



# Article SkyroadAR: An Augmented Reality System for UAVs Low-Altitude Public Air Route Visualization

Junming Tan <sup>1,2,3</sup>, Huping Ye <sup>1,3,4</sup>, Chenchen Xu <sup>1,3,4</sup>, Hongbo He <sup>1,2,3</sup> and Xiaohan Liao <sup>1,2,3,4,\*</sup>

- State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; tanjm.19b@igsnrr.ac.cn (J.T.); yehp@igsnrr.ac.cn (H.Y.); xucc.14s@igsnrr.ac.cn (C.X.); hehb.20b@igsnrr.ac.cn (H.H.)
- <sup>2</sup> College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
- <sup>3</sup> Key Laboratory of Low Altitude Geographic Information and Air Route, Civil Aviation Administration of China, Beijing 100101, China
- <sup>4</sup> The Research Center for UAV Applications and Regulation, Chinese Academy of Sciences, Beijing 100101, China
- \* Correspondence: liaoxh@igsnrr.ac.cn; Tel.:+86-10-6488-9502

Abstract: Augmented Reality (AR) technology visualizes virtual objects in the real environment, offering users an immersive experience that enhances their spatial perception of virtual objects. This makes AR an important tool for visualization in engineering, education, and gaming. The Unmanned Aerial Vehicles' (UAVs') low-altitude public air route (Skyroad) is a forward-looking virtual transportation infrastructure flying over complex terrain, presenting challenges for user perception due to its invisibility. In order to achieve a 3D and intuitive visualization of Skyroad, this paper proposes an AR visualization framework based on a physical sandbox. The framework consists of four processes: reconstructing and 3D-printing a sandbox model, producing virtual scenes for UAVs Skyroad, implementing a markerless registration and tracking method, and displaying Skyroad scenes on the sandbox with GPU-based occlusion handling. With the support of the framework, a mobile application called SkyroadAR was developed. System performance tests and user questionnaires were conducted on SkyroadAR; the results showed that our approachs to tracking and occlusion provided an efficient and stable AR effect for Skyroad. This intuitive visualization is recognized by both professional and non-professional users.

**Keywords:** augmented reality; the UAVs' low-altitude public air route; sandbox; markerless registration and tracking; GPU-based virtual-real occlusion handling; SkyroadAR

# 1. Introduction

Augmented Reality (AR) is a "human-centered" visualization technology that differs from Virtual Reality (VR), which creates independent immersive 3D virtual environments that are detached from the physical world. AR can overlay virtual objects in real-time onto the displayed image of the real environment through mobile device screens or headmounted displays, and provide interactive experiences that combine both virtual and real elements according to user behavior [1]. This allows users to enhance their spatial thinking and orientation capabilities for virtual objects without leaving the physical world. AR can be applied in various visualization scenarios, ranging from indoor to outdoor spaces. Examples include a virtual teaching sandbox [2], geographic process simulation [3], indoor and outdoor AR navigation [4,5], and underground pipeline inspection [6].

In UAV operation and aviation-related fields, AR technology has been widely used in assisted flight and simulation. For example, the augmented reality head-up display (AR HUD) has been used to project some important parameters to the screen in front of the pilot, such as altitude, speed, and navigation information [7], or the panel of the UAV's image



Citation: Tan, J.; Ye, H.; Xu, C.; He, H.; Liao, X. SkyroadAR: An Augmented Reality System for UAVs Low-Altitude Public Air Route Visualization. *Drones* 2023, 7, 587. https://doi.org/10.3390/ drones7090587

Academic Editor: Vaios Lappas

Received: 7 August 2023 Revised: 7 September 2023 Accepted: 15 September 2023 Published: 19 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). transmission system [8]. Flightradar24, a real-time flight information service provider, uses mobile AR technology to display information such as the departure, destination, flight number, flight altitude, and airspeed of the flight in the corresponding sky orientation [9]. In the indoor environment, AR can be used to superimpose the scaled-down virtual scene into the real environment and observe the overall situation of the virtual scene from a global perspective. For example, C Liu et al. presented the reconstructed UAV flight environment for 3D visualization, and specified the moving position of the UAV target through gestures. The position of the virtual UAV is displayed in the flight environment expressed by AR [10].

In recent years, civilian UAVs have experienced an explosive growth in the applications market. The large number of low-altitude UAVs flying in disorder poses risks to public safety. However, the current means of UAV supervision and air traffic management cannot adapt to the millions of drones in the future sky. Hence, UAVs' low-altitude public air route (also known as UAVs' Skyroad) is proposed as a forward-looking low-altitude traffic management solution [11]. In our study, low altitude is defined as 300 m AGL.

Different from the traditional road traffic, the UAVs' Skyroad is a digital traffic infrastructure in mid-air above the ground surface. The intuitive visualization and interaction involved in Skyroad plays an important role in the effectiveness of decision-making from planning to operation. At present, the visualization of the air route is mainly based on the 2D or 3D GIS platform developed by the general graphics library. For example, He et al. [12] used ArcGIS software to visualize the spatial distribution of the UAV logistics air route network in the 2D map. For users to understand the 2D visual expression, a strong spatial abstraction ability is often required, which makes it unsuitable for non-professional users. Zhang et al. [13] used the 3D WebGIS platform (i.e., Cesium) to visualize collision-free path planning for the UAVs. Although a computer-aided 3D rendering can display the multi-dimensional information of Skyroad, because it creates a virtual scene through 3D projection onto a 2D screen, users cannot intuitively perceive the spatial relationship of Skyroad by using mouse clicks to change the perspective, which increases the information processing burden.

The use of GIS in aviation has traditionally focused on combining dynamic and static map visualization with basic spatial analysis. Skyroad, a forward-looking transportation infrastructure for large-scale UAVs operations, is in the early research stage. To meet new demands, AR + GIS visualization is crucial for assisting decision-making throughout Skyroad's lifecycle. At present, there is still a lack of research on the AR visualization of Skyroads, and there are some shortcomings in applying traditional AR visualization methods to Skyroad. For example, the use of AR in large outdoor scenes is not suitable for displaying Skyroad from a global perspective, while indoor AR is mostly displayed on a miniature model or a card with quick response (QR) code technology using marker-based tracking, which is easily constrained by the environment and the location of markers. In addition, virtual objects are always rendered first in the AR scene, while less consideration is given to virtual-real occlusions in the environment, which reduces the layered sense of AR 3D expression.

In order to present the UAVs' Skyroad in the most macroscopic and realistic way, this paper proposes an AR visualization framework for Skyroad based on a physical sandbox model. We developed an innovative model-based marker-less tracking and virtual-real occlusion handling method, and built a prototype system (UAVs' Skyroad AR Visualization System, SkyroadAR) on mobile devices, and verified its effectiveness and usability through a comparative testing of system performance and user questionnaires.

## 2. Related Work

# 2.1. AR Sandbox Visualization

An important application of indoor AR is to display virtual elements of different scales on real objects or planes to understand and perceive virtual elements in an all-round way. It is widely used in interior design, games, education, etc. The well-known furniture company IKEA has launched an AR application, IKEA Place, where users can place virtual furniture in their furniture library at home to visually view the effect, guiding users to purchase suitable furniture [14]. Neven A et al. [15] developed the ARSC card, which presents the course content in virtual 3D form on the card to help students carry out visual practice and master knowledge in a new and effective, interactive way. Bobrich displays a 3D DEM terrain model on a paper map marked with ARToolKit, and users can interact with the map in their own way [16].

