A Combination of Geomatic Techniques for Modelling the Urban Environment in Virtual Reality

A combination of geomatic techniques for modelling the urban environment in Virtual Reality

Claudio Spadavecchia[1[0000-0003-2087-9828]], Elena Belcore[1[0000-0002-3592-9384]], Vincenzo Di Pietra[1[0000-0001-7501-1183]], Nives Grasso[1[0000-0002-9548-6765]]

[1]DIATI, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
claudio.spadavecchia@polito.it; elena.belcore@polito.it; vincenzo.dipietra@polito.it; nives.grasso@polito.it

Abstract. Historical city centres represent invaluable cultural and architectural heritage, often facing challenges such as neglect, urbanisation, and dwindling visitor engagement. Italian city centres are usually rich in historical buildings, statues and open-sky arts, causing difficulties in changing the urbanisation tissue and promoting infrastructural works in front of the citizens. Therefore, it is fundamental not only to document and survey the current state with geomatics techniques but also to propose engaging and educational experiences for citizens exploiting the potential of VR. In fact, VR offers an immersive and interactive medium that can bridge the gap between the past and the present, revitalising these urban gems while preserving their historical significance. This paper focuses on two aspects of historical city centres' digital documentation: (a) the geomatics integrated survey exploiting LiDAR and UAV photogrammetric techniques and (b) the process of optimising geomatics products into VR experiences. A big issue in transposing dense digital models into a VR environment is the information's weight, which prevents the app from running smoothly. We aim to comprehensively examine the potential, challenges, and best practices for processing surveying data at optimal efficiency with respect to a VR environment.

Keywords: VR environment, 3D model, Cultural Heritage, LiDAR, UAS, Geomatic.

1 Introduction

Historic city centres represent cultural and architectural treasures of immeasurable value [1]. They serve as custodians of history, reflecting traces of bygone eras through their architecture, streets, and monuments. These places captivate not only with their aesthetic beauty but also with their ability to tell stories, revealing the evolution of societies over time. The preservation of historic city centres is essential for safeguarding cultural diversity, promoting intercultural understanding, and providing a tangible connection to our past [2]. Furthermore, they contribute to sustainable tourism, fostering the conservation and appreciation of our heritage riches [3, 4]. Nowadays, Virtual Reality (VR) plays a pivotal role in the preservation and promotion of cultural heritage

[5–7]. It offers an immersive platform that allows individuals to step back in time and explore historical and cultural sites with unprecedented realism and interactivity. VR enables the creation of faithful digital replicas of heritage sites, artefacts, and artworks, ensuring their conservation for future generations [8–10]. Moreover, it enhances educational experiences by providing students and enthusiasts with dynamic, hands-on learning opportunities, making history and culture more engaging and accessible [11–13]. Additionally, VR contributes to the accessibility of cultural heritage, enabling individuals with physical limitations or geographical constraints to experience and appreciate these treasures [14]. By blending technology with culture, VR revolutionises how we engage with our past, fostering a deeper connection to our heritage and encouraging its preservation for posterity.

Virtual reality is a powerful technology that promises to change our lives [15] by simulating a virtual environment that immerses users to the extent that they have the feeling of "being there" [16]. In recent years, virtual reality has made a comeback mainly in the field of gaming and entertainment and is supported by low production costs and affordable retail prices; however, this technology is nowadays widely used in various research fields such as medicine [17–20], military purposes [21], and education [22]. More recently, this technology has been used for cultural heritage domain [23, 24] and tourist and informational purposes, such as virtual museum visits [25, 26], where users can virtually move between art installations and inspect them.

Recent studies have proposed procedures for taking advantage of three-dimensional photogrammetric or LiDAR models and recreating virtual reality environments [27, 28], mainly focusing on mesh generation and importing in a virtual environment. Nevertheless, to the best of our knowledge, no study describes the complete workflow to follow in a detailed and exhaustive way to acquire a georeferenced three-dimensional model to be imported and visualised in VR. Photogrammetry is the process of deriving metric information about an object through measurements made on object photographs [29]. This well-established methodology is widely used in civil engineering [30, 31], mapping [32] and 3D modelling [33, 34]. 3D modelling is a process that starts from data acquisition and ends with the final 3D virtual model [33], and it can now be considered one of the most important and attractive products of photogrammetry [35, 36]. Traditional 3D models are generated with aerial and/or terrestrial images [33, 37] according to consolidated Structure from Motion/Multi-View Stereo (SfM-MVS) workflows [38, 39]. However, with the advent of LiDAR (Light Detection and Ranging) as a commonly used instrument in the remote sensing field, new three-dimensional models are generated using this approach [40, 41] or by combining it with the photogrammetric procedure and according to standardised procedures [33, 42–44]. Nowadays, generating a 3D model through an integrated approach is preferred because of the ease of acquisition, mainly due to the photogrammetric part, and the simultaneous high accuracy due to the LiDAR survey.

When discussing VR experience with headset devices, it is important to build applications optimised enough to run smoothly on their target devices and avoid discomfort for the user [45]. This is particularly true when the developed application must reproduce digitally real objects to let the user interact with them [46]. Optimising a VR project is a complex task involving developers and artists to simulate the complex reality

in a virtual environment. While modelling from scratch a digital asset unrelated to the physical world is a task related to digital artists, reconstructing existing objects, structures or environments is nowadays demanded to the potentiality of the photogrammetric process [47, 48]. With photogrammetry, creating extremely high-quality assets for non-VR usage is possible, mainly required in the engineering architecture and environment fields [49]. The workflow applied for photogrammetric reconstruction is primarily devoted to producing cartographic products like Digital Elevation Models (DTM) and Orthomosaic Maps. Point clouds are sometimes used to perform specific tasks like semantic segmentation, classification or 3D mapping [50]. In this panorama, the optimisation parameters are totally different concerning VR 3D modelling, as the photogrammetric model must ensure accurate geometric reconstruction for measuring and design and confident radiometric response for monitoring and documentation.

In this study, our goal is to fill the gap mentioned above, proposing a standardised workflow that describes the entire process necessary for the creation of a complete virtual environment, starting from the acquisition phase, passing through data processing and the generation of a three-dimensional model, and finalising it with export into virtual reality for educational purposes. In particular, we will go into depth about the photogrammetric products and highlight the difficulties encountered and the challenges to be overcome to optimise the result in the best possible way regarding quality and processing time. Specifically, we will propose a Virtual Reality outdoor environment recreating a city portion of Turin (Italy), including Carlo Alberto Square, Carignano Square, and Carignano Palace. Due to the historical value of this area, the final goal is to enhance and promote the cultural heritage from an artistic point of view, giving virtual reality a tourist, educational, and didactic purpose. This study is a natural continuation of our previous study [44], in which we had deepened the survey and post-processing phase but in which the implementation in a virtual environment had only been hinted at.

