discusses the performance of the prototype system and the specific applications to which this research can contribute and presents the limitations of this research.

6.1. Prototype system performance

In the prototype system, we showed that users could experience AR with model-based occlusion handling on a web browser from the perspective of a drone that automatically navigates along a preset route. The AR rendering was performed by a server PC, allowing detailed city 3D models to be handled. Our system uses universal technologies, such as virtual cameras and the screen sharing functions of online meeting applications, to enable internet-based system integration of AR and drones without specific SDKs.

The AR system was used at two outdoor sites to visualize the design target buildings. Occlusion handling accuracies of 0.820 and 0.786 were achieved at sites A and B, respectively. The IoU at which a building is adequately detected is 0.5 (Jabbar *et al.*, 2017), and the prototype exceeded this value, although the IoU did decrease over time (Figs 16 and 24). In addition, the occlusion handling accuracy for site A using the LoD1 model was higher than that of site B using the LoD2 model, which was presumably due to the vast scope of the verification target at site B. Because the scope of the verification target is large from a bird's-eye view, a comparison of the LoD1 and LoD2 models after expanding the scope of the verification target will be required in the future.

The system outputs the video in a web browser on an iPad, a common device that users may own. For sites A and B, the frame rate of the AR video that the users viewed was about 30 fps. This is higher than the 15 fps that is needed for humans to perceive video comfortably (Chen & Thropp, 2007). The latency was about 3 s, and thus the 3D model of the design target displayed on the AR device corresponded to what was displayed on the real-world video about 3 s before. Because the AR rendering with time synchronization and occlusion handling was performed by the server PC, we assumed that there was no gap

Table 7: Details of the field test at site B.
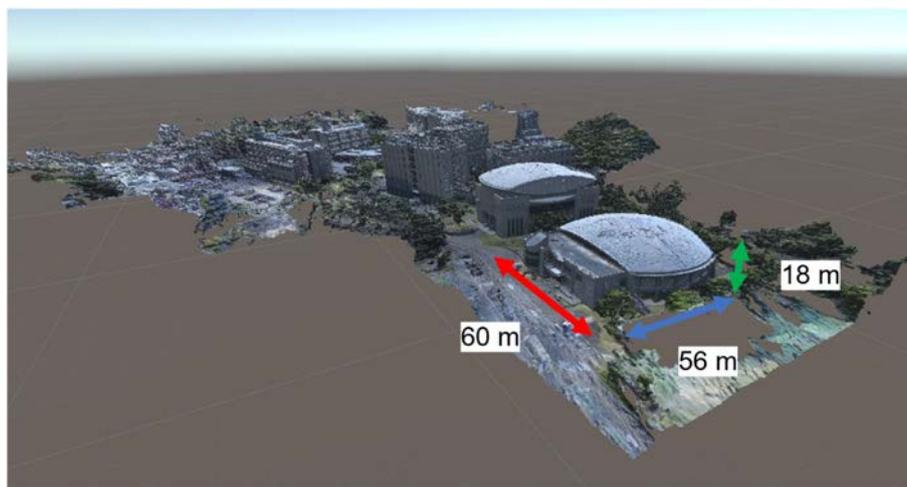
Field test	Detail	
B	Prerecording date	06:00 JST on 30 August 2021
	Postprocessing date	07:00 JST on 2 October 2021
	Field test location	Osaka University Suita Campus, Suita, Japan

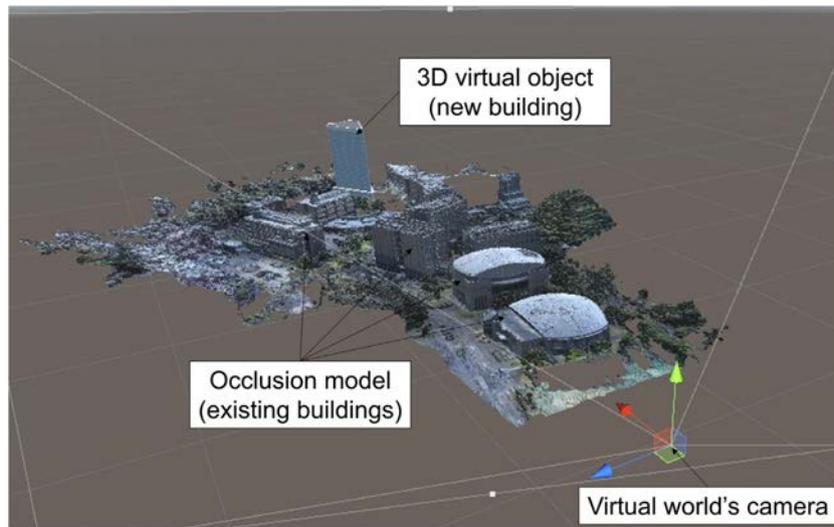**Figure 17:** LoD2 occlusion model created in Agisoft Metashape.

