Creation of a Virtual Tour .Exe Utilizing Very High-Resolution RGB UAV Data

Umut Gunes SEFERCİK, Taskin KAVZOGLU, Mertcan NAZAR, Can ATALAY, Muhammed MADAK

Chief in Editor
Prof. Dr. Cem Gazioğlu

Co-Editors  Prof. Dr. Dursun Zafer Şeker, Prof. Dr. Şinasi Kaya,
Prof. Dr. Ayşegül Tank and Assist. Prof. Dr. Volkan Demir

Editorial Committee (December 2022)
Assoc. Prof. Dr. Abdullah Aksu (TR), Assoc. Prof. Dr. Uğur Algancı (TR),
Prof. Dr. Levent Bat (TR), İrșad Bayırhan (TR), Assist. Prof. Dr. Tolga Bakırman (TR), Prof. Dr. Bülent Bayram (TR), Prof. Dr. Nuray Çağılar (TR), Prof. Dr. A. Evren Erginal (TR), Assoc. Prof. Dr. Cüneyt Erenoğlu (TR),
Prof. Dr. Çiğdem Göksel (TR), Prof. Dr. Manik Kalubarmé (IN), Dr. Hakan Kaya (TR), Assist. Prof. Dr. Serkan Kükrer (TR), Assoc. Prof. Dr. Maged Marghany (MY), Prof. Dr. Hasan Özdemir (TR), Prof. Dr. Barış Salıhoğlu (TR),
Assist. Prof. Dr. Başak Savun-Hekimoğlu (TR), Prof. Dr. Elif Sertel (TR), Prof. Dr. Füsun Balık Şanlı (TR),
Dr. Duygu Ülker (TR), Assoc. Prof. Dr. Ömer Suat Taşkın (TR), Assist. Prof. Dr. Tuba Ünsal (TR),
Assist. Prof. Dr. Sibel Zeki (TR)

Creation of a Virtual Tour .Exe Utilizing Very High-Resolution RGB UAV Data

Umut Gunes Sefercik¹, Taskin Kavzoglu¹, Mertcan Nazar*¹, Can Atalay²ID, Muhammed Madak³ID

¹Gebze Technical University, Faculty of Engineering, Department of Geomatics Engineering, Kocaeli, TURKIYE
²Ecevit University, Faculty of Engineering, Department of Geomatics Engineering, Zonguldak, TURKIYE
³Department of Digitalization, Gizil Enerji A.S., Istanbul, TURKIYE
* Corresponding author: M. Nazar
E-mail: mnazar@gtu.edu.tr
Received 13.04.2022
Accepted 22.08.2022


Abstract

In the last decades, developments in game engine technology led to a raised attraction to the virtual reality (VR) and augmented reality (AR) concepts which offer users an interactive synthetic environment. Also, with the travel limitations of the current COVID-19 pandemic, VR tour applications that visualize the geospatial data gained popularity more than ever. In this study, a three-dimensional (3D) VR tour application was created for Gebze Technical University (GTU) Campus by integrating unmanned aerial vehicle (UAV) data into an artificial environment by using the cross-platform game development engine Unity. For creating high-quality 3D models of the Campus, different imaging geometries and flight altitudes were applied. The aerial photos were achieved with a ground sampling distance (GSD) of ≤2.2 cm with a 20 megapixel (MP) Sony Exmor RGB camera. Point cloud processing and the generation of high-quality 3D products were carried out by structure from motion (SfM) based software Agisoft Metashape. Using 86 well-distributed ground control points (GCPs), geometric correction accuracy of ±2 cm (~0.9 pixels) was reached as root mean square error (RMSE). Generated 3D models were imported into the Unity environment and the negative influence of high polygon data on the application performance was reduced by applying occlusion culling and space subdivision rendering optimization algorithms. The visual potential of the VR was improved by adding 3D individual object models such as trees, benches, and arbors. For enhancing the information content of the VR tour, interactive information panels including the building metadata such as building name, block name, and total floor area were placed. Finally, a first-person player was implemented for a realistic VR experience.