The AR sandbox developed in recent years has replaced the traditional AR card or AR, and become a powerful tool in the fields of education and geographic design. At first, it consisted of a table as the carrier of a virtual sand layer to display virtual terrain [2], and then, based on the terrain landscape formed by real sand, it displayed virtual contour information [17]. Three-dimensional printing technology has greatly reduced the cost of making 3D models, and is becoming an important way to build AR sandboxes. Gong Jianhua et al. proposed to enhance the process visualization method in the geographical environment, print out the geographical environment models such as buildings and mountain models, and superimpose the virtual fire emergency evacuation process on the building model [18], as well as providing a virtual simulation of flood diffusion effect in the mountain model [3]. Tuzun et al. [19] superimposed 3D virtual architectural models on Lego models to elaborate future scenarios. The AR sandbox provides a novel way to express and study different geographical developments and phenomena in real time. More and more evidence is showing that AR sandboxes are beneficial to improve users' spatial thinking [20]. However, most of the visualization methods mentioned above aim to miniaturize the physical environment objects into a small-scale sandbox or other carriers. The level of detail in the displayed virtual information is limited by the size of the sandbox, and it is usually impossible to observe from multiple angles and distances.

## 2.2. Registration and Tracking for AR

The main function of AR registration and tracking is to track the changes in the target position in the real scene to obtain the sensor pose relative to the target in real time and determine the spatial relationship between the virtual object and the target. Tracking and registration technology has always been a research hotspot and a difficulty in AR [21].

Sensor-based registration and tracking aims to capture and display the position and posture of the device in real time through the signal received by the hardware sensor. For example, outdoor AR applications usually use the global positioning system of mobile devices to obtain location information, and combine the inertial navigation system to obtain attitude [22]. The location of the mobile device is obtained indoors using the Wi-Fi number intensity ranging model, and the orientation sensor of the device is fused to obtain the attitude, thereby realizing AR tracking and registration [23]. Although the tracking and registration technology of sensors is simple and fast, it relies too much on hardware devices and is susceptible to measurement noise caused by external factors.

Image-based tracking and registration are currently the mainstream in AR, which uses advanced computer vision algorithms to extract artificial features or natural features in real scene images, and calculates the current position and orientation of the sensor in real time according to the features in the image [24]. A common marker-based tracking method is to estimate the 6DOF pose of the camera relative to the marker by placing an easily identifiable marker in the environment and detecting the marker in real time through the camera image. This method is widely used in indoor AR applications [25,26]. The pose estimation obtained by the marker-based tracking is usually accurate, but it requires the marker to always appear in the user's field of view and the marker must maintain full visibility. On the other hand, marker-less tracking mainly tracks the camera pose by obtaining natural features such as corners, lines, and planes in the real world to find the correspondence between the 3D world and the 2D image coordinates. The flexibility of the method makes it compatible with indoor and outdoor environment [27]. These features can be pre-learned templates [28], or real-time reconstruction and perception acquisition of the environment based on simultaneous localization and mapping (SLAM) and deep

learning [29]. The marker-less tracking can solve the problem of needing to place markers in advance, but has high requirements for equipment computing power and algorithm efficiency, especially when using the online reconstruction learning method, which can surpass the limitations of environmental scenarios in any unfamiliar environment, but current mobile devices cannot meet such a high computational cost.

# 2.3. Occlusion Handling in AR

As another key AR technology, the virtual-real occlusion handling affects the authenticity of AR visualization. Occlusion handling aims to judge the sequence between the virtual object and the target object on the basis of tracking and registration. The key is to obtain the depth of the target object in the scene and compare it with the depth of the virtual object.

Occlusion handling based on image segmentation mainly aims to separate the foreground and background of the obstruction in the image through contour extraction, and track the obstruction area in subsequent frames. Cordes et al. [30] proposed an image segmentation method, which takes the foreground pixels marked by the user as input, and tracks these foreground regions to occlude virtual objects in the synthetic scene. Tian et al. used the interactive segmentation method [31] and the automatic disparity map extraction method [32] to obtain the contour of the real object, and redrew the pixels of the foreground area of the obstruction on the synthetic image to show the correct occlusion relationship. Although segmentation can provide real-time occlusion processing, it can only treat the whole obstruction as a foreground area, the effect of mutual occlusion is not good, and when the outline of the obstruction changes significantly, the method will fail.

Depth-based occlusion handling refers to obtaining the depth of the real scene, either offline or online, and comparing it with the depth of the virtual object to deal with occlusion. Off-line 3D reconstruction based on multi-view stereo vision MVS is the main method of generating scene depth maps. For example, Kim et al. [33] proposed an occluded object edge depth estimation algorithm based on hierarchical disparity estimation. The edges of the disparity estimation are regularized and converted into depth maps. Methods based on offline 3D reconstruction can effectively deal with mutual forms of occlusion, but are not suitable for dynamically changing scenes. Lzadi [34] and Newcombe et al. [35] proposed the KinectFusion framework, using Microsoft Kinect with a 3D structured light sensor to obtain depth maps in real time. The active visual sensors that provide depth measurements have high requirements for algorithms and hardware equipment, and cannot be run on most mobile devices equipped with monocular cameras. In recent years, with the development of computer vision and AI technology, the real-time 3D reconstruction and depth estimation technology used in monocular SLAM [36,37] has become more common in AR, but due to algorithm and hardware limitations, it is currently only applicable to depth estimations in indoor scenes and cannot run on mobile devices.

## 3. Materials and Methods

## 3.1. SkyroadAR Framework

The planned Skyroad shuttles in the complex low-altitude geographical environment, and the user can intuitively express the overall operation concept of the air routes and their relationship to place attachment, assist stakeholders in route planning, display the planned global air routes in the physical sandbox reconstructed from a real geography environment, and meet the simplicity, efficiency, accuracy and interactivity requirements of AR visualization. This paper proposes the overall technical framework for the development of the SkyroadAR visualization system. This is shown in Figure 1, it is divided into four main parts:

(1) The basic geographic information data of the study area are used to obtain the geographical constraint elements that affect air route planning, such as buildings, roads, and water bodies, restore the geometric shape, texture and color of the elements in the scene

in proportion through 3D-printing technology, and establish a physical sandbox model of the study area. As shown in the red block in Figure 1; see Section 3.2 for specific methods.

(2) Based on geographic information data, multi-level air-route planning (backbone, trunk, branch, terminal) [38] and the iterative construction of an air route network in the urbanization region [39] are executed. The generated waypoint data are imported into the 3D animation software in a specific file exchange format to create a multi-scene and multi-level virtual animation scene for UAVs' low-altitude public air routes. This is shown in the blue block in Figure 1; see Section 3.3 for specific methods.

(3) The 3D virtual model of the sandbox is generated by multi-view 3D reconstruction. Using a model-based marker-less tracking and registration method, in the offline stage, the multi-view reference images of the virtual model are collected, and the camera position and attitude information are recorded in the extensible markup language (xml). In the online stage, the key frame image collected in real time is template-matched with the gradient information of the reference image in the offline stage, and the roughly estimated pose information is obtained from the xml. By matching the key frame image with the local natural feature descriptor of the reference image, the relationship between the 2D and 3D corresponding points is obtained to accurately calculate the camera pose. This is shown in the green block in Figure 1; see Section 3.4 for specific methods.