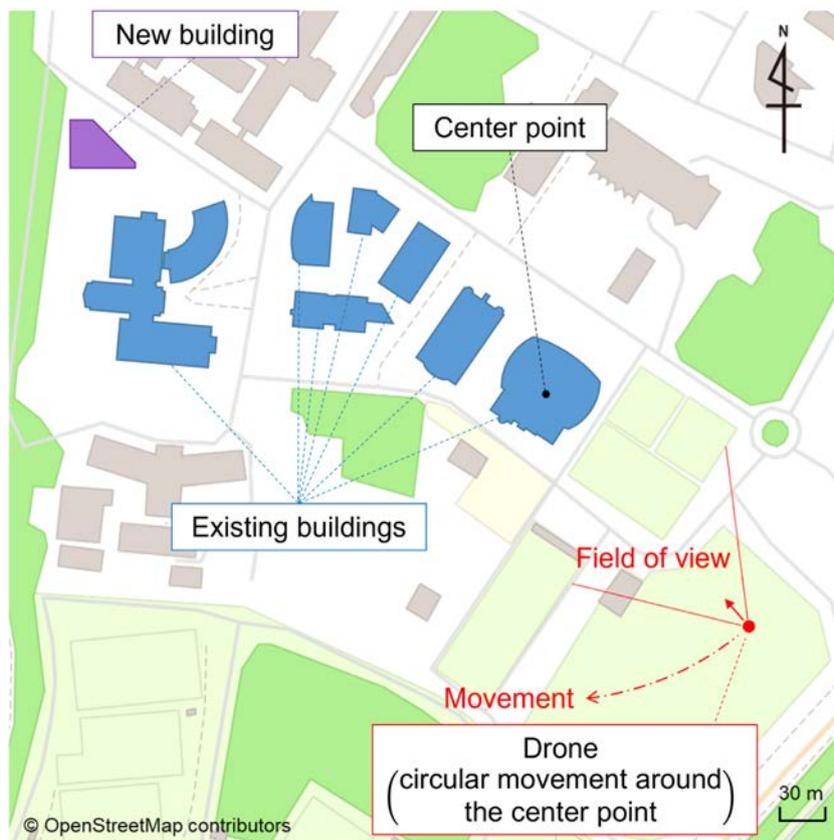**Figure 18:** Virtual world in the site B field test.

Figure 19: Map of the existing and design target (new) buildings in the site B field test.

between the real-world video and the 3D model of the design target caused by the latency.

The frame rate of the virtual scene on the server PC at site A was approximately 3000 fps, which was much higher than the virtual scene frame rate of approximately 700 fps in the earlier version of this paper (Kikuchi *et al.*, 2021). In the verification experiment at site B, which used an LoD2 model with a polygon count of 2.09×10^7, the virtual scene frame rate was also about 200 fps. The frame rate of the whole system was 30 fps, which is higher than the 15 fps threshold.

6.2. Limitations

There were several limitations in this research. The area around the boundary between the existing building and the 3D model of the design target was not accurately occluded, and thus the