Keywords: UAV, SFM, 3D Textured Mesh Model, Virtual Reality, Unity

Introduction

Physical reality is a concept in which humanity has deep knowledge and is in constant interaction. Through continuous connection with the physical environment, humans learn to foresee features of the physical world and outcomes of certain events. However, digital computers, offering users a look into a “mathematical wonderland”, are used to have an understanding of unfamiliar concepts (Sutherland, 1965). Computer systems provide the opportunity for users to completely immerse themselves inside an interactive digital environment through virtual reality (VR) technology, a specific section of computer graphics (Brooks, 1999). Advancements in game engines caused an increased interest in VR technology, including various applications both in commercial and scientific fields. VR systems grow into a capable tool for the visualization of geospatial data and the generation of high-quality three-dimensional (3D) models for virtual tourism (Guttentag, 2010; Huang et al., 2016; Yung and Khoo-Lattimore, 2019), archeology (Bruno et al., 2010; Rua and Alvito, 2011; Bruno et al., 2018), cultural heritage documentation (Giatatzes et al., 2001; Carrozzino and Bergamasco, 2010; Bekele et al., 2018), forensic applications (Ma et al., 2010; Sieberth et al., 2019), physical therapy (Hoffman et al., 2000; Schmitt et al., 2011), exposure therapy (Powers and Emmelkamp, 2008; Carl et al., 2019) and so on. The current COVID-19 pandemic, forcing people to stay at home as a safety measure against the virus, restricted social activities. Due to limitations in the interaction between people, virtual tours which allow users to immersive themselves inside a place in a synthetic environment, have become popular. In addition, on a larger scale produced virtual tours can be integrated into the metaverse, the Internet of next-generation in which users can interact with other users and software applications through avatars in a 3D virtual environment (Duan et al., 2021). While metaverse offers an opportunity to reach a wider audience for VR applications it facilitates the development of VR technology and provides connectivity with web-based applications.

Unmanned aerial vehicles (UAVs), remotely controlled unmanned flying platforms, have become an efficient tool in photogrammetric studies. Due to offering high-resolution, periodical, and low-cost data, UAVs are large in demand both in several scientific and commercial applications. Also, breakthroughs in the technological
field progressively improve the efficiency of UAVs. Day by day, UAV data became indispensable in several applications such as cultural heritage documentation (Eisenbeiss and Zhang, 2006; Lo Brutto et al., 2014; Martínez-Carricondo et al., 2020; Karakaş and Altuğun, 2020), archeology (Lin et al., 2011; Smith et al., 2014; Campana, 2017), architectural modelling (Achille et al., 2015; Silva et al., 2020), natural hazard monitoring (Hirokawa et al., 2007; Gazioglu et al., 2017; Annis et al., 2020; Ulu et al., 2020; Bayırhan and Gazioglu, 2020; İncekara and Şeker, 2021), water quality analysis (Ozdoğan et al., 2021) and precision (digital) agriculture (Bendig et al., 2013; Tsoulos et al., 2019). The most significant reason for this large demand is easily generated high-quality 3D dense point clouds and textured mesh models by means of high-resolution UAV aerial photos using structure from motion (SfM) based image matching software such as Agisoft Metashape, Pix4D, and RealityCapture (Sefercik et al., 2019; Tan and Li, 2019; Pepe and Costantino et al., 2021).

In this study, the creation of a 3D VR tour application for the Gebze Technical University (GTU) Campus was purposed by integrating high-quality UAV-based 3D models into a virtual environment. Accordingly, by means of high-resolution UAV data, qualified 3D textured mesh models were generated in SfM-based software Agisoft Metashape and integrated into the virtual environment by the cross-platform game development engine Unity. Created 3D VR tour application will provide the opportunity for student candidates, academicians, researchers, and other visitors to remotely visit and get information about the Campus. Also, users will take part in an immersive and realistic virtual environment providing convenience during the COVID-19 pandemic. Furthermore, generated 3D Campus model can be used for the planning and construction of upcoming buildings and landscapes. The study consists of five different sections. In the next section, the study area and specifications of utilized materials are described. In section 3, the methodology of 3D textured mesh model generation and 3D VR tour creation are given. Section 4 shows the results and discussions followed by the conclusion and future remarks.