(4) Based on the depth information of the 3D virtual sandbox model, the template buffer in the GPU real-time rendering pipeline is used to make a transparent mask template, and the fusion effect of virtual-real occlusion is realized through template testing and depth testing. This is shown in the yellow block in Figure 1; see Section 3.5 for specific methods.



**Figure 1.** Methodological framework of SkyroadAR. The red block represents Section 3.2 physical sandbox construction, the blue block represents Section 3.3 AR Scene Production for UAVs' Skyroad, the green block represents Section 3.4 Model-Based Markless Registration and Tracking and the yellow block represents Section 3.5 GPU-Based Virtual-Real Occlusion Handling.

## 3.2. Physical Sandbox Construction

Representing the environment in the form of miniature entities can display abstract geographical cognition in the form of a 3D model and enhance the multi-scale spatial imagination. With the development of 3D printing technology, the use of GIS and CAD software and high-precision inkjet technology to quickly generate a realistic sandbox and reduce the cost of manual sandbox production has been greatly welcomed. This paper takes an area of about 25 square kilometers in Shekou, Nanshan District, Shenzhen, Guangdong Province, China, as the research area, and builds a physical sandbox with a size of  $3.6 \text{ m} \times 3 \text{ m}$ .

## 3.2.1. GIS Data Preparation and Processing

The low-altitude geographical environment of UAVs' flight is complex and changeable, and the factors affecting route safety in the geographical environment are defined as routesensitive constraints, such as terrain, building, road, water body, vegetation, and power lines (or poles). Firstly, determine the geographic boundaries of the planned UAV Skyroad. Collect basic geographic information data such as a digital elevation model (DEM) from the open data sharing platform, and obtain building boundaries, roads, vegetation, water



bodies and other information based on intelligent interpretation technology for high-resolution remote sensing images [40], as shown in Figure 2.

**Figure 2.** Basic GIS data of the study area: (**a**) image; (**b**) DEM (Different colors represent different elevations, refer to the ribbon label); (**c**) building; (**d**) road.

# 3.2.2. CAD Data Generation

Except for the DEM data stored in the raster image format, mapped by the 2D grayscale symbology, most of the basic GIS data are vector boundary information stored in the shapefile (.shp) file or json file format. A few, such as oblique photographic models and laser point cloud data, are directly stored in 3D data format. Before 3D printing, it is necessary to unify the format into the stereolithography file (STL).

Use the DEMto3D plug-in in QGIS software to project the DEM raster data of the research area into the general transverse Mercator coordinate system, determine the model scale according to the output printing size set by the user and adjust the elevation exaggeration coefficient to meet the horizontal and vertical visual balance. The STL file is generated by calculating the grid vertex coordinates.

For the vector boundary data, leverage the Supermap software to extrude the 3D model according to the height attribute, and export the model in STL format. The specific

process of converting GIS data to CAD data is shown in Figure 3, and the generated CAD data of the study area are shown in Figure 4.



Figure 3. The flowchart of GIS data to CAD data.



**Figure 4.** The CAD data: (a) 2D CAD: The red lines represent the road, the black lines represent the buildings, and the other color represent the extent of other elements; (b) 3D CAD: The green represent 3D building and the pink represent the lower buildings.

# 3.2.3. Sandbox Printing

The 3D printer prints a section with a certain micro-thickness and a specific shape each time, and bonds it layer by layer to form a 3D model. Since the STL file format only stores the discrete triangular patch information on the surface of the CAD model, its simple data structure is suitable for sliced 3D printers, which use specific materials (such as foam boards) to print various geographically constrained elements onto the white model, on which the texture and color extracted from remote sensing images are painted, and sound, light, and electrical systems are added to make a vivid physical sandbox, as shown in Figure 5.



Figure 5. The sandbox 3D printing: (a) the white model; (b) the colored model.

# 3.3. AR Scene Production for UAVs' Skyroad

In order to express the concept of low-altitude public routes on the physical sandbox, it is necessary to plan air routes based on GIS data and create a virtual air route scene by simulating the flight of the UAVs along the Skyroad. By integrating the virtual scene with the physical sandbox, the Skyroad planning and application effect of UAVs can be realistically displayed.

# 3.3.1. Low-Altitude Flight Environmental Modeling for UAVs

For the complex and changeable low-altitude geographical environment, Skyroad planning needs to comprehensively consider the specific geographical elements that are to be avoided (such as buildings and power lines) and important flight conditions (such as noise an privacy) so as to set different altitude levels and safety intervals for the route. Intelligent interpretation technology is used to extract the information of various sensitive constraints, combined with risk assessment to determine the boundaries of geo-fences, while parametric modeling methods are used to establish a 3D representation model of the low-altitude flight environment. The geographical fence is divided into four levels, the airworthiness zone, the buffer zone, the warning zone, and the no-fly zone, and the level of welcoming UAVs in the area is gradually reduced, as shown in Figure 6.



**Figure 6.** Modeling of UAVs' low-altitude flight environment: (**a**) constraint element extraction; (**b**) 3D geofencing.

## 3.3.2. Multi-Level and Regional Air Route Network Planning

According to the IEEE standard for planning UAVs' low-altitude public air routes based on massive geographic information, the construction of national-scale UAV Skyroad is divided into four levels according to their operating capabilities (backbone routes, trunk routes, branch routes, and terminal (regional) routes [11]), which can be obtained by multiple geographic constraints and the path search algorithm [38]. At the regional scale, the initial regional airway network is generated by vertically lifting the ground road network, and the airway network is constructed and iteratively improved in five steps under the condition of utilizing or avoiding geographical constraint elements [39], as shown in Figure 7.



**Figure 7.** A multi-level Skyroad network indicated by yellow line: (**a**) backbone air-routes (country level); (**b**) trunk air-routes (province level); (**c**) branch level air-routes (city level); (**d**) terminal level air-routes (region level) [39].

# 3.3.3. Animation of Virtual Skyroad Scene

In order to dynamically express the Skyroad scene of UAVs' operation, a multi-level route network is designed according to the application requirements, and used to verify altitude conversion and approach rules [41]. For normal operation scenarios, such as the urban loading of passengers or goods, and logistics distribution, a general air route with two-way operation lanes is set. For special scenarios, such as sightseeing and tourism, power (or road) inspection, survey and mapping, and cross-sea logistics transportation, through setting up both dedicated and public air routes to achieve staggered peak operation, information technology can be used to share the Skyroad's right of way, and make rational use of airspace resources. As shown in Figure 8, the scale-transformed virtual sandbox model can be used to establish the UAVs flight animation in the scene, and the 3D models of point of interest (POI) and geo-fence can be added to vividly display a virtual Skyroad scene.



**Figure 8.** Modeling of virtual Skyroad scene: (**a**) virtual sandbox scale transformation; (**b**) POI modeling: the red texts indicate POIs information; (**c**) Skyroad modeling: Red indicates low-altitude geography, yellow lines are higher altitude air routes, and blue lines are lower altitude air routes; (**d**) geofence modeling: Red geo-fences are areas where drone flights are restricted or prohibited.