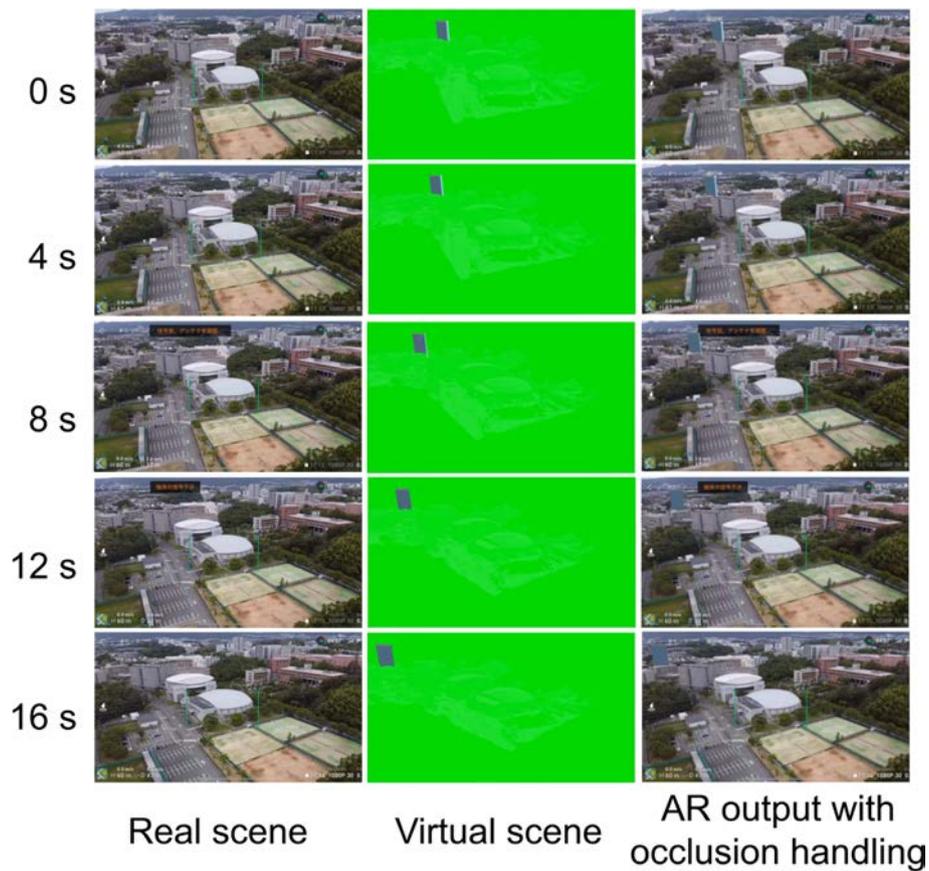Figure 20: The scene of the site B field test (40 ± 1 m altitude).

Figure 21: Snapshots from the field test at site B (60 ± 1 m altitude).

occlusion handling accuracy needs to be improved (Figs 25 and 26). The existing building and the 3D model of the design target were tilted (Fig. 26) due to the distortions inherent in the wide-angle lens cameras (the drone camera and the virtual world's camera). The IoU decreased over time, presumably because the position and direction of the drone's camera and of the virtual world's camera were not corrected, and thus their misalignment increased over time (Figs 16 and 24).
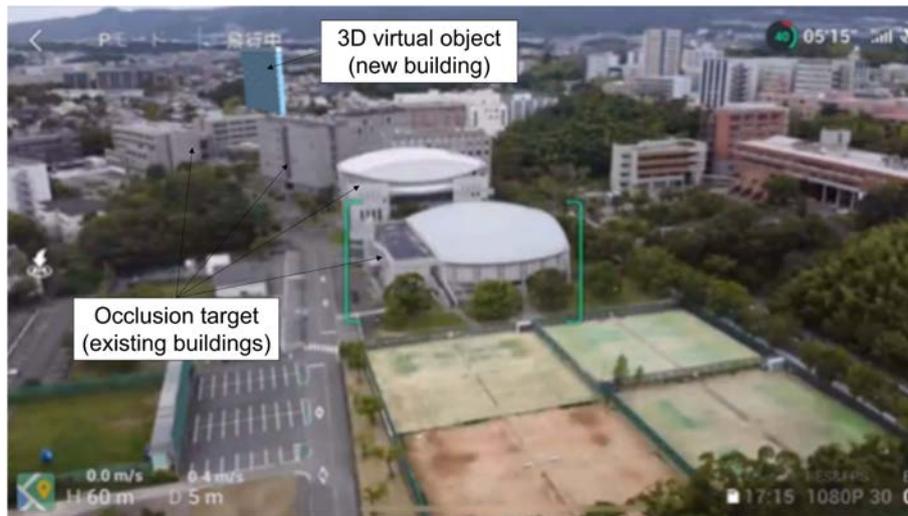
Figure 22: Example of the results from the site B field test (60 ± 1 m altitude).
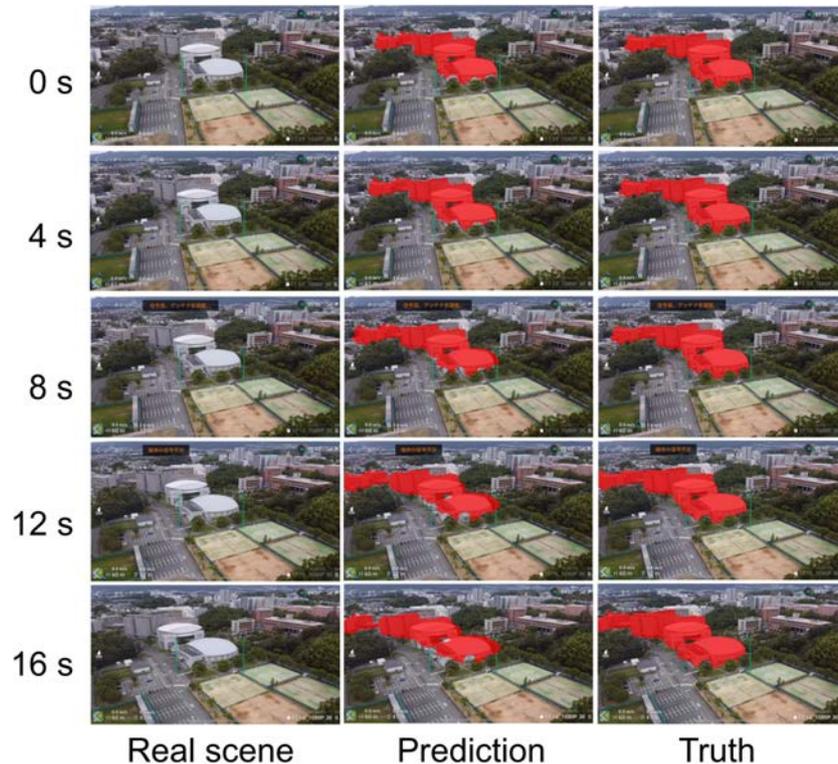