2 Materials and methods

In this chapter, the study area is first introduced, and the methodology adopted is subsequently described in depth in terms of sensors and instrumentation used and how it was integrated to obtain a coherent final product, placing greater emphasis on photogrammetric processing and optimization for VR experience.

2.1 Survey area

The survey area, as shown in Figure 1, extends for about 21,000 m^2, and it includes Carignano Square, Carignano Palace, Carlo Alberto Square, Principe Amedeo Street and Cesare Battisti Street (both in the extension between Carignano Square and Carlo Alberto Square), Carlo Alberto Street (between Carlo Alberto Square and Po Street) and Po Street (between Castello Square and Giambattista Bogino Street). As highlighted in [44], the survey and the high-resolution 3D model generation were performed with the aim of implementing the model in a virtual environment and providing an

interactive tool for the enhancement of the survey area and innovatively advertise citizens and tourist (with a greater level of detail on Palazzo Carignano and the statue of Carlo Alberto). Additional goals were (i) to provide support for the renovation of Carlo Alberto Square, where a station for the new underground city metro line will be realised; (ii) to measure the exact location of the statue dedicated to Carlo Alberto (placed in the homonymous square) which, to preserve its integrity, will be removed during the construction phase and subsequently placed back in the exact location; and (iii) to carry out further analyses concerning possible road interferences between the surface road network (in particular public transport by rail) and the planned underground road.

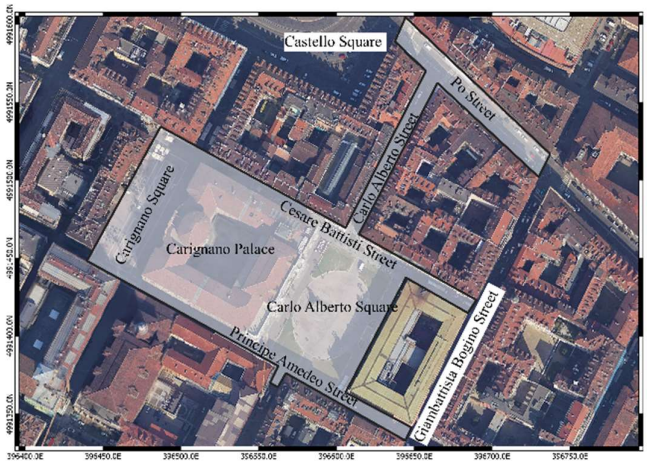

Fig. 1. Study area. EPSG: 32632 [44].

2.2 Materials

Table 1 summarises the sensors used during the survey. A DJI Mini Mavic drone was used to acquire rooftop images; the Riegl VZ-400i laser scanner combined with a Nikon D800E digital camera was used to acquire a dense coloured point cloud to reconstruct the façade geometries with high-resolution; finally, the texture generation in narrow streets was entrusted to a Ricoh Theta V 360° camera.

The topographic phase of the survey, fundamental to defining a standard reference coordinate system for harmonising data, was done with the support of two GNSS receivers (Leica GS14 and Leica GS18), which can receive both the GPS and GLONASS constellations. Also, a Leica Geosystem Image Station total station and prisms were needed to create a polygonal scheme consisting of 14 station points required to georeference the point cloud and the images acquired.

Table 1. Specifications and purpose of the sensors used in the in-situ survey.

Sensor	Characteristics		Purpose
DJI Mini Mavic	Resolution	12 Mpx	Roofs modelling
	Focal Length	4 mm	
	Sensor	1/2.3" CMOS	
RIEGL VZ-400i	Measure technique	ToF	High-detailed reconstruction of facade geometries
	Operative distance	0.5 – 800 m	
	FOV	100°/360°	
	Frequency	100/1200 Hz	
	Accuracy	5/3 mm	
	Size	206 x 308 mm	
NIKON D800E	Resolution	36.2 Mpx	High-detailed texture generation
	Focal Length	60 mm	
	Sensor	35,9 x 24,0 mm CMOS	
Ricoh Theta V	Resolution	12 Mpx (x2)	Texture generation for narrow streets
	Focal Length	1 mm	
	Sensor	1/2.3 CMOS (x2)	
Leica Geosystem Image Station	Angle measure accuracy 1 Hz and V	0.5" (0.15 mgon)	Lidar and images georeferencing
Prisms	Reflective	Yes	
GS14 and GS18 GNSS receivers	GNSS Systems	Multi-frequency	

The VR experience populated with the digital model obtained by the integrated survey was developed using Unity software (Unity Technologies) and deployed in the Meta Quest Pro headset. Unity is a cross-platform game engine able to support Android platforms and customisation for headset VR applications. For this project, the Unity Editor version 2021.3.19f1 has been used.

Meta Quest Pro is a mixed reality (MR) headset developed by Reality Labs, a division of Meta Platforms (formerly Facebook, Inc.). Meta Quest Pro has a resolution of 1832*1920 pixels per eye and a claimed field of view of 106° (horizontal) × 96° (vertical). Table 2 summarises the technological features of the headset.

Table 2. Technical specs of Meta Quest Pro headset.

Sensor component	Characteristics
Operating system	Android
Headset processor	Qualcomm Snapdragon XR2+
Headset dimension	265 mm (length) x 127 mm (height) x 196 mm (width); 722g (weight)
Display	LCD 1800 x 1920 per eye @ 72-90 Hz
Graphics	Adreno 650
Eye-tracking	5x IR imaging sensor (12° FOV)
Headset-tracking	10x imaging camera for 6 DoF SLAM inside-out
Storage memory	Western Digital 256GB NAND flash
Memory (RAM)	Micron Technology 12GB LPDDR5, SDRAM
Controller processor	2x Qualcomm Snapdragon 662, Arm-based, octo-core
Controller tracking	SLAM based on 5 imaging cameras

2.3 Methods

Addressing the integration and optimisation process of 3D photogrammetric products for VR use requires dividing the workflow into two main steps: the execution of the integrated survey, exploiting the methods and techniques of geomatics, and the processing and manipulating of derived data for the VR experience.