## 3.4. Model-Based Markless Registration and Tracking

Since the Skyroad scene is represented by a 3D Cartesian coordinate system that has nothing to do with the real world, to superimpose the virtual Skyroad on the physical sandbox, we need to align the size of the virtual scene and the real scene, and obtain the spatial relationship between the camera and the target in real time through a change in the target position in the real scene. This relationship is represented by Formula (1).

$$\lambda m' = P * M' \tag{1}$$

 $\lambda$  denote the scale factor, and a 3D point M' is corresponds to the homogeneous coordinate of a 2D point m' on the image through the pose matrix P, which is equal to k \* [R|t], where R is  $3 \times 3$  rotation matrix and t denote translation vector ( $3 \times 1$  column matrix); together, they constitute the camera's extrinsic parameter matrix, while k denote the intrinsic parameters matrix that occurs independent of motion, as shown in Formula (2).

$$k = \begin{bmatrix} \alpha_x & s & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$
(2)

Here,  $\alpha_x$  and  $\alpha_y$  depend on the focal length and the size of the photosensitive element,  $(u_0, v_0)$  is the principal point of the camera, and the camera tilt factor *s* approximately equals zero.

The intrinsic parameter is stable after camera calibration, so recovering the camera's extrinsic parameter accurately is the primary concern of AR registration and tracking. In this paper, a model-based tracking method is adopted. The physical sandbox is reconstructed to generate a virtual 3D model. In the offline stage, reference images of various viewpoints with different angles and distance are collected by simulating the real camera movement and pose parameter information of the virtual camera, and edges and natural feature points in the reference image are also recorded. In the online stage, the extrinsic parameter matrix of the camera is recovered by template-matching and feature-matching on the real-time 2D image and the reference image.

## 3.4.1. 3D Reconstruction of Physical Sandbox

The multi-view images are collected by shooting around the sandbox with the camera, and the feature points of each image are extracted using the scale-invariant feature transformation (SIFT) operator. Sparse point-line feature-matching is employed to calculate the internal and external orientation elements, and they are optimized using bundle adjustment. The 3D point cloud of the target is acquired based on dense matching, and finally a 3D model is generated through triangulation and texture mapping. We add real scale constraints in the reconstruction process so that the ratio of the reconstructed virtual sandbox to the real physical sandbox is 1:1, which is used as the scale benchmark of the virtual Skyroad scene to achieve virtual-real scale matching. The sandbox 3D reconstruction is shown in Figure 9.



**Figure 9.** 3D reconstruction of physical sandbox: (**a**) sparse point cloud generation by orientation estimation; (**b**) reconstructed 3D model of virtual sandbox.

# 3.4.2. Offline Training of Gradient-Based Contour Directions and ORB Features

In the offline training phase, by simulating the real camera motion parameters, the 3D world coordinates of the virtual sandbox model are projected to the image plane to generate reference views under different view angles, and the gradient directions of the contours in the reference views and the oriented FAST and rotated BRIEF (ORB) features are extracted, saved and encoded.

Assume that the virtual sandbox is located at the center of the spherical coordinate system formed by the camera orbit motion, and the optical axis of the camera always passes through the center. By specifying the interval of longitude  $\lambda$ , latitude  $\varphi$  and distance *d* in the spherical parameters, the viewing range covered by the camera under different view angles and different distances is limited to a certain spherical hexahedron, as shown in Figure 10.



Figure 10. The spherical hexahedron of offline sampling.

We draw on the approach of hierarchical tree proposed by Wiedemann [42]; the views are roughly sampled at higher image pyramid levels to speed up recognition in the online stage. View sampling starts from the lowest image pyramid level, and calculates the similarity between views at adjacent camera positions by applying a contour similarity measurement (Formula (5)). The pair of views with the highest similarity is merged into a new view, and the similarity between its neighbors is calculated. This process is repeated until the similarity is below a certain threshold. As shown in Figure 11a, the views remaining after merging are stored at the lowest (original) hierarchical level. In order to derive the view of the next higher hierarchical level, the similarity constraint is relaxed by reducing the image resolution while continuing to merge according to the similarity measure, as shown in Figure 11b–d. The subviews are views of the lower pyramid level that have been merged to obtain views at the current pyramid level or that cannot be merged. Each view stores references to all subviews via a hierarchical tree structure. This information is used during the online phase to query a given view at a higher pyramid level, and at the lower pyramid level to refine the match.



**Figure 11.** Views at various levels are stored in the hierarchical tree structure. (**a**) is the original view, while (**b**) is a subview created by merging (**a**) based on the similarity threshold. Similarly, (**c**,**d**) are subviews of the previous view.

The outline template for the view is extracted in the hierarchy tree. Denoising and preprocessing are completed by Gaussian fuzzy low-pass filtering, using the linear parallel multi-modal LINE-MOD template-matching method [43], while the contours of the sandbox are calculated by the Sobel operator, and the Phase function is used to calculate the three-channel (red, green and blue) gradient direction of each pixel of the contours. The maximum value is taken, as shown in Formula (3), which represents the gradient of image *I* at position *x*, and *ori* denotes radians.

$$I_{g}(x) = ori(\hat{C}(x)) \tag{3}$$

where

$$\hat{C}(x) = \underset{C \in \{R,G,B\}}{\operatorname{argmax}} \|\frac{\partial C}{\partial x}\|$$
(4)

In order to use the rich texture information of the sandbox to obtain accurate matching, the extraction of ORB feature points and the generation of feature descriptors are also performed in the offline stage, and the extracted feature points are back-projected onto the 3D model of the object using the pose of the virtual camera to determine the world coordinates of the corresponding 3D points. Finally, the pose, the contour template, the ORB feature points and feature descriptors, and the corresponding 3D world coordinates of each reference view are stored in the XML file. Figure 12 shows the contour template and extracted ORB feature points under a certain view.



**Figure 12.** Offline training of a certain view: (**a**) contour template, extracted based on LINE-MOD; (**b**) feature points (in red), extracted based on ORB.

## 3.4.3. Online Matching and Pose Optimization

In the online stage, by matching the similarities between the offline training template image and the current input image, the 3D pose of the physical sandbox relative to the camera is roughly determined, and the range of ORB feature points is narrowed down to accurately estimate the pose.

Recognition starts from the top level of the input image pyramid. All 2D contour templates at this level are searched using Formula (5) to compute the similarity measure between the input image and the contour template.

$$\varepsilon(I,T,c) = \sum_{r \in p} \max_{t \in R(c+r)} |\cos(ori(O,r) - ori(I,t))|$$
(5)

where  $\varepsilon(I, T, c)$  is the similarity between input image *I* and template *T*. ori(O, r) denotes the gradient direction at position *r* in the template image *O* expressed in radians; ori(I, t) is the radian representation of the gradient direction at position *t* in the input image *I*; *p* is the list of position *r*, and T = (O, p) is the outline template of the object.

$$t \in R(c+r) = [c+r-\frac{\tau}{2}, c+r+\frac{\tau}{2}][c+r-\frac{\tau}{2}, c+r+\frac{\tau}{2}]$$
(6)

where *R* represents the area with the offset *c* at position *r* as the center and  $\tau$  as the radius of the neighborhood.