Figure 23: Occlusion handling accuracy in the site B field test (60 ± 1 m altitude).

The drone's navigational route and direction are predefined for AR alignment. The user cannot change the viewpoint during AR execution. Consequently, free-flying drones do not support our proposed method. Using information from the drone's internal sensors requires a drone-specific SDK, which is limited to model-dependent system configurations. If drone- or AR-specific SDKs are not used, the positional information for the drone's camera in flight is required to link the position and direction of the drone's camera with those of the virtual world's camera in real time. The proposed method also re-

quires accurate 3D modeling of the structures around the flight area.

7. Conclusions

We developed a digital-twin approach to landscape visualization that uses a detailed city model to achieve both first-person and overhead views in outdoor AR with occlusion handling. We used universal technologies, such as virtual cameras and the screen-sharing functions of online meeting applications, to achieve an
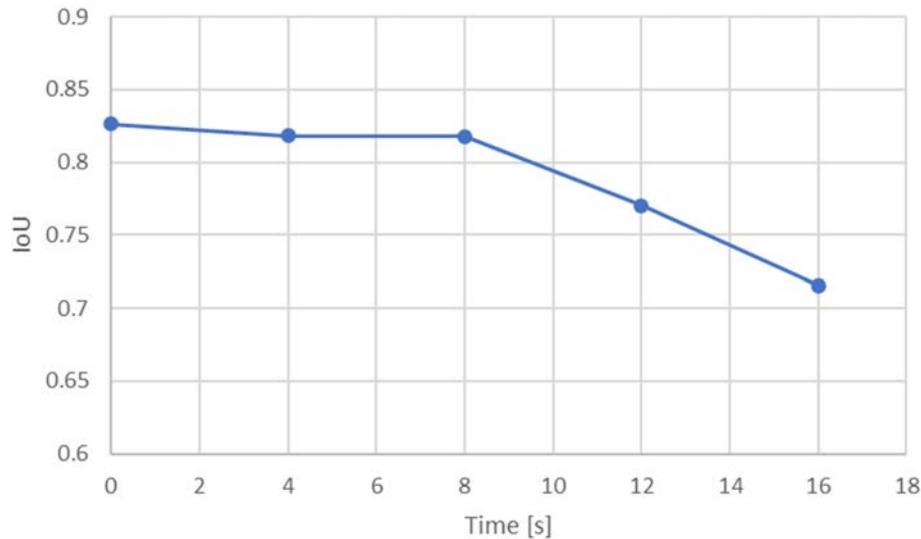
Figure 24: Occlusion handling accuracy as a function of time in the site B field test.