Study Area and Materials

Study Area

The study area is the GTU Campus, located in the western part of the Gebze district of Kocaeli Metropolitan, Turkey. The area has several land cover classes such as forest, agricultural, built-up, road, and water bodies. Especially in the northern part of the Campus, dense forests make the 3D modelling of the forest understory difficult. The area is approx. 2.5 km² and the orthometric elevation of the bare ground is between 1-50 m. Figure 1 shows the location of Kocaeli Province in Turkey and the UAV orthomosaic of the GTU Campus in the geographic coordinate system and WGS84 datum.

Materials

The UAV flights were completed by DJI Phantom IV Pro V2.0 optical UAV which has a 20 MP spatial resolution Sony Exmor RGB camera. The mobile ground control points (GCPs), consisting of plus-shaped two 0.25 × 1 m polycarbonates, were used for the absolute orientation and they were measured by CHC i80 Global Navigation Satellite System (GNSS) receivers. All of the utilized equipment, shown in Figure 2, is available at GTU Geomatics Engineering Department's Advanced Remote Sensing Technology Laboratory (ARTLAB).
Table 1 shows the specifications of the utilized DJI Phantom IV Pro V2.0 optical UAV and CHC i80 GNSS receiver. With the advantage of an 8.6 mm focal length of a 20 MP Sony Exmor camera, the UAV enables to achieve data from the target area with approx. 1 cm ground sampling distance (GSD) applying 40 m flying altitude.

Table 1. Specifications of DJI Phantom IV Pro V2.0 UAV and CHC i80 GNSS receiver

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DJI Phantom IV Pro V2.0 UAV</strong></td>
<td></td>
</tr>
<tr>
<td>Camera</td>
<td>4K, HD, 1080p, 1”, effective pixel resolution 20 MP</td>
</tr>
<tr>
<td>Gimbal</td>
<td>3-axis (pitch, roll, yaw)</td>
</tr>
<tr>
<td>Flight duration</td>
<td>Max. 30 minutes</td>
</tr>
<tr>
<td>Weight</td>
<td>1375 g</td>
</tr>
<tr>
<td>Speed</td>
<td>Max 20 m/s in S-mode</td>
</tr>
<tr>
<td>Wind speed resistance</td>
<td>Max. 10 m/s</td>
</tr>
<tr>
<td>Operating temperature</td>
<td>0° to 40°C</td>
</tr>
<tr>
<td>Outdoor positioning module</td>
<td>GPS/GLONASS dual</td>
</tr>
<tr>
<td>Hover accuracy range</td>
<td>± 0.1 m V, ± 0.5 m H (Vision)</td>
</tr>
<tr>
<td></td>
<td>± 0.3 m V, ± 1.5 m H (GPS)</td>
</tr>
<tr>
<td><strong>CHC i80 GNSS Receiver</strong></td>
<td></td>
</tr>
<tr>
<td>GNSS technology</td>
<td>GPS, GLONASS, GALILEO, BeiDou, SBAS, NavIC</td>
</tr>
<tr>
<td>Operating system</td>
<td>Linux</td>
</tr>
<tr>
<td>Working modes</td>
<td>Static, VRS RTK, UHF RTK, all surveying modes</td>
</tr>
<tr>
<td>Internal memory</td>
<td>32 GB</td>
</tr>
<tr>
<td>Positioning accuracy RTK</td>
<td>± 0.8 cm H, ± 1.5 cm V with initialization reliability &gt;99.9%</td>
</tr>
<tr>
<td>Battery</td>
<td>Dual; Static up to 10 h, Cellular receive only up to 9h, UHF receive/transmit up to 6h</td>
</tr>
</tbody>
</table>

![Fig. 3. Sub-stages of 3D textured mesh model generation in Agisoft (upper) and 3D VR tour creation in Unity (lower).](image)

![Fig. 4. Flight plans prepared in Pix4Dcapture over GTU Campus.](image)
In addition, stereo-modelling of approx. 20 ha area is possible within the limits of one safe flight.