The matching similarity measure at position p is computed in the input image using a sliding window approach, and the contour template parameters (position, rotation, scale)

that exceed the similarity threshold are stored in the matching candidate list. At the next lower hierarchical tree level, refinement is performed by computing the similarity measure between the 2D contour templates of all subviews and the current pyramid level input image, and the scope of template parameters in the candidate list is limited to the immediate neighbors of the parent match. This process is repeated until all matching candidates have been traced down to the lowest pyramid level. Through hierarchical trees and image pyramids, this approach can greatly speed up the matching.

After the key frame image that is most similar to the pose of input image is obtained through template-matching, a more accurate camera pose is estimated based on the method of ORB natural feature point-matching [44]. By inputting the 2D image and the corresponding 3D feature points in the key frame, the pose matrix [R|t] in Formula 1 of the current camera is accurately computed based on the PnP algorithm [45]. Utilizing the continuity between adjacent frames, the Lucas Kanade sparse optical flow algorithm [46] is used to predict the coordinates of the image feature points in the next frame during the tracking process to reduce the time spent repeatedly extracting and identifying features of the entire image and improve the tracking speed and real-time performance of the algorithm. The template-matching process will only be reactivated when the number of matching feature points is less than the set threshold, for example, the view angle or distance changes greatly.

## 3.5. GPU-Based Virtual-Real Occlusion Handling

After determining the spatial relationship between the camera and the sandbox, it is necessary to superimpose the created virtual Skyroad scene on the physical sandbox, and in order to render the virtual scene more realistically, it is necessary to obtain the correct occlusion relationship between them. The virtual sandbox model can be used to obtain the depth information of the physical sandbox in the scene and realize the virtual-real occlusion handling through the template test and depth test of the GPU graphics rendering pipeline. The specific process is shown in Figure 13.



Figure 13. Flow chart of virtual-real occlusion.

Firstly, the depth map of the reconstructed 3D sandbox model can be obtained according to the current pose of the camera, and the camera projection matrix can be used to transform the 3D points in the visible surface into the camera coordinate system. This can be set not to output any color (that is, transparent display), thereby generating a transparent mask, and its distance from the camera (that is, the depth value) can be stored.

Next, traversal is used to obtain all the pixels in the current view. If it is determined that the pixel is inside the outline of the sandbox, the pixel value is set to 1; otherwise, it is set to 0, so we obtain the occlusion template.

Then, a template test is performed on the pixel value of the virtual scene in the fragment shader, the pixel with a value of 1 is set to pass the template test according to the occlusion template, and the RGB color information of the pixel is stored in the color cache. Pixels that fail the stencil test are discarded.

Finally, a depth test is performed on the pixel that has passed the template test, and its depth value is compared with the one that already been stored. If the new pixel depth value is smaller than the original one, the new pixel value will replace the original pixel value. The RGB information of the pixel is stored in the color cache so as to draw the virtual Skyroad scene blocked by the transparent mask, as shown in Figure 14.



**Figure 14.** The effect of the transparent mask frm a certain view angle: (**a**) the virtual sandbox and Skyroad scene; (**b**) the virtual route after being blocked by the transparent mask.

## 4. Results

# 4.1. SkyroadAR Visualization Effect

Figure 15a is a schematic diagram of the user using SkyroadAR, and the UAVs' lowaltitude public air routes, which are superimposed on the physical sandbox reconstructed from the environment of Shekou in Nanshan District of Shenzhen City (Figure 15b). The blue and green corridors in Figure 15c are the trunk routes and branch routes of the urban area. The operating speed and traffic capacity designed for the trunk routes are higher than those of the branch routes, so the height and width of trunk routes are larger. AR expression is carried out for the geo-fence area composed of constraint elements. The yellow area is the warning area of Nanshan National Park, and the red area is the no-fly area of Shenzhen Bay Port.



**Figure 15.** The effect of SkyroadAR visualization: (**a**) user holds mobile device for multi-directional observation; (**b**) the physical sandbox of Shenzhen; (**c**) virtual Skyroad scene overlay.

Based on the planned Skyroad, a number of operating rules and application scenarios of Skyroad are dynamically displayed with simulated animation effects such as the altitude conversion between multi-level routes, meeting UAVs at intersections, cross-sea logistics distribution, and bridge inspections, as shown in Figure 16.



**Figure 16.** UAVs' Skyroad application scenarios and concept display of traffic rules, different colored lines represent different types of routes, and wide routes and narrow routes are connected by altitude conversion (red dot): (a) altitude conversion in terminal area; (b) intersection meeting rules; (c) cross-sea logistics route; (d) inspection route.

# 4.2. Tracking Testing

The method described in this paper can be compared with the model target module in the commercial software Vuforia. The same mobile device is used to record the time consumed by pose estimation of each frame through a camera video at different view angles, different distances, and different scales, as shown below in Figure 17. The tracking registration from the center of the sandbox (frame 0–120) is initialized, and moved smoothly to the right side of the sandbox before zooming the observation distance (frame 120–200), then quickly moving to the left side of the sandbox (frame 200–320), and using the hand for full occlusion (frame 320–360) and for partial occlusion (frame 360–400). The tracking time corresponding to this 400-frame video is shown in the upper part of Figure 17.

# 4.3. Occlusion Testing

By turning the virtual-real occlusion function in the local and global scene on or off, the AR effect and can be compared and the real-time frame rate recorded to evaluate the efficiency of occlusion, as shown in Figure 18. In the local scene (Figure 18b), since the green Skyroad is not occluded, it is always displayed in front of the building, resulting in incorrect spatial cognition. In Figure 18c, by using the reconstructed sandbox model to compare the depth value of the current position, the green Skyroad behind the building is occluded so as to obtain the correct AR effect.



**Figure 17.** A 400-frame video sequence records during the process of moving and zooming from different view angles and distances from the sandbox, and records the time consumed in the tracking and registration of each frame: (**a**) initialization; (**b**) smoothly moving to the right side of the sandbox; (**c**) quickly moving to the left side of the sandbox; (**d**) fully blocking; (**e**) partially blocking with the palm.



**Figure 18.** Red boxes and arrows indicate sample occlusion of local scenes. As designed in Figure 8c, the blue lines denote primary urban skyroad with higher altitudes and faster speeds, while the green lines represent secondary skyroad within urban communities with lower altitudes and slower speeds. (a) local view of the physical sandbox; (b) without occlusion handling; (c) after occlusion handling.

In the global scene, after the occlusion handling is turned on, the alternate occlusion effect is not affected. As shown in Figure 19c, after the building blocks the green Skyroad, it can still be blocked by the blue Skyroad passing in front of it. The frame rate before and after the occlusion function is maintained within the range of 35~45 frames per second (fps), that is, the smoothness of visualization is guaranteed while the occlusion is correctly handled.



**Figure 19.** Red boxes and arrows indicate sample occlusion of global scenes, the skyroad represented by blue lines and green lines are the same as those in Figure 18. As designed in Figure 8d, the yellow block indicates a flight restricted area and the red block indicates a flight prohibited area.. (**a**) global view of the sandbox; (**b**) without occlusion handling; (**c**) after occlusion handling.