Integrated survey. The primary challenge in creating a high-detail, multi-scale, and versatile 3D model of a city lies in finding a compromise between the time spent on data acquisition and data processing. This involves a thorough analysis of the relationship between the area to be surveyed, the time required for surveying, and the careful selection of appropriate sensors to ensure sufficient detail based on the survey objectives. Since the investigated area is characterised by various difficulties (i.e., walking area, flight restrictions, numerous architectural details), it was necessary to plan the survey carefully, in particular: (i) the UAS flight planning required careful consideration of factors such as proper overlap, correct resolution, adequate camera orientation, optimal acquisition time to minimise shadows on facades, and adhering to safety requirements in light of the high tourist activity in the area; (ii) the location, number, and resolution of TLS scans were carefully set to achieve a continuous surface with a consistent point density; (iii) the methodology for acquiring photos for texture reconstruction was constrained by the varying illumination conditions on the facades, adding to the complexity of the task. Consequently, multi-sensor and multi-scale approaches were necessary to fulfil these requirements. Some areas within the surveyed region may have greater significance, such as the Carignano Palace facade and Carlo Alberto's statue, compared to other buildings in Cesare Battisti and Principe Amedeo Streets. Combining optical and LiDAR sensors was considered to balance resolution demands and available resources. Figure 2 shows the workflow followed in this study.

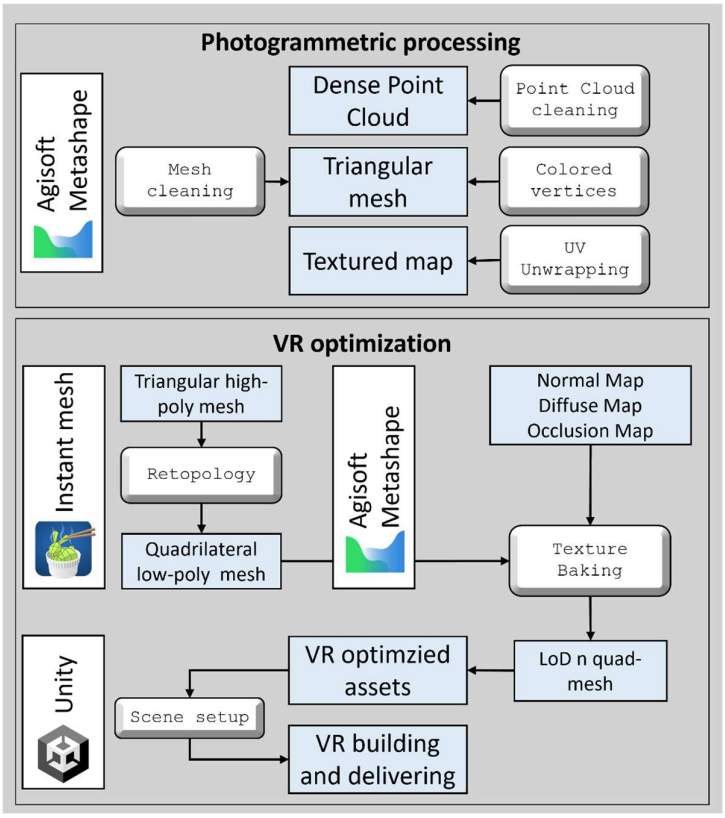

Fig. 2. Workflow of the paper.

In-situ survey. The topographic network needed for georeferencing the acquired data was established comprising four vertices strategically positioned to ensure ample satellite visibility. The coordinates of these vertices were measured using GNSS instrumentation in rapid-static mode (Leica GS14 and Leica GS18), with each point stationed for one hour. The total station was also used to determine the position of specific reference points. This process led to the creating of a polygonal scheme comprising 14 station points. In addition, circular retroreflective artificial targets with a 5 cm radius were strategically placed along the facades and measured using the total station (Figure 3). Due to the reflectivity information, these targets can be easily identified within the LiDAR point clouds. Furthermore, natural points with distinct features were also measured (Figure 3). Altogether, a total of 264 points were measured to enable the georeferencing of both LiDAR and photogrammetric acquisitions.

Fig. 3. Example of a retroreflective target (left) and natural point (right) [44].

42 LiDAR scans were needed to obtain a complete point cloud of the area. The acquisition frequency was set equal to 600 kHz (equivalent to an acquisition speed of approximately 250,000 points per second), and the angular resolution was carefully set to ensure a point was captured approximately every 6 mm at a distance of 10 m from the station point. The combination with the NIKON digital camera allows the assigning of colour information to the point cloud.

The DJI Mini Mavic UAS was employed to conduct a high-resolution survey of the roofs with both nadiral and oblique views, ensuring complete coverage of the entire area and gathering detailed information about ceilings and walls for integration into the LiDAR model. The flight height was set at 40 meters above the ground to achieve a Ground Sample Distance (GSD) of less than 2 centimetres. In total, 1,599 images were acquired during the survey.

Digital images from the ground at a high resolution, particularly for the facades of buildings, were needed to guarantee the high quality of the 3D model textures. To achieve this, approximately 400 images were captured using the NIKON D800E camera, employing a nadiral angle as close to the facades as possible. This data collection focused on the buildings facing the two squares and included Carlo Alberto's statue.

Data processing. The GNSS data was post-processed in relative mode using the data from the GNSS Interregional Positioning Service (SPIN 3 GNSS) permanent network within the Leica Geo Office (LGO) software. For the reference cartographic grid, the UTM-WGS84 projection was chosen. However, a conventional local isometric system (MTL2 ISO250) was adopted to avoid typical cartographic representation distortions. Regarding the GNSS heights, they were measured above the ellipsoid and subsequently converted into orthometric heights using the "ITALGEO2005" geoid undulations model provided by the Italian Geographic Military Institute (Istituto Geografico Militare, IGM).

The compensation of the network using MicroSurvey Star*Net software resulted in an estimated coordinates standard deviation (RMSE) of less than 1 cm. To georeference the UAS photogrammetric products, specific Ground Control Points (GCPs) were marked and measured using GNSS receivers. The NRTK-GNSS method, with real-time centimetric precision, was utilised for these measurements, thanks to the corrections received from a local network of permanent SPIN 3 GNSS service stations. Photographic points easily identifiable in drone images, such as pavement edges, road markings, and corners of maintenance holes, were used for this purpose. The field-acquired

coordinates were then exported using LGO software to obtain the final coordinates in the national geodetic reference system ETRF2000 with 32N UTM projection.