Table 8: Specifications of the prototype used in the site B field test.

3D model	Number of polygons	2.09×10^7
	Number of vertices	1.38×10^7
Total system	Data amount (MB)	8.46×10^2
	Output frame rate (fps)	30

Table 9: Accuracy of the occlusion handling in the site B field test.

IoU targets	IoU
Occlusion target–occlusion model	0.786

Table 11: Latency in the site B field test.

Communication between devices	Latency (s)
Controller–server PC	0.248
Server PC–AR device	3.223

internet-based system integration of AR and a drone without specific SDKs. Users could view the aerial perspective AR video with occlusion handling both on-site and off-site by using an internet browser.

In the verification experiments, we built a prototype system to evaluate our method and visualized a 3D model of the design target using several existing buildings as occlusion models in front of the new building. The occlusion handling accuracy of the prototype system was measured using IoU and was about 0.8. To improve the accuracy of occlusion handling, a method is required for linking the location information of the drone's camera with that of the virtual world's camera in real time. The ultimate objective is to build an AR method that includes a digital twin, one that allows the digital twin (virtual world) to respond continuously to the physical twin (real world) by linking the lo-

cation and sensor information of the real and virtual worlds via the drone.

The frame rate of the entire prototype system was about 30 fps even with the LoD1 and LoD2 occlusion models, which should be compatible with AR rendering that uses a detailed city model as a tool for visualizing a city digital twin. However, the latency was about 3 s and requires improvement. The noticeable latency arose because the system has a one-to-many structure in which one controller delivers to many clients via a single server PC. A new communication method, such as 5G, would allow for the use of a system structure with low video transmission latency between the controller and the client device.

This study makes two main contributions to improving, compared to the previous version of this study (Kikuchi *et al.*, 2021), the practicality of AR as a tool for visualizing the city digital twin. First, the system design means that it can be used outdoors, requiring only a drone, a controller (smartphone), and an AR device (smartphone) at a construction site. Second, the AR rendering with model-based occlusion handling is performed by a PC on a server, enabling a large city 3D model to be handled.

This research could help support business development and could improve reviewing in the planning and design stage. By

Table 10: Internet speed in the site B field test.

Device	Internet speed (Mbps)	
	Download	Upload
Controller internet communication speed	89.7	27.4
Server PC internet communication speed	748	874
AR device communication speed	79.8	18.9

Figure 25: Incorrect occlusion handling in the site A field test (40 ± 1 m altitude).

Figure 26: Incorrect occlusion handling in the site B field test (60 ± 1 m altitude).

using our AR system, stakeholders could check the 3D model of the proposed building, which is occluded, on a real scale and from the multiple angles given by the first-person and overhead viewpoints. In addition, the system can be easily operated using a drone, a controller (smartphone), and an AR device (smartphone) without transporting a high-performance PC to the project site. Furthermore, multiple remote users, at a head office or home, for example, could view the AR videos simultaneously on common devices, such as smartphones or PCs, using web browsers. Thus, regardless of location, stakeholders would be able to both understand and share proposed changes to the building environment during the planning and design stage, which would contribute to accelerating large-scale building construction projects. In addition, usability testing with end users, including citizens, is necessary for facilitating the acceleration of such construction projects in the future.

Interaction with light is one of the most important features in realistic 3D modeling. Use cases of light intensity visualization for interaction with light have been identified in the Virtual Singapore (2017) and PLATEAU (2020) city digital twins. Visualization of light changes to the design object and realistic shadows on neighboring buildings is a topic for future research. In addition, the method could be used to develop city-scale traffic and flood simulations from first-person and overhead viewpoints in AR by using a city model with LoD3 models placed on it. The proposed method is limited to predefined drone camera routes and directions and does not support free flight. In the case of the GPS-based free-flight method using a specific SDK, the system configuration depends on the specific SDK and is not versatile. The method of mounting a cell phone on a drone can permit free-flight AR by using a versatile GPS that is not dependent on a specific SDK.