# 4.4. User Questionnaire

The user experience of SkyroadAR is evaluated through a user questionnaire. The evaluation of the usability and usefulness of AR visualization depends on the subjective experience of users, so it is necessary to rationally design metrics for statistics and scientifically guide subsequent research and development. This paper designs a questionnaire based on the principles of the Technology Acceptance Model (TAM) [47]. The questionnaire is divided into eight scoring questions based on three dimensions: perceived usability, perceived ease of use, and intention to use. Table 1 provides a list of items used to measure each dimension, with 5-point Likert scales coded as 1—strongly disagree, 2—disagree, 3—neutral, 4—agree, 5—strongly agree.

Dimension	No	Question	
Perceived usefulness	Q1 Q2	It was intuitive to help you understand the concept of UAVs Skyroad SkyroadAR provided a fast and precise tracking	
	Q3	The occlusion between the virtual Skyroad and sandbox is accurate	
Perceived ease of use	Q4	Learning to use SkyroadAR would be easy for me	
	Q5	Interaction with SkyroadAR would be flexible	
	Q6	SkyroadAR can do what I want it to do	
Intention to use	Q7 Q8	I intent to use SkyroadAR in the near future I intent to check the availability of SkyroadAR in the near future	

**Table 1.** Questionnaire designed based on TAM.

The survey was sent out to 70 system users; their age group was 15–60 years old, and they were GIS professional graduate students, teaching staff or others. We converted the returned questionnaires into excel form, and calculated the average score of each question, as shown in Figure 20.





A descriptive analysis was performed on the questionnaire data using SPSS software. Since this is a 5-level scale, the theoretical average score is 3, which was used as the test value for a single sample t-test. Table 2 provides the test results.

Question	Т	SD	p
Q1	27.12	0.49	< 0.01
Q2	13.10	0.72	< 0.01
Q3	14.39	0.63	< 0.01
Q4	7.92	0.95	< 0.01
Q5	0.66	0.90	>0.05
Q6	-4.18	0.94	< 0.01
Q7	16.86	0.64	< 0.01
Q8	4.32	1.16	< 0.01

Table 2. A single sample t-test scale of the SkyroadAR user questionnaire.

T represents the t value of *t*-test. SD represents the standard deviation. *p* stands for the statistical *p*-value. DF denotes the degrees of freedom; N denotes the number of samples. DF = N - 1 = 69.

## 5. Discussion

## 5.1. The Efficiency of Tracking

As shown in Figure 17, the markerless tracking method employed by SkyroadAR (red line) demonstrates faster initialization during the first stage of tracking compared to the Vuforia model target module tracking method (green line). This is because Vuforia directly adopts the ORB feature point-matching, while SkyroadAR narrows down the search range of ORB feature points and reduces the time consumed in feature point-matching through hierarchical tree-based contour template-matching. After successful tracking, the camera moves smoothly to the right side of the sandbox. The sparse optical flow method is used for follow-up tracking in SkyroadAR, which is robust to distance and scale changes under stable lighting conditions. Vuforia also employs a similar processing method, making the overall time-consuming trend of the second stage similar.

After swiftly moving to the left side of the sandbox, the algorithm encounters a loss of tracking due to significant camera movement, resulting in a scarcity of feature points. SkyroadAR reverts to template matching for registration. Because of the hierarchical tree structure we used, the search was only performed in the vicinity of the subview, resulting in lower time consumption compared to the first stage. However, this was slightly higher than the time consumption of Vuforia since Vuforia integrates the attitude sensor data of mobile devices for tracking and registration. However, due to the cumulative error of the attitude sensors, the accuracy of tracking declined, while the virtual and real fusion effect of SkyroadAR remained stable.

When using the palm for full occlusion, feature points are completely lost, leading to an increase in tracking time. However, when the palm is partially removed, SkyroadAR utilizes the key frame before occlusion to narrow the search range of ORB feature points, resulting in faster registration compared to Vuforia.

## 5.2. The Effect of Occlusion Handling

According to the questionnaire results, users gave an average score of 4.09 for the hypothetical question Q3 on the occlusion effect. This indicates that the occlusion handling in this paper reduces the false spatial cognition of users and enhances scene realism. The occlusion effectively utilizes prior depth information of the reconstructed model and GPU-based shader rendering, ensuring correct and fast occlusion without compromising user experience. However, this method has limitations. The occlusion effect relies on the 3D model's fineness and tracking accuracy. Due to pre-modeling accuracy, occlusion may not be ideal in local fine structure areas, resulting in fragmented and jagged occlusion. Additionally, pre-modeling limits AR flexibility and cannot handle dynamic occlusion.

## 5.3. System Usability and Usefulness Evaluation

The user experience questionnaire results show mean scores for seven questions that are higher than the theoretical mean, with narrow gaps around the average value (standard deviations ranging from 0.49 to 1.16). The single-sample t-test indicates a significant difference between the average score of Q1 (T = 27.12, p < 0.05) and the theoretical average score, demonstrating that most users agree with the assumption that SkyroadAR is an intuitive and effective visualization system. Q2 (T = 13.1, p < 0.05) and Q3 (T = 14.39, p < 0.05) indicate that the acceptance of tracking and occlusion handling effects surpasses the average. SkyroadAR excels in perceived usability.

In terms of perceived ease of use, Q5 (T = 0.66, p > 0.05) suggests that users perceive the ease of interaction to be slightly higher than the theoretical average score, with minimal difference. However, Q6 (T = -4.18, p < 0.05) indicates that the current software interaction has certain limitations, differing significantly from the assumption, and users desire more interactive functions. This is because the primary objective of SkyroadAR is to visualize the operating concept of UAVs' low-altitude public air route. Due to the substantial difference between mobile device operation and traditional keyboard-mouse interaction, desired interactive functions like a precise interactive air-route design module have not been developed. Q7 (T = 16.86, p < 0.05) and Q8 (T = 4.32, p < 0.05) indicate user affirmation regarding the useful role of SkyroadAR in the forward-looking research and development of Skyroad. Both professional and non-professional users find this valuable for UAV regulation in this innovative approach.

## 5.4. The Opportunities and Challenges of SkyroadAR

We introduced a novel AR visualization framework for UAVs' low-altitude public air route. It is a geovisualization method that not only applies to Skyroad but also benefits other domains with spatio-temporal information, including trajectory data analysis, environmental simulation, and GIS data visualization. Furthermore, the method does not require the use of a specific AR device. Other head-mounted display (HMD) systems, such as Hololens, can also be used as a carrier for visualization. These characteristics ensure that ARSkyroad is a promising and feasible AR visualization method. On the other hand, geographic information data should not only be used to reconstruct the sandbox, but should be integrated with IMU, RTK-GPS and other multi-sensor data to increase the intelligence of AR full-space scene registration and tracking, and advanced technologies such as cloud rendering should be used to further improve the rendering authenticity, accuracy and efficiency.

Since the SkyroadAR prototype aimed to display the planned UAVs' Skyroad on the physical sandbox, it primarily addressed markerless tracking and virtual-real occlusion, leaving room for improvement in AR interaction. In the future, alignment with domain

experts' needs and targeted user evaluations will enhance software quality. For instance, in research on urban crowdsensing for complex tasks like UAV route planning [48] and resource scheduling [49], using SkyroadAR visualization to verify the results offers an innovative avenue for in-depth exploration.