**Methodology**

The utilized methodology has two main stages, 3D textured mesh model generation in Agisoft Metashape Professional software and 3D VR tour creation in the Unity game engine. Figure 3 shows the sub-processing stages, completed in Agisoft and Unity. The GCP measurements are also presented under Agisoft sub-stages.

**Flight Planning and GCP Measurements**

The flight plans of the 2.5 km² Campus area were prepared in Pix4Dcapture software as polygonal, bundle-grid, and circular considering the land use and land cover properties of individual flight zones. While polygonal flights are preferred for forest and agricultural zones, bundle-grid and circular flights were planned for built-up zones. In bundle-grid and circular flying modes, 70° camera viewing angle (off-nadir 20°) was preferred while nadir view is applied in polygonal flights. Regarding the geometry of the flight plans, 86 mobile GCPs were established and measured using the continuously operating reference stations (CORS) real-time kinematic (RTK) GNSS technique. In UAV flights, minimum front and side overlap ratios were 80% and 60%, respectively, and the flying altitude was chosen 80 m for both polygonal and bundle grid flights and 30 m for circular flights. With 32 UAV flights, 8333 aerial photos were obtained with ≤2.2 cm GSD. Figure 4 illustrates the prepared flight plans over the study area.

**Orientation of Aerial Photos**

The orientation procedure was carried out in two steps, generating a sparse point cloud by initial (mutual) alignment of aerial photos and absolute orientation utilizing GCPs. For processing a huge number of aerial photos (>2000), a high-capacity workstation and chunk-based study are necessary. In this study, aerial photos were processed in different chunks by a high-end HP Z4 G4 workstation which has an Intel® Xeon® W-2133 CPU @ 3.60 GHz, NVIDIA Quadro P2000 GPU, and 64 GBs of RAM. Agisoft Metashape is a SfM-based software for photogrammetric evaluation and 3D geospatial data generation. The SfM approach is a low-cost and adequate photogrammetric method for geospatial data visualization with the ability to reconstruct 3D geometry in high resolution from a series of overlapping offset photos (Westoby et al., 2012). In SfM-based image matching, 3D point cloud data is generated using different views of sequencing 2D data such as aerial photos and satellite imagery. Figure 5 presents an instance of a circular sequencing photo acquisition for generating a 3D point cloud using the SfM algorithm.

In the aerial photos captured with oblique viewing geometry in bundle-grid and circular flights, it was necessary to filter background pixels to eliminate the effect of noise in the orientation. In addition, the accuracy of orientation was increased by removing the influence of points with low correlation. Therefore, a masking process was conducted by selecting the background pixels from the aerial photos. Figure 6 shows a sample of the masking process.

---

Fig. 5. An example of circular sequencing photo acquisition for the SfM technique

Fig. 6. Aerial photo before (a) and after (b) masking

After the masking process, alignment of aerial photos was initiated and sparse point clouds with a total size of >73 million points were obtained. Then, GCP coordinates were imported as a reference into the Agisoft Metashape Professional for absolute orientation. GCPs were selected on aerial photos by checking the root mean square error (RMSE) values simultaneously to avoid inaccurate selection. To carefully select the center of a GCP for a reliable orientation, aerial photos were zoomed into a suitable level as shown in Figure 7.
The geometric accuracy was calculated according to selected GCPs as RMSE in both meter and pixel values using equation 1. Finally, the achieved geometric accuracy was ±2 cm (~0.9 pixels). In the following RMSE formula for i camera position $\hat{X}_i$, $\hat{Y}_i$, $\hat{Z}_i$ represent estimated values, and $X_i$, $Y_i$, $Z_i$ shows the actual input values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{X}_i - X_i)^2 + (\hat{Y}_i - Y_i)^2 + (\hat{Z}_i - Z_i)^2}{n}}$$ (1)

**Generation of Dense Point Clouds**

The sparse point cloud represents the topography and objects in low-resolution 3D vector data. However, a more realistic and complex characterization is required for producing high-quality 3D models. Therefore, by utilizing depth maps, built by extracting depth information from aerial photos with sufficient overlap, densely populated point clouds were generated with a total size of approx. 1.5 billion. Despite the relatively high RAM capacity (64 GB) of the workstation, dense point cloud generation took 5-12 hours per chunk. In generated dense point clouds, the occurrence of noisy parts is expected because of distortions in imaging geometry and the influence of land cover classes such as forest, vegetation, and water bodies. The dense point clouds were filtered by fencing and classifying noisy points. A general view of generated dense point clouds before and after the filtering is given in Figure 8.