Acknowledgments

Author contributions

Conceptualization, Naoki Kikuchi and Tomohiro Fukuda; supervision, Nobuyoshi Yabuki; project administration, Tomohiro Fukuda; investigation; Naoki Kikuchi; formal analysis, Naoki Kikuchi and Tomohiro Fukuda; software, Naoki Kikuchi; methodology, Naoki Kikuchi Tomohiro Fukuda and Nobuyoshi Yabuki; validation, Naoki Kikuchi; data curation, Naoki Kikuchi;

Conflict of interest statement

All the authors declare that there is no actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations.

References

Amazon. (2013). Amazon prime air. https://www.amazon.com/Amazon-Prime-Air/b?ie=UTF8&node=8037720011. Accessed on: October 9, 2021.

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355.

Batty, M. (2018). Digital twins. *Environment and Planning B: Urban Analytics and City Science*, 45(5), 817–820. https://doi.org/10.1177/2399808318796416.

Chen, J. Y. C., & Thropp, J. E. (2007). Review of low frame rate effects on human performance. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 37, 1063–1076. https://doi.org/10.1109/TSMCA.2007.904779.

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M., & Yamu, C. (2020). Urban digital twins for smart cities and citizens: The case study of Herrenberg, Germany. *Sustainability*, 12(6), 1–17. https://doi.org/10.3390/su12062307.

Deng, T., Zhang, K., & Shen, Z. J. M. (2021). A systematic review of a digital twin city: A new pattern of urban governance toward smart cities. *Journal of Management Science and Engineering*, 6(2), 125–134. https://doi.org/10.1016/j.jmse.2021.03.003.

Dou, S. Q. Q., Zhang, H. H. H., Zhao, Y. Q. Q., Wang, A. M. M., Xiong, Y. T. T., & Zuo, J. M. M. (2019). Research on construction of spatio-temporal data visualization platform for GIS and BIM fusion. In *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives* (Vol. 42, pp. 555–563). https://doi.org/10.5194/isprs-archives-XLII-3-W10-555-2020.

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., & Kim, D. (2020). DepthLab: Real-time 3D interaction with depth maps for mobile augmented reality. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (UIST '20), Association for Computing Machinery* (pp. 829–843). https://doi.org/10.1145/3379337.3415881.

Elloumi, M., Dhaou, R., Escrig, B., Idoudi, H., & Saidane, L. A. (2018). Monitoring road traffic with a UAV-based system. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 1–6). https://doi.org/10.1109/WCNC.2018.8377077.

Evangelidis, K., Papadopoulos, T., & Sylaiou, S. (2021). Mixed reality: A reconsideration based on mixed objects and geospatial modalities. *Applied Sciences*, 11(5), 2417. https://doi.org/10.3390/app11052417.

Fast.com, Internet Speed Test. https://fast.com/ja/. Accessed on: October 9, 2021.

Fukuda, T., Yokoi, K., Yabuki, N., & Motamedi, A. (2019). An indoor thermal environment design system for renovation us-

ing augmented reality. *Journal of Computational Design and Engineering*, 6(2), 179–188. https://doi.org/10.1016/j.jcde.2018.05.007.

Gallacher, D. (2017). Drone applications for environmental management in urban spaces: A review. *International Journal of Sustainable Land Use and Urban Planning*, 3(4), 1–14. https://doi.org/10.24102/IJSLUP.V3I4.738.

Gimeno, J., Casas, S., Portalés, C., & Fernández, M. (2018). Addressing the occlusion problem in augmented reality environments with phantom hollow objects. In *IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (pp. 21–24). https://doi.org/10.1109/ISMAR-Adjunct.2018.00024.

Goudarznia, T., Pietsch, M., & 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, 244–251. https://doi.org/10.14627/537629025.

Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication, White Paper* (pp. 1–7).

Ham, Y., & Kim, J. (2020). Participatory sensing and digital twin city: Updating virtual city models for enhanced risk-informed decision-making. *Journal of Management in Engineering*, 36(3), 1–12. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000748.

Haynes, P., Hehl-Lange, S., & Lange, E. (2018). Mobile augmented reality for flood visualization. *Environmental Modelling and Software*, 109, 380–389. https://doi.org/10.1016/j.envsoft.2018.05.012.

Herman, L., Juřík., V., Snopková, D., Chmelík, J., Ugwitz., P., Stachoň, Z., Šašinka, Č., & Řezník, T. (2021). A comparison of monoscopic and stereoscopic 3D visualizations: Effect on spatial planning in digital twins. *Remote Sensing*, 13(15), 2976. https://doi.org/10.3390/rs13152976.