## 6. Conclusions

In this paper, we proposed an innovative AR sandbox visualization framwork for UAVs' low-altitude public air route, and developed the SkyroadAR prototype system. The framework provides an intuitive and effective visualization for the forward-looking low-altitude digital transport infrastructure. We examined the key technologies in sandbox reconstruction, Skyroad scene production, tracking and registration, and virtual-real occlusion. The system's usability, ease of use, and user intentions were verified through system performance experiments and user questionnaires. The experimental results demonstrate that superimposing the virtual Skyroad scene on the physical sandbox offers an intuitive and efficient environment for expressing UAVs' low-altitude public air route. The improved LINE-MOD template-matching method was based on hierarchical trees and image pyramids, enhancing tracking speed and accuracy. The transparent mask created by the sandbox effectively handles occlusion in the GPU graphics-rendering pipeline, enhancing the user's AR experience with low computational cost.

Nevertheless, the user questionnaire feedback indicates system shortcomings in the interaction function, with the current AR primarily focusing on expressing the Skyroad concept on the sandbox. In future research, specific outdoor UAV air-route AR visualization should be conducted, and the potential of AR methods in assisting UAVs' low-altitude management and applications should be further explored.

In addition, with the development of AI and space technology, AR and GIS visualization will not be limited to sandbox applications. An intelligent perception and understanding of real geographical scenes will be crucial for future AR in fields like robotics and autonomous driving. AR autonomous positioning technology combined with GIS map semantics to improve human–machine coupling indoor and outdoor spatial cognition is expected to achieve better results. Combining AR, AI and GIS map semantics enhances indoor and outdoor spatial cognition and accelerates the development of maps from 2D to 3D, and eventually to 4D, enabling comprehensive geographic expression.

**Author Contributions:** Conceptualization, J.T. and X.L.; methodology, J.T. and H.Y.; software, J.T. and C.X.; validation, J.T. and X.L.; resources, J.T. and H.H.; data curation, J.T. and H.H.; writing—original draft preparation, J.T.; writing—review and editing, J.T., H.Y., X.L., C.X. and H.H.; visualization, J.T.; funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** The Strategic Priority Research Program of the Chinese Academy of Sciences (grant number: XDA28050200).

**Data Availability Statement:** Data available in a publicly accessible repository that does not issue DOIs. Publicly available datasets were analyzed in this study. This data can be found here: https://www.earthexplorer.usgs.gov/, and here: https://www.openstreetmap.org/.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

- 1. Azuma, R.T. A survey of augmented reality. *Presence Teleoperators Virtual Environ*. **1997**, *6*, 355–385. [CrossRef]
- Reed, S.E.; Kreylos, O.; Hsi, S.; Kellogg, L.H.; Schladow, G.; Yikilmaz, M.B.; Segale, H.; Silverman, J.; Yalowitz, S.; Sato, E. Shaping watersheds exhibit: An interactive, augmented reality sandbox for advancing earth science education. In Proceedings of the AGU Fall Meeting Abstracts, San Francisco, CA, USA, 15–19 December 2014; p. ED34A-01.
- 3. Zhang, G.; Gong, J.; Li, Y.; Sun, J.; Xu, B.; Zhang, D.; Zhou, J.; Guo, L.; Shen, S.; Yin, B. An efficient flood dynamic visualization approach based on 3D printing and augmented reality. *Int. J. Digit. Earth* **2020**, *13*, 1302–1320. [CrossRef]
- Rehman, U.; Cao, S. Augmented-Reality-Based Indoor Navigation: A Comparative Analysis of Handheld Devices Versus Google Glass. *IEEE Trans. Hum.-Mach. Syst.* 2017, 47, 140–151. [CrossRef]

- Asraf, S.M.H.; Hashim, A.F.M.; Idrus, S.Z.S. Mobile Application Outdoor Navigation Using Location-Based Augmented Reality (AR). J. Phys. Conf. Ser. 2020, 1529, 022098. [CrossRef]
- Fenais, A.; Ariaratnam, S.T.; Ayer, S.K.; Smilovsky, N. Integrating Geographic Information Systems and Augmented Reality for Mapping Underground Utilities. *Infrastructures* 2019, 4, 60. [CrossRef]
- Zhang, M. Optimization analysis of AR-HUD technology application in automobile industry. In Proceedings of the Journal of Physics: Conference Series, Diwaniyah, Iraq, 21–22 April 2021; p. 012062.
- 8. Ruano, S.; Cuevas, C.; Gallego, G.; García, N. Augmented Reality Tool for the Situational Awareness Improvement of UAV Operators. *Sensors* 2017, *17*, 297. [CrossRef]
- Show Us Your Best Augmented Reality View Photos to Win a Free Flightradar24 Subscription. Available online: https:// www.flightradar24.com/blog/show-us-your-best-augmented-reality-view-photos-to-win-a-free-flightradar24-subscription/ (accessed on 28 January 2023).
- Liu, C.; Shen, S. An Augmented Reality Interaction Interface for Autonomous Drone. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; pp. 11419–11424.
- 11. Liao, X.; Xu, C.; Yue, H. Research on UAV Low-altitude Public Air Route Planning Based on Geographic Information. *Unmanned Veh.* **2018**, *2*, 45–49. (In Chinese)
- 12. He, H.; Ye, H.; Xu, C.; Liao, X. Exploring the Spatial Heterogeneity and Driving Factors of UAV Logistics Network: Case Study of Hangzhou, China. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 419. [CrossRef]
- 13. Zhang, N.; Zhang, M.; Low, K.H. 3D path planning and real-time collision resolution of multirotor drone operations in complex urban low-altitude airspace. *Transp. Res. Part C Emerg. Technol.* **2021**, *129*, 103123. [CrossRef]
- 14. Ozturkcan, S. Service innovation: Using augmented reality in the IKEA Place app. J. Inf. Technol. Teach. Cases 2021, 11, 8–13. [CrossRef]
- Sayed, N.A.M.E.; Zayed, H.H.; Sharawy, M.I. ARSC: Augmented Reality Student Card. In Proceedings of the 2010 International Computer Engineering Conference (ICENCO), Cairo, Egypt, 27–28 December 2010; pp. 113–120.
- 16. Bobrich, J.; Otto, S. Augmented maps. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2002, 34, 502–505.
- Sánchez, S.Á.; Martín, L.D.; Gimeno-González, M.Á.; Martín-Garcia, T.; Almaraz-Menéndez, F.; Ruiz, C. Augmented reality sandbox: A platform for educative experiences. In Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality, Salamanca, Spain, 2–4 November 2016; pp. 599–602.
- Gong, J.; Li, W.; Zhang, G.; Shen, S.; Huang, L.; Sun, J. An Augmented Geographic Environment for Geo-process Visualization: A Case of Crowd Evacuation Simulation. *Acta Geod. Cartogr. Sin.* 2018, 47, 1089–1097. (In Chinese)
- Tuzun Canadinc, S.; Yan, W. 3D-Model-Based Augmented Reality for Enhancing Physical Architectural Models. In Proceedings of the 40th Conference on Education and Research in Computer Aided Architectural Design in Europe (eCAADe 2022), Ghent, Belgium, 13–16 September 2022.
- George, R.; Howitt, C.; Oakley, G. Young children's use of an augmented reality sandbox to enhance spatial thinking. *Child. Geogr.* 2020, 18, 209–221. [CrossRef]
- 21. Ma, L.F.; Huang, T.Q.; Wang, J.; Liao, H.E. Visualization, registration and tracking techniques for augmented reality guided surgery: A review. *Phys. Med. Biol.* 2023, *68*, 04TR02. [CrossRef]
- 22. Fenais, A.; Ariaratnam, S.T.; Smilovsky, N. Assessing the Accuracy of an Outdoor Augmented Reality Solution for Mapping Underground Utilities. *J. Pipeline Syst. Eng. Pract.* 2020, *11*, 04020029. [CrossRef]
- 23. Chen, X.; Li, H.; Zhou, C.Y.; Liu, X.; Wu, D.; Dudek, G. Fidora: Robust WiFi-Based Indoor Localization via Unsupervised Domain Adaptation. *IEEE Internet Things J.* 2022, *9*, 9872–9888. [CrossRef]
- 24. Zhang, X.; Fronz, S.; Navab, N. Visual marker detection and decoding in ar systems: A comparative study. In Proceedings of the International Symposium on Mixed and Augmented Reality, Darmstadt, Germany, 30 September–1 October 2002; pp. 97–106.
- Arifitama, B.; Syahputra, A.; Permana, S.D.H.; Bintoro, K.B.Y. Mobile Augmented Reality for Learning Traditional Culture Using Marker Based Tracking. In Proceedings of the 2nd International Conference on Informatics, Engineering, Science, and Technology (INCITEST 2019), Bandung, Indonesia, 18 July 2019.
- Rabbi, I.; Ullah, S. Extending the Tracking Distance of Fiducial Markers for Large Indoor Augmented Reality Applications. *Adv. Electr. Comput. Eng.* 2015, 15, 59–64. [CrossRef]
- Duan, L.Y.; Guan, T.; Yang, B. Registration Combining Wide and Narrow Baseline Feature Tracking Techniques for Markerless AR Systems. *Sensors* 2009, 9, 10097–10116. [CrossRef]
- Lin, L.; Wang, Y.T.; Liu, Y.; Xiong, C.M.; Zeng, K. Marker-less registration based on template tracking for augmented reality. *Multimed. Tools Appl.* 2009, 41, 235–252. [CrossRef]
- Song, J.; Kook, J. Visual SLAM Based Spatial Recognition and Visualization Method for Mobile AR Systems. *Appl. Syst. Innov.* 2022, 5, 11. [CrossRef]
- Cordes, K.; Scheuermann, B.; Rosenhahn, B.; Ostermann, J. Occlusion Handling for the Integration of Virtual Objects into Video. In Proceedings of the VISAPP, Rome, Italy, 24–26 February 2012; Volume 2, pp. 173–180.
- Tian, Y.; Guan, T.; Wang, C. Real-Time Occlusion Handling in Augmented Reality Based on an Object Tracking Approach. Sensors 2010, 10, 2885–2900. [CrossRef]