**Generation of 3D Textured Mesh Models**

Dense point cloud data visualize the ground and objects in non-continuous point-based vector format, so for models to be more realistic and detailed, it is essential to transform them into the 3D solid form. Furthermore, a more immersive 3D VR tour can be created by displaying objects and surfaces in a continuous format without any gaps or holes. So, using filtered dense point clouds 3D mesh models were generated. In addition, interpolation was enabled for filling gaps that may appear because of the limited number of points. Moreover, due to the number of predicted faces (triangles) and vast size of polygonal data, mesh decimation was conducted and the number of generated mesh faces was reduced to approx. 430 million faces. However, generated 3D meshes contained blocks of shapes with rough geometry and isolated polygons. To solve the aforementioned problem a manual mesh filtering procedure was realized. To obtain 3D mesh models similar to their counterparts in physical reality, it is essential to apply high-resolution texture graphics to them. In Agisoft, generated 3D meshes are visualized in the shaded mode which depicts objects with color but in low resolution compared to a texture graphic. So, utilizing UAV aerial photos a high-resolution texture was generated and applied to generated 3D mesh models (Figure 9).

**3D VR Tour Creation**

As mentioned before, for creating a 3D VR tour application, generated 3D textured mesh models were integrated into an artificial environment using the Unity game engine. Unity is a cross-platform game development system, operating in the C# programming language. Utilizing the Unity engine, games for websites, computers, mobile devices, console systems, and console platforms can be developed for entertainment or education. For importing 3D meshes to Unity, it was essential to export them from the Agisoft Metashape in a proper format for efficiency and practicality. Accordingly, 3D mesh models were exported in Wavefront OBJ format and imported in Unity which supports OBJ format for 3D models. As a conventionally utilized 3D data format Wavefront OBJ stores both point coordinates and polygonal data and it is supported by a lot of 3D computer-aided design (CAD) software (Kato and Ohno, 2009). A material that contains texture and lighting properties of the shader
object, is required to be created in Unity for proper visualization of imported textures along with the 3D meshes. Moreover, using file inspector window materials were edited by selecting shader type as unlit/texture for providing smooth views of textures over meshes. Using the inspector, the max size was selected as 16384 for high-resolution texture presentation. Realistic lighting was produced by adding a directional light for rendering shadows of objects over the environment. Because of the sheer number of faces in imported 3D models and thus the large size of polygonal data, occlusion culling and space subdivision rendering optimization algorithms were applied using tools in Unity. For improving the performance and overcoming the hardware bottlenecks, occlusion culling algorithms are utilized to recognize and render visible areas of 3D structures so that objects placed on the backside of a specific viewpoint are not rendered (Coorg and Teller, 1997; Sefercik et al., 2021).

Applied optimization algorithms have a minor impact on processing time, making the whole process a lot more efficient. Furthermore, to improve visualization capability and offer users a more realistic VR experience, premade 3D object models such as lighting poles, arbors, benches, and trees were added to the application. In addition, for allowing visitors to have information about Campus buildings during the VR tour, textual metadata including building name, block number, total floor area, and total usage area was added using interactive information panels with pop-up window ability. To construct an immersive VR tour experience, users should be able to walk around buildings and objects as if they are walking in real life. Therefore, a player with a first-person camera was added to the application for a realistic virtual experience. The first-person player consists of a character controller component, a capsule mesh, scripts for player movement, and a camera programmed to follow the movement. A collider was also added as a component to the first-person player for allowing the player to detect and contact the faces of generated 3D mesh models in the Unity environment (Jafri et al., 2017). Finally, an executable file (EXE) was built in Unity for the 3D VR tour application.
Results

Figure 11 shows the generated sparse and dense point clouds and the final 3D textured mesh model of the GTU Campus area produced by combining all generated 3D mesh models. In the Figure, the superiority of the 3D textured raster mesh model against vector point clouds in realistic description potential is clear. Using the first-person player with an orientable camera function, users can easily adjust their movements and look at where they desire using basic keyboard and mouse controls.

![Sample scenes from the 3D VR tour application in the Unity game engine from the viewpoint of the visitor](image)

Figure 12 shows the views of the first-person player in sample scenes of created virtual tour application. In these views, the environment looks very clear by means of a high-resolution GSD achieved by a 20 MP RGB camera. Figure 13 illustrates a sample interactive information panel. When the visitor looks at the information sign (i in the circle), the information panel with textual metadata opens.

![Sample scenes from the 3D VR tour application in the Unity game engine from the viewpoint of the visitor](image)

Discussion of the Results

While checking the 3D models some gaps are noticed especially on the thin cover walls of the building roofs (Figure 14). The main reason for this situation is non-continuous vertical geometry where the alignment results in a small number of points due to low correlation. Hence, an insufficient number of points are used in interpolation which causes gaps. These gaps were filled by using the hole filling tool in Agisoft Metashape software.

![Gaps on building rooftop surroundings](image)

Optical imaging systems are not able to penetrate dense forest or vegetation areas as in laser scanning (lidar) or large wavelength synthetic aperture radar (SAR) systems due to operating with the passive sensing principle. That means sensing dense forest or vegetation understory is impossible with a nadir view. That’s why it is of utmost importance to capture dense forest and vegetation areas using different viewing geometries and oblique camera
angles. Along with the frequently adopted polygonal flight geometry with nadir-view, it is vital to conduct different UAV flight modes such as bundle-grid and circular flights. Figure 15 shows an instance of a dense vegetation problem on the GTU Campus. Due to the lack of points under dense trees, a bicycle road cannot be described correctly in a 3D mesh model. In similar cases, we recommend two solutions if the problem cannot be solved by optical UAV flights. The first solution is a lidar UAV flight in the problematic area. However, the laser scanner should have multiple returns (more than two) to penetrate dense vegetation. The second solution is to have dense point clouds with colorful terrestrial laser scanning (TLS) in problematic areas and merge them with UAV dense point clouds and re-generation of the 3D textured mesh model.

Conclusion

In this study, the GTU Campus area was captured by using high-resolution UAV aerial photos for the purpose of producing high-quality 3D models which were integrated into a virtual domain to create a 3D VR tour application. Obtained aerial photos were oriented in SfM-based Agisoft Metashape software using 86 GCPs and ±2 cm (~0.9 pixels) geometric accuracy was calculated as RMSE. After the orientation process, sparse and dense point clouds were generated. However, dense point clouds had some noise because of moving objects and the influence of land cover classes such as forest, so a filtering procedure was applied. Adopting filtered dense point clouds, solid 3D mesh models were produced. Then, using high-resolution aerial photos, high-quality texture graphics were constructed and applied to 3D meshes for a realistic display. VR integration was done using the Unity game engine for importing 3D models into a virtual domain and rendering optimization algorithms were applied for better performance. Moreover, premade 3D object models of lighting poles, arbors, benches, and trees were placed in a virtual environment for creating a genuine VR tour. Also, building name, block number, total floor, and usage area were added as textual metadata to the buildings in the form of an interactive information panel. Lastly, a first-person player was implemented for users to freely walk around the Campus and gather information about buildings. Overall, integrating 3D textured models, generated from high-resolution UAV data, to the virtual environment by means of the Unity game engine a qualified 3D VR tour application was developed. This application will be available through the GTU website for visitors. Additionally, the 3D model of the Campus can be used for planning, management, and construction activities. Due to having real-world coordinates the product can be integrated into a larger dataset in the scope of a building information modeling (BIM) study or for creating a 3D smart city model.

Acknowledgements

We would like to thank Gebze Technical University for supporting this research in the scope of a scientific research project (ID: 2020-A-105-42).

References


Bayrhyth, I., Gazioglu, C. (2020). Use of Unmanned Aerial Vehicles (UAV) and Marine Environment Simulator in Oil Pollution Investigations, Baltic J. Modern Computing, 8(2), 327-336, DOI: 10.22364/bjmc.2020.8.2.08