Holynski, A., & Kopf, J. (2018). Fast depth densification for occlusion-aware augmented reality. *ACM Transactions on Graphics*, 37(6), 1–11. https://doi.org/10.1145/3272127.3275083.

Hong, I., Kuby, M., & Murray, A. T. (2018). A range-restricted recharging station coverage model for drone delivery service planning. *Transportation Research Part C: Emerging Technologies*, 90, 198–212. https://doi.org/10.1016/j.trc.2018.02.017.

Jabbar, A., Farrawell, L., Fountain, J., & Chalup, S. K. (2017). Training deep neural networks for detecting drinking glasses using synthetic images. In *International Conference on Neural Information Processing* (pp. 354–363). https://doi.org/10.1007/978-3-319-70096-0_37.

Kasperi, J., Edwardsson, M. P., & Romero, M. (2017). Occlusion in outdoor augmented reality using geospatial building data. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology. Association for Computing Machinery* (Vol. 30, pp. 1–10). https://doi.org/10.1145/3139131.3139159.

Kido, D., Fukuda, T., & Yabuki, N. (2021). Assessing future landscapes using enhanced mixed reality with semantic segmentation by deep learning. *Advanced Engineering Informatics*, 48, 101281. https://doi.org/10.1016/j.aei.2021.101281.

Kikuchi, N., Fukuda, T., & Yabuki, N. (2021). Landscape visualization by integrating augmented reality and drones with occlusion handling to link real and virtual worlds – Towards city digital twin realization. In *Proceedings of the 39th eCAADe Conference* (Vol. 2, pp. 521–528). http://papers.cumincad.org/cgi-bin/works/paper/ecaade2021_036. Accessed on: October 9, 2021.

Kim, J. I., Li, S., Chen, X., Keung, C., Suh, M., & Kim, T. W. (2021). Evaluation framework for BIM-based VR applications in design phase. *Journal of Computational Design and Engineering*, 8(3), 910–922. https://doi.org/10.1093/jcde/qwab022.

Koch, V., Ritterbusch, S., Kopmann, A., Müller, M., Habel, T., & von Both, P. (2011). Flying augmented reality: Supporting planning and simulation analysis by combining mixed reality methods using multicopter and pattern recognition. In *Proceedings of the 29th eCAADe Conference* (Vol. 1, pp. 843–849). http://cumincad.scix.net/cgi-bin/works/Show?_id=ecaade2011_064. Accessed on: October 9, 2021.

Li, W., Han, Y., Liu, Y., Zhu, C., Ren, Y., Wang, Y., & Chen, G. (2018). Real-Time location-based rendering of urban underground pipelines. *ISPRS International Journal of Geo-Information*, 7(1), 1–17. https://doi.org/10.3390/ijgi7010032.

Liu, D., Xia, X., Chen, J., & Li, S. (2021). Integrating building information model and augmented reality for drone-based building inspection. *Journal of Computing in Civil Engineering*, 35, 1–40. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000958.

Microsoft HoloLens. (2016). https://www.microsoft.com/en-us/hololens/. Accessed on: October 9, 2021.

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12), 1321–1329. https://search.ieice.org/bin/summary.php?id=e77-d_12_1321. Accessed on: October 16, 2021.

Mohammadi, N., & Taylor, J. (2019). Devising a game theoretic approach to enable smart city digital twin analytics. In *Hawaii International Conference on System Sciences (HICSS)* (pp. 1995–2002). https://doi.org/10.24251/HICSS.2019.241.

Motlagh, N. H., Bagaa, M., & Taleb, T. (2017). UAV-Based IoT platform: A crowd surveillance use case. *IEEE Communications Magazine*, 55(2), 128–134. https://doi.org/10.1109/MCOM.2017.1600587CM.

Offenhuber, D., & Seitinger, S. (2014). Over the rainbow: Information design for low-resolution urban displays. In *Proceedings of the 2nd Media Architecture Biennale Conference: World Cities (MAB '14). Association for Computing Machinery* (pp. 40–47). https://doi.org/10.1145/2682884.2682886.

Open Geospatial Consortium. (2021). OGC city geography markup language (CityGML) 3.0 conceptual model users guide, 20-066, Version 3.0. https://docs.ogc.org/guides/20-066.html. Accessed on: October 9, 2021.