- 32. Tian, Y.; Guan, T.; Wang, C. An automatic occlusion handling method in augmented reality. *Sens. Rev.* 2010, 30, 210–218. [CrossRef]
- Kim, H.; Sohn, K. Hierarchical depth estimation for image synthesis in mixed reality. In Proceedings of the Stereoscopic Displays and Virtual Reality Systems X, Santa Clara, CA, USA, 21–24 January 2003; pp. 544–553.
- 34. Izadi, S.; Kim, D.; Hilliges, O.; Molyneaux, D.; Newcombe, R.; Kohli, P.; Shotton, J.; Hodges, S.; Freeman, D.; Davison, A.; et al. KinectFusion: Real-time 3D reconstruction and interaction using a moving depth camera. In Proceedings of the 24th annual ACM Symposium on User Interface Software and Technology, Santa Barbara, CA, USA, 16–19 October 2011; pp. 559–568.
- Newcombe, R.A.; Izadi, S.; Hilliges, O.; Molyneaux, D.; Kim, D.; Davison, A.J.; Kohi, P.; Shotton, J.; Hodges, S.; Fitzgibbon, A. KinectFusion: Real-time dense surface mapping and tracking. In Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality, Basel, Switzerland, 26–29 October 2011; pp. 127–136.
- 36. Davison. Real-time simultaneous localisation and mapping with a single camera. In Proceedings of the Ninth IEEE International Conference on Computer Vision, Nice, France, 13–16 October 2003; Volume 1402, pp. 1403–1410.
- 37. Klein, G.; Murray, D. Parallel Tracking and Mapping for Small AR Workspaces. In Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, Nara, Japan, 13–16 November 2007; pp. 225–234.
- Xu, C.; Liao, X.; Yue, H.; Lu, M.; Chen, X. Coustuction of a UAV Low-alititude Public Air Route based on an Improved Ant Colony Algorithm. J. Geo-Inf. Sci. 2019, 21, 570–579. (In Chinese)
- 39. Xu, C.; Ye, H.; Yue, H.; Tan, X.; Liao, X. Iterative construction of UAV low-altitude air route network in an urbanized region: Thoretical system and technical roadmap. *Acta Geogr. Sin.* **2020**, *75*, 917–930. (In Chinese)
- 40. Liao, X.; Xu, C.; Ye, H.; Tan, X.; Fang, S.; Huang, Y.; Lin, J. Critical infrastructures for developing UAVs' applications and low-altitude public air-route network planning. *Bull. Chin. Acad. Sci.* **2022**, *37*, 977–988. (In Chinese) [CrossRef]
- 41. Qu, W.; Xu, C.; Tan, X.; Tang, A.; He, H.; Liao, X. Preliminary Concept of Urban Air Mobility Traffic Rules. *Drones* 2023, 7, 54. [CrossRef]
- 42. Wiedemann, C.; Ulrich, M.; Steger, C. Recognition and Tracking of 3D Objects. In Proceedings of the Pattern Recognition, Berlin/Heidelberg, Germany, 10–13 June 2008; pp. 132–141.
- Hinterstoisser, S.; Cagniart, C.; Ilic, S.; Sturm, P.; Navab, N.; Fua, P.; Lepetit, V. Gradient Response Maps for Real-Time Detection of Textureless Objects. *IEEE Trans. Pattern Anal. Mach. Intell.* 2012, 34, 876–888. [CrossRef]
- 44. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.
- 45. Hare, S.; Saffari, A.; Torr, P.H.S. Efficient online structured output learning for keypoint-based object tracking. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 1894–1901.
- 46. Lucas, B.D.; Kanade, T. An iterative image registration technique with an application to stereo vision. In Proceedings of the IJCAI'81: 7th International Joint Conference on Artificial Intelligence, Vancouver, BC, Canada, 24–28 August 1981.
- 47. Davis, F.D. A technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results; Massachusetts Institute of Technology: Cambridge, MA, USA, 1985.
- Xiang, C.; Zhou, Y.; Dai, H.; Qu, Y.; He, S.; Chen, C.; Yang, P. Reusing delivery drones for urban crowdsensing. *IEEE Trans. Mob. Comput.* 2021, 22, 2972–2988. [CrossRef]
- 49. Xiang, C.; Cheng, W.; Zheng, X.; Wu, T.; Fan, X.; Wang, Y.; Zhou, Y.; Xiao, F. Enabling Cost-effective Wireless Data Collection by Piggybacking on Delivery Drones in Agriculture. *ACM Trans. Sens. Netw.* **2023**. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.