The data collected from the laser scanner and the topographic survey underwent post-processing using Riegl's RiScan Pro software. To achieve relative registration of scans and georeferencing in the absolute coordinate system, retroreflective targets were employed. Each target was associated with the actual coordinates measured during the topographic survey. The initial registration of scanning positions was performed semi-automatically based on voxel analysis. The identified targets in the scans served as additional observations for relative registration between scans, enhancing the accuracy of scan location estimation. Subsequently, further optimisation was carried out for relative registration between scans and georeferencing. This optimisation considered all the acquisitions made and available targets, GNSS measurements, and altitude measurements from the inertial platform. During this optimisation phase, flat patches representing the environment's surfaces were extracted from each scan. An iterative process was then used to seek homologous planes between different scans and establish their correspondence. This procedure permanently corrected the relative positions between the scans while simultaneously estimating the absolute position of the scans. As raw scan data typically lacks radiometric information, the Riegl VZ-400i relies on its digital camera, mounted on top of the instrument, to attribute colour information to each pixel. However, for this process, calibration data of the camera is required. Although the internal calibration of the camera is known in advance, recalibration of external parameters is necessary whenever the camera is mounted on the instrument. These external calibration parameters can be estimated by matching common points between the scan and the images acquired at the same station point. After this calibration process, each point in the cloud is coloured according to the corresponding pixels of the assigned calibrated images. Finally, the georeferenced point cloud, as shown in Figure 4, was exported in .las format.

Fig. 4. Registered point cloud [44].

Regarding the photogrammetric procedure, all the images were processed using Agisoft Metashape Professional, a commercial software based on Structure-from-Motion (SfM) techniques. In the initial phase, a fully automatic alignment between images was performed to reconstruct the relative positions. A photogrammetric camera calibration was conducted to determine essential parameters such as the correct focal

distance, main point position, sensor distortions, and radial and tangential optics. To generate a dense point cloud, the algorithm computes depth maps, which are derived through the application of dense stereo-matching techniques. These depth maps are computed by considering the relative exterior and interior orientation parameters of overlapping image pairs, estimated through bundle adjustment. Multiple depth maps are generated pairwise for each camera and subsequently merged to form a consolidated depth map. This merging process leverages the surplus information in the overlapping regions to eliminate inaccuracies in in-depth measurements. To improve accuracy, 10 Ground Control Points (GCPs) were manually selected and marked on various images. Among these, five were chosen as checkpoints (CPs) for validation purposes. The block compensation process resolved the residuals on GCPs and CPs, leading to the refinement of the camera calibration. As a result, the images were accurately positioned and oriented in space, ensuring alignment with the isometric coordinate system. To create the textured 3D model of the study area, the dense point clouds obtained from the procedures mentioned above were integrated, resulting in a comprehensive model of the urban environment. The dense cloud was then divided into seven distinct portions to be processed separately. A full resolution was chosen for the main buildings: Carignano Palace, the National Library, and Carlo Alberto's statue. However, 1 to 5 centimetres resolution was used for buildings facing the square and terrain. The dense georeferenced models served as a basis for generating the 3D mesh, focusing on achieving a high-quality level maximising the number of triangles for optimal representation.

Photogrammetric processing and optimisation for VR experience. This section is dedicated to applying photogrammetry in VR development, outlining the steps to bridge the divide between digital reconstruction and the optimisation necessary for virtual reality. The optimisation of a VR project is a multifaceted endeavour that requires collaboration between developers and artists to replicate intricate real-world scenarios within a virtual environment faithfully. The proposed workflow focuses on the photogrammetric processing and optimisation aspects of geometric and radiometric reconstruction. Other parameters unrelated to the photogrammetric procedure are not considered for this work. Shadowing, aliasing, mipmapping, and other optimisation aspect are strictly related to the digital artist choice and requires specific competence related to the game design and VR tools developers.

The proposed workflow can be synthesised in the following bullet points list and described in Figure 2:

- Dense polygonal mesh model generation and detailed texturisation in Metashape;
- Mesh refinement in Metashape;
- Mesh Decimation, Retopology and Texture baking;
- Scene setup, controller input mapping, and mesh collider in Unity.

Dense polygonal mesh model generation and detailed texturisation in Metashape. Using photogrammetry, it is possible to create an extremely accurate model of real-world objects. Typically, a photogrammetric survey provides geometrically and

radiometrically reliable data for documentation, modelling, simulation and analysis. Creating digital assets for VR using photogrammetry follows different requirements and constraints instead. Usually, in engineering, architecture, archaeology and geoscience, the mesh generation is a step that can be avoided, opting for the direct use of dense cloud as the basis for DEM and orthomosaic. In the case of digital assets for VR experiences, the dense polygonal mesh generation is mandatory and will be used to populate the scene.

Mesh refinement in Metashape. The photo-consistent mesh refinement operation allows for iteratively recovering details (bas-relief or ditch) using a variational refinement method described in [51]. The result is a mesh that automatically and adaptively adjusts to image resolution. The required parameters are (i) quality, (ii) iterations and (iii) smoothness. The first defines the image downsizing with the same principle of image alignment; the second describes the number of iterations of the variational method; the third represents a classical smoothness threshold. Starting from the high-resolution polygonal mesh, the refined mesh is produced with an increased balance between noise suppression and feature recovery.

Mesh Decimation, Retopology and Texture baking. The refined mesh obtained in the previous step is now the most detailed and optimised mesh obtainable from the photogrammetric workflow. Unfortunately, dense polygonal meshes are unsuitable and not remarkably manageable for VR software due to the data size and the complexity. Moreover, while most photogrammetric applications require only access to the 3D model and inspect it through viewer software, VR experiences require simulating a player moving closer to the object.

Therefore, it is common to produce several versions of the model, from low-poly to high-poly, interchanged during the user's movement. It is evident that creating several models of the same asset at different resolutions is a time-consuming process. Therefore, several methods have been developed to reshape the model, decreasing its size while keeping its complexity and details. These methods consist of reducing the number of polygons without losing details and maintaining the geometrical information of the original mesh. Mesh decimation and retopology methods are the most valuable methods to solve this task. In this work, both methods have been applied to evaluate which one allows the better result in terms of cost/benefits. Mesh decimation is utilised to diminish the complexity of a model by substituting a high-resolution mesh with a lower-resolution alternative while preserving the capability to represent the object's geometry accurately. Metashape software implements a specific tool that requires defining the target polygon number.

While Mesh Decimation applies a decimation algorithm to reduce the number of triangular faces, the Retopology method operates on a mesh's topology to obtain a cleaner and simplified polygonal mesh composed of more easily manageable shapes (usually quadrangular faces). This automatic technique is implemented in several software programs, which allows the conversion of the triangular mesh into a quad-mesh that guarantees an effective interaction with texturing and animation software,

contributing to minimising the dimensions of the model and their rendering time. This conversion has been demanded by the Instant Meshes software [52].

Regardless of the method used to simplify the high-resolution mesh, the next step in the workflow is baking the diffusion maps and normal maps. Baking involves transferring the textures generated from the high-resolution model to all other low-poly models. This task requires consistency between the UV space of the texture and the 3D space of the model. Keeping as much quality as possible before the baking process is fundamental. Diffuse maps are texture maps that show the colour information reprojected by the images to the model. Normal maps complement diffuse maps and are texture maps that can give the appearance of additional geometry or details on flat surfaces and are used to make a low-poly mesh appear high-poly. Both maps are required to provide detailed material texture with depth perception.

In this workflow, we start by creating the high-resolution diffuse and normal texture in Metashape software. Then, we independently upload the quad-mesh or decimated mesh and bake the diffuse and normal maps. The procedure can be repeated for each Level of Detail designed for the VR experience.

Scene setup, controller input mapping, and mesh collider in Unity. Once all meshes are obtained in different LoD, converting them into scene-ready meshes with the correct orientation, pivot and uv set is required. This process can be done directly in Metashape during the export operations. The software chosen to develop the VR experience is Unity (references). The virtual set in Unity is called scene, and it hosts all the elements composing the project. A basic VR scene must contain an XR Origin, which defines the tracking data reference system. It allows the alignment of the tracking data acquired by the headset and aligned in a real-world reference system to the reference system defined in the unity environment. The XR Origin contains children's objects like the user's headset (Main Camera object) or the hand-held device (XR controller). Fundamental Objects are also the locomotion system to move around the scene and an input system to interact with the 3D objects. The test VR experience proposed in this work is a virtual tour without interaction of the portion of Turin city obtained with the integrated survey. Therefore, it was required to map the user's movement (walking, climbing) and to set up the mesh collider. Once the VR environment was correctly populated, the developed application was built and delivered for Android.

3 Results and discussion

The results of the study conducted in this paper are shown and discussed in this chapter. Please refer to [44] for further information regarding the integrated survey, while the results regarding the integration of VR are described in an exhaustive manner. For illustrative purposes (also on a visual level) we have chosen to show the analyses regarding Carlo Alberto’s statue.

3.1 Integrated survey

The survey's final result consists of an integrated point cloud with about 5 billion points containing radiometric and LiDAR intensity information. The following checks were performed to validate the registration and georeferencing procedure accuracy concerning LiDAR and photogrammetric.

Concerning the LiDAR cloud, the accuracy of the scanning registering procedure was evaluated through the estimated deviations between the targets identified in the scans and their known coordinates (measured with the total station). The results of this analysis are shown in Table 3, and an RMSE of approximately 4 cm between points can be observed.

Table 3. Estimated residues between targets identified in scans and known coordinates [44].

	dX [m]	dY [m]	dZ [m]	Dist. [m]
Min	-0.08946	-0.07368	-0.13934	0.001692
Max	0.088073	0.105132	0.087002	0.156342
Mean	0.000671	-0.00033	0.002233	0.032446
RMSE	0.018788	0.020184	0.029044	0.040049

Concerning the photogrammetric survey, the final output was validated by analysing the estimated residuals on the control points and checkpoints (Table 4). The acquired measurements reflect the general geometric precision of the photogrammetric models produced using the Structure from Motion (SfM) technique. The point cloud comprehensively portrays the structures' rooftops and vertical dimensions. However, it lacks density in the lower sections of the buildings within narrow streets, where distortion is apparent even in oblique images. Consequently, a choice was made to reduce the quantity of ground-based photogrammetric data collection in these regions and instead utilise LiDAR data for 3D reconstruction.

Table 4. Results of the photogrammetric block adjustment [44].

	X (m)	Y (m)	Z (m)
RMSE	0.016	0.013	0.027
10 GCPs error	0.019	0.017	0.022
5 CPs error	0.016	0.013	0.027

It is important to remember that extraneous information is frequently captured during the data collection procedure. This issue is particularly pronounced in urban settings, where the raw scan data becomes distorted due to various interferences, such as highly reflective surfaces (like headlights and water), transparent materials (such as glass), and moving elements (for instance, people and vehicles). Among these factors, moving elements posed the most significant source of disturbance: pedestrians moving through the area during the measurement process led to disruptions in the point clouds. Initially, all individuals and unneeded objects within the scene were manually eliminated from the point cloud. Subsequently, the remaining noise was mitigated through the Statistical Outlier Removal (SOR) tool, which identifies and removes points likely

unrelated to the modelled surfaces. This technique involves computing the average distance between each point and its neighbouring points (using six neighbours in this study). Points that exceed the average distance, multiplied by a coefficient derived from the standard deviation (0.5 in this instance), are filtered out. Despite these filtering procedures, managing the extensive point cloud still presented challenges and demanded substantial time due to its high point density. The point cloud was then systematically downsampled to facilitate the analysis and modelling tasks based on the area of interest.

3.2 VR application

Among the seven portions processed in the previous section, Carlo Alberto's statue has been selected for the following parametrisation analysis. The dense point cloud obtained by the photogrammetric processing was a three-band 8-bit cloud composed of 47,405,562 points obtained by setting the quality parameter to "high", which downscale the images by a factor of 4. The processing time was 5 hours and 32 minutes, producing a file of 706.05 MB. The dense cloud was used to compute the polygonal mesh model, obtaining 9,481,010 faces and 4,746,181 vertices in 3 hours and 40 minutes. The vertices were colourised again with three bands of 8-unit colour information, and the entire mesh was texturised with a high-quality 8,192 x 8,192, 4 bands, uint8 texture. The obtained 3D mesh focused on achieving a high-quality level, maximising the number of triangles for optimal representation.

Fig. 5. Dense point cloud (left), dense triangulated mesh (middle) and textured model (right) of Carlo Alberto's statue.

Further processing steps have been applied in Metashape to obtain the mesh required for subsequent steps described in Figure 5. Firstly, using the Metashape tool "refine mesh", an iterative process was used to recover surface details like bas-relief or ditch. Then, the texture was computed again on the refined model at the highest quality. The resulting 3D mesh was composed of 11,736,644 faces, 5,873,952 vertices and texturised with an 8,192 x 8,192, 4 bands, 8-unit image. The required processing time for refinement and texturing was 20 minutes, and the result was stored in 629.80 MB space.

The result was exported both in .obj and in.ply file format, with the texture stored in .jpg (Figure 6).

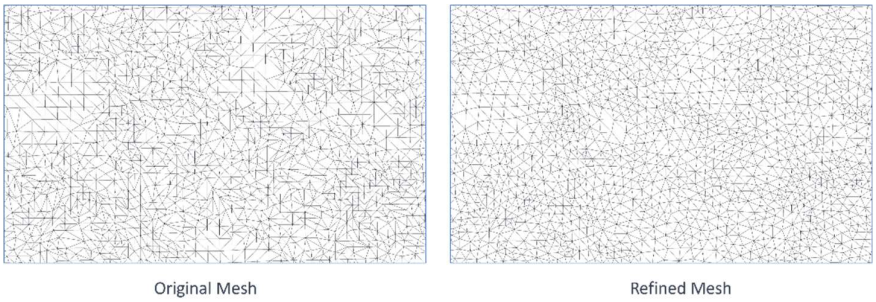

Fig. 6. Triangular Mesh structure: Original (left) and Refined (right).

Retopology. The refined 3D polygonal mesh model was imported into Instant Mesh software to perform an automatic retopology to minimise the model dimensions by transforming the mesh into a quadrangular faces model (Figure 7). In the case of vertex singularities detected by the algorithm, a semi-automatic brush tool has been used to change the polygons' direction. Figure 8 shows different Levels of Details of the quadratic mesh, among which LoD 4 was selected to preserve the details of the original mesh, although obtaining a model light enough to be managed in VR software.

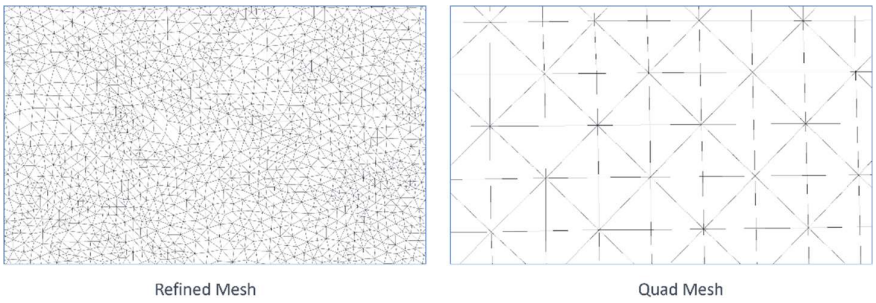

Fig. 7. Mesh obtained in Instant Mesh software at LoD 4.

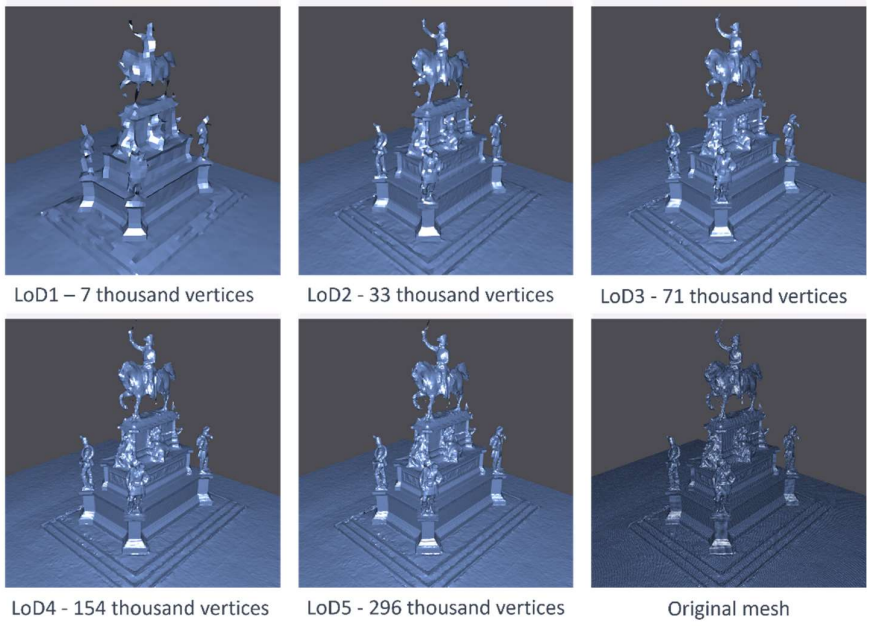

Fig. 8. Different LoD obtained in Instant Mesh software.

The normal map obtained in Metashape can now be applied to the low-poly model obtained by retopology. In VR software, this gives the appearance of fine surface details without the need for a high polygon count.

Unity. After setting up the VR scene, the controller input mapping and the mesh collider, the app was tested on the Meta Quest Pro device and the performance parameters were analysed. We used an XPS 15 7590 Windows 11 computer (Intel(R) Core(TM) i9-9980HK CPU @ 2.40GHz, NVIDIA GeForce GTX 1650, 32 GB RAM) to run the VR experience and display it on a Meta Quest Pro (Model: DK94EC) through a wire connection (QuestLink) with a fresh rate of 90 Hz. These are reported in Table 5, while Figure 9 shows an example of the interactive VR environment scene to explore Carlo Alberto's statue.

Table 3. Performance of the VR application.

Graphics	22.4 FPS
CPU:	Main 44.7 ms render thread 21.0 ms
Batches:	125
Tris:	1.3 M
Vert:	708.8k
Screen:	890x451 – 4.6 MB
SetPass calls:	125

Also, an entire scene of the integrated survey has been set up. The workflow followed was the same, but in order to maintain the optimal user experience, the Carigano Palace and the National Library were further divided in function of the floors. This resulted in generating a mesh with the maximum number of triangles for the lower portions of the facades while limiting the model's quality on the subsequent floors.

Fig. 9. VR environment scene of Carlo Alberto's statue and Palazzo Carignano.

4 Conclusions

In this work, we propose a comprehensive workflow for building a VR application using digital assets obtained by optimising photogrammetric and LiDAR 3D models. The workflow is applied on a high-detail, multi-scale 3D model of a city, which has been obtained through a multi-sensor integrated survey, exploiting UAV photogrammetry and TLS. VR is an innovative technology widely used to create faithful digital replicas of heritage sites and create immersive and engaging experiences. Historical city centres are also resistant to modification through the construction of modern infrastructure due to the presence of historical budlings and the opposition of many inner-city communities; therefore, VR applications can show how the construction of new infrastructure can change the neighbourhood's appearance while respecting the historical heritage. If in our previous study we limited ourselves to describing the data collection and processing phase in order to obtain a 3D model, in this study we also examined the precautions to be followed to implement this model correctly and efficiently in a virtual environment.

The importance of integrating geomatic techniques has emerged once again, and it will play a fundamental role in the following years since virtual reality is experiencing a quick increase of interest from private and public bodies for the innovative advertising and tourism of cultural heritage or museums. Optimising procedures to produce digital

assets is fundamental to exploit the 3D survey potentiality without stressing headset hardware.

The results are excellent; however, further tests are needed to validate the proposed workflow in other environments (e.g., indoor, natural areas, etc.) and to describe and solve all the difficulties that may be encountered as the scenarios vary.

References

1. Klamer, A.: The Value of Cultural Heritage. In: Hutter, M. and Rizzo, I. (eds.) Economic Perspectives on Cultural Heritage. pp. 74–87. Palgrave Macmillan UK, London (1997). https://doi.org/10.1007/978-1-349-25824-6_5.
2. Ogden, S.: Understanding, Respect, and Collaboration in Cultural Heritage Preservation: A Conservator's Developing Perspective. Library Trends. 56, 275–287 (2007). https://doi.org/10.1353/lib.2007.0056.
3. Lussetyowati, T.: Preservation and Conservation through Cultural Heritage Tourism. Case Study: Musi Riverside Palembang. Procedia - Social and Behavioral Sciences. 184, 401–406 (2015). https://doi.org/10.1016/j.sbspro.2015.05.109.
4. Araoz, G.F.: Preserving heritage places under a new paradigm. Journal of Cultural Heritage Management and Sustainable Development. 1, 55–60 (2011). https://doi.org/10.1108/20441261111129933.
5. Gaitatzes, A., Christopoulos, D., Roussou, M.: Reviving the past: cultural heritage meets virtual reality. In: Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage. pp. 103–110. ACM, Glyfada Greece (2001). https://doi.org/10.1145/584993.585011.
6. Bekele, M.K., Pierdicca, R., Frontoni, E., Malinverni, E.S., Gain, J.: A Survey of Augmented, Virtual, and Mixed Reality for Cultural Heritage. J. Comput. Cult. Herit. 11, 1–36 (2018). https://doi.org/10.1145/3145534.
7. Skublewska-Paszkowska, M., Milosz, M., Powroznik, P., Lukasik, E.: 3D technologies for intangible cultural heritage preservation—literature review for selected databases. Herit Sci. 10, 3 (2022). https://doi.org/10.1186/s40494-021-00633-x.
8. Cantatore, E., Lasorella, M., Fatiguso, F.: VIRTUAL REALITY TO SUPPORT TECHNICAL KNOWLEDGE IN CULTURAL HERITAGE. THE CASE STUDY OF CRYPTOPORTICUS IN THE ARCHAEOLOGICAL SITE OF EGNATIA (ITALY). Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLIV-M-1–2020, 465–472 (2020). https://doi.org/10.5194/isprs-archives-XLIV-M-1-2020-465-2020.
9. Younes, G., Kahil, R., Jallad, M., Asmar, D., Elhajj, I., Turkiyyah, G., Al-Harithy, H.: Virtual and augmented reality for rich interaction with cultural heritage sites: A case study from the Roman Theater at Byblos. Digital Applications in Archaeology and Cultural Heritage. 5, 1–9 (2017). https://doi.org/10.1016/j.daach.2017.03.002.
10. Canciani, M., Conigliaro, E., Del Grasso, M., Papalini, P., Saccone, M.: 3D SURVEY AND AUGMENTED REALITY FOR CULTURAL HERITAGE. THE CASE STUDY OF AURELIAN WALL AT CASTRA PRAETORIA IN ROME.

Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLI-B5, 931–937 (2016). https://doi.org/10.5194/isprs-archives-XLI-B5-931-2016.
11. Economou, T., Vosinakis, S.: Mobile Augmented Reality Games As An Engaging Tool For Cultural Heritage Dissemination: A Case Study. (2018). https://doi.org/10.5281/ZENODO.1214569.
12. Valtolina, S., Franzoni, S., Mazzoleni, P., Bertino, E.: Dissemination of Cultural Heritage Content through Virtual Reality and Multimedia Techniques: A Case Study. In: 11th International Multimedia Modelling Conference. pp. 214–221. IEEE, Honolulu, HI, USA (2005). https://doi.org/10.1109/MMMC.2005.36.
13. Mortara, M., Catalano, C.E., Bellotti, F., Fiucci, G., Houry-Panchetti, M., Petridis, P.: Learning cultural heritage by serious games. Journal of Cultural Heritage. 15, 318–325 (2014). https://doi.org/10.1016/j.culher.2013.04.004.
14. Sun, Q., Wei, L.-Y., Kaufman, A.: Mapping virtual and physical reality. ACM Trans. Graph. 35, 1–12 (2016). https://doi.org/10.1145/2897824.2925883.
15. LaValle, S.M.: Virtual reality. Cambridge University Press (2023).
16. Bowman, D.A., McMahan, R.P.: Virtual Reality: How Much Immersion Is Enough? Computer. 40, 36–43 (2007). https://doi.org/10.1109/MC.2007.257.
17. Pensieri, C., Pennacchini, M.: Overview: Virtual Reality in Medicine. JVWR. 7, (2014). https://doi.org/10.4101/jvwr.v7i1.6364.
18. Riener, R., Harders, M.: Introduction to Virtual Reality in Medicine. In: Virtual Reality in Medicine. pp. 1–12. Springer London, London (2012). https://doi.org/10.1007/978-1-4471-4011-5_1.
19. Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., Wang, X., Jiang, Q.: Application of virtual reality technology in clinical medicine. Am J Transl Res. 9, 3867–3880 (2017).
20. McCloy, R., Stone, R.: Science, medicine, and the future: Virtual reality in surgery. BMJ. 323, 912–915 (2001). https://doi.org/10.1136/bmj.323.7318.912.
21. Liu, X., Zhang, J., Hou, G., Wang, Z.: Virtual Reality and Its Application in Military. IOP Conf. Ser.: Earth Environ. Sci. 170, 032155 (2018). https://doi.org/10.1088/1755-1315/170/3/032155.
22. Tzanavari, A., Tsapatsoulis, N. eds: Affective, Interactive and Cognitive Methods for E-Learning Design: Creating an Optimal Education Experience. IGI Global (2010). https://doi.org/10.4018/978-1-60566-940-3.
23. Häkkilä, J., Hannula, P., Luiro, E., Launne, E., Mustonen, S., Westerlund, T., Colley, A.: Visiting a virtual graveyard: designing virtual reality cultural heritage experiences. In: Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia. pp. 1–4. ACM, Pisa Italy (2019). https://doi.org/10.1145/3365610.3368425.
24. Donghui, C., Guanfa, L., Wensheng, Z., Qiyuan, L., Shuping, B., Xiaokang, L.: Virtual reality technology applied in digitalization of cultural heritage. Cluster Comput. 22, 10063–10074 (2019). https://doi.org/10.1007/s10586-017-1071-5.
25. Giangreco, I., Sauter, L., Parian, M.A., Gasser, R., Heller, S., Rossetto, L., Schuldt, H.: VIRTUE: a virtual reality museum Experience. In: Proceedings of the 24th International Conference on Intelligent User Interfaces: Companion. pp. 119–120.

ACM, Marina del Ray California (2019). https://doi.org/10.1145/3308557.3308706.
26. Lee, H., Jung, T.H., Tom Dieck, M.C., Chung, N.: Experiencing immersive virtual reality in museums. Information & Management. 57, 103229 (2020). https://doi.org/10.1016/j.im.2019.103229.
27. Vincke, S., De Lima Hernandez, R., Bassier, M., Vergauwen, M.: IMMERSIVE VISUALISATION OF CONSTRUCTION SITE POINT CLOUD DATA, MESHES AND BIM MODELS IN A VR ENVIRONMENT USING A GAMING ENGINE. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-5/W2, 77–83 (2019). https://doi.org/10.5194/isprs-archives-XLII-5-W2-77-2019.
28. Dhanda, A., Reina Ortiz, M., Weigert, A., Paladini, A., Min, A., Gyi, M., Su, S., Fai, S., Santana Quintero, M.: RECREATING CULTURAL HERITAGE ENVIRONMENTS FOR VR USING PHOTOGRAMMETRY. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XLII-2/W9, 305–310 (2019). https://doi.org/10.5194/isprs-archives-XLII-2-W9-305-2019.
29. Mikhail, E.M., Bethel, J.S., McGlone, J.C.: Introduction to modern photogrammetry. New York:Wiley (2001).
30. Zhao, S., Kang, F., Li, J.: Concrete dam damage detection and localisation based on YOLOv5s-HSC and photogrammetric 3D reconstruction. Automation in Construction. 143, 104555 (2022). https://doi.org/10.1016/j.autcon.2022.104555.
31. Maas, H.-G., Hampel, U.: Photogrammetric Techniques in Civil Engineering Material Testing and Structure Monitoring. photogramm eng remote sensing. 72, 39–45 (2006). https://doi.org/10.14358/PERS.72.1.39.
32. Bemis, S.P., Micklethwaite, S., Turner, D., James, M.R., Akciz, S., Thiele, S.T., Bangash, H.A.: Ground-based and UAV-Based photogrammetry: A multi-scale, high-resolution mapping tool for structural geology and paleoseismology. Journal of Structural Geology. 69, 163–178 (2014). https://doi.org/10.1016/j.jsg.2014.10.007.
33. Remondino, F., El-Hakim, S.: Image-based 3D Modelling: A Review: Image-based 3D modelling: a review. The Photogrammetric Record. 21, 269–291 (2006). https://doi.org/10.1111/j.1477-9730.2006.00383.x.
34. Kwiatek, K., Tokarczyk, R.: Immersive photogrammetry in 3D modelling: Fotogrametria immersyjna w modelowaniu 3D. geom. 9, 51 (2015). https://doi.org/10.7494/geom.2015.9.2.51.
35. Singh, S.P., Jain, K., Mandla, V.R.: VIRTUAL 3D CITY MODELING: TECHNIQUES AND APPLICATIONS. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-2/W2, 73–91 (2013). https://doi.org/10.5194/isprsarchives-XL-2-W2-73-2013.
36. Zhu, Q., Hu, M., Zhang, Y., Du, Z.: Research and practice in three-dimensional city modeling. Geo-spatial Information Science. 12, 18–24 (2009). https://doi.org/10.1007/s11806-009-0195-z.
37. Sato, T., Kanbara, M., Yokoya, N.: Outdoor scene reconstruction from multiple image sequences captured by a hand-held video camera. In: Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent

Systems, MFI2003. pp. 113–118. IEEE, Tokyo, Japan (2003). https://doi.org/10.1109/MFI-2003.2003.1232642.
38. Pepe, M., Fregonese, L., Crocetto, N.: Use of SfM-MVS approach to nadir and oblique images generated throught aerial cameras to build 2.5D map and 3D models in urban areas. Geocarto International. 37, 120–141 (2022). https://doi.org/10.1080/10106049.2019.1700558.
39. Smith, M.W., Carrivick, J.L., Quincey, D.J.: Structure from motion photogrammetry in physical geography. Progress in Physical Geography: Earth and Environment. 40, 247–275 (2016). https://doi.org/10.1177/0309133315615805.
40. Tse, R.O.C., Gold, C., Kidner, D.: 3D City Modelling from LIDAR Data. In: van Oosterom, P., Zlatanova, S., Penninga, F., and Fendel, E.M. (eds.) Advances in 3D Geoinformation Systems. pp. 161–175. Springer Berlin Heidelberg, Berlin, Heidelberg (2008). https://doi.org/10.1007/978-3-540-72135-2_10.
41. Wang, C., Wen, C., Dai, Y., Yu, S., Liu, M.: Urban 3D modeling with mobile laser scanning: a review. Virtual Reality & Intelligent Hardware. 2, 175–212 (2020). https://doi.org/10.1016/j.vrih.2020.05.003.
42. Ramos, M.M., Remondino, F.: Data fusion in Cultural Heritage – A Review. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-5/W7, 359–363 (2015). https://doi.org/10.5194/isprsarchives-XL-5-W7-359-2015.
43. Sahin, C., Alkis, A., Ergun, B., Kulur, S., Batuk, F., Kilic, A.: Producing 3D city model with the combined photogrammetric and laser scanner data in the example of Taksim Cumhuriyet square. Optics and Lasers in Engineering. 50, 1844–1853 (2012). https://doi.org/10.1016/j.optlaseng.2012.05.019.
44. Grasso, N., Spadavecchia, C., Di Pietra, V., Belcore, E.: LiDAR and SfM-MVS Integrated Approach to Build a Highly Detailed 3D Virtual Model of Urban Areas: In: Proceedings of the 9th International Conference on Geographical Information Systems Theory, Applications and Management. pp. 128–135. SCITEPRESS - Science and Technology Publications, Prague, Czech Republic (2023). https://doi.org/10.5220/0011760800003473.
45. Chang, E., Seo, D., Kim, H.T., Yoo, B.: An Integrated Model of Cybersickness: Understanding User's Discomfort in Virtual Reality. JOK. 45, 251–279 (2018). https://doi.org/10.5626/JOK.2018.45.3.251.
46. Zhdanov, A.D., Zhdanov, D.D., Bogdanov, N.N., Potemin, I.S., Galaktionov, V.A., Sorokin, M.I.: Discomfort of Visual Perception in Virtual and Mixed Reality Systems. Program Comput Soft. 45, 147–155 (2019). https://doi.org/10.1134/S036176881904011X.
47. Obradović, M., Vasiljević, I., Đurić, I., Kićanović, J., Stojaković, V., Obradović, R.: Virtual Reality Models Based on Photogrammetric Surveys—A Case Study of the Iconostasis of the Serbian Orthodox Cathedral Church of Saint Nicholas in Sremski Karlovci (Serbia). Applied Sciences. 10, 2743 (2020). https://doi.org/10.3390/app10082743.
48. PortalÃ