PLATEAU. (2020). Ministry of Land, Infrastructure, Transport and Tourism. https://www.mlit.go.jp/plateau/. Accessed on: October 9, 2021.

Roxas, M., Hori, T., Fukiage, T., Okamoto, Y., & Oishi, T. (2018). Occlusion handling using semantic segmentation and visibility-based rendering for mixed reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology (VRST 2018)* (Vol. 20, pp. 1–8). https://doi.org/10.1145/3281505.3281546.

Ruohomäki, T., Airaksinen, E., Huuska, P., Kesäniemi, O., Martikka, M., & Suomisto, J. (2018). Smart city platform enabling digital twin. In *International Conference on Intelligent Systems (IS)* (pp. 155–161). https://doi.org/10.1109/IS.2018.8710517.

Schrotter, G., & Hürzeler, C. (2020). The digital twin of the city of Zurich for urban planning. *Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88, 99–112. https://doi.org/10.1007/s41064-020-00092-2.

Shahat, E., Hyun, C. T., & Yeom, C. (2021). City digital twin potentials: A review and research agenda. *Sustainability*, 13, 1–20. https://doi.org/10.3390/su13063386.

Tian, Y., Long, Y., Xia, D., Yao, H., & Zhang, J. (2015). Handling occlusions in augmented reality based on 3D reconstruction method. *Neurocomputing*, 156, 96–104. https://doi.org/10.1016/j.neucom.2014.12.081.

Tomkins, A., & Lange, E. (2019). Interactive landscape design and flood visualisation in augmented reality. *Multimodal Technologies and Interaction*. 3(2), 1–13. https://doi.org/10.3390/mti3020043.

Tomkins, A., & Lange, E. (2021). Where the wild things will be: Adaptive visualisation with spatial computing. *Journal of Digital Landscape Architecture*, 6, 140–147. https://doi.org/10.14627/537705011.

Unal, M., Bostanci, E., Sertalp, E., Guzel, M. S., & Kanwal, N. (2018). Geo-location based augmented reality application for cultural heritage using drones. In *Proceedings of the 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)* (pp. 1–4). https://doi.org/10.1109/ISMSIT.2018.8567073.

Unal, M., Bostanci, E., & Sertalp, E. (2020). Distant augmented reality: Bringing a new dimension to user experience using drones. *Digital Applications in Archaeology and Cultural Heritage*, 17, 1–12. https://doi.org/10.1016/j.daach.2020.e00140.

Valentin, J. et al. (2018). Depth from motion for smartphone AR. *ACM Transactions on Graphics*. 37(6), 1–19. https://doi.org/10.1145/3272127.3275041.

Virtual Singapore. (2017). Singapore Government. https://www.nrf.gov.sg/programmes/virtual-singapore. Accessed on: October 9, 2021.

Wen, M., & Kang, C. (2014). Augmented reality and unmanned aerial vehicle assist in construction management. In *International Conference on Computing in Civil and Building Engineering* (pp. 1570–1577). https://doi.org/10.1061/9780784413616.195.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, 179, 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021.

White, G., Zink, A., Codecá, L., & Clarke, S. (2021). A digital twin smart city for citizen feedback. *Cities*, 110, 1–11. https://doi.org/10.1016/j.cities.2020.103064.

Wiberg, A. H., Løvhaug, S., Mathisen, M., Tschoerner, B., Resch, E., Erdt, M., & Prasolova-Førland, E. (2019). Visualisation of KPIs in zero emission neighbourhoods for improved stakeholder participation using virtual reality. *IOP Conference Series: Earth and Environmental Science*, 323, 012074. https://doi.org/10.1088/1755-1315/323/1/012074.

Yan, L., Fukuda, T., & Yabuki, N. (2019). Integrating UAV development technology with augmented reality toward landscape tele-simulation. In *Proceedings of the 24th CAADRIA Conference* (Vol. 1, pp. 423–432). http://papers.cumincad.org/cgi-bin/works/paper/caadria2019_411. Accessed on: October 9, 2021.

Zollmann, S., Langlotz, T., Grasset, R., Lo, W. H., Mori, S., & Regenbrecht, H. (2020). Visualization techniques in augmented reality: A taxonomy, methods and patterns. *IEEE Transactions on Visualization and Computer Graphics*, 27(9), 3808–3825. https://doi.org/10.1109/TVCG.2020.2986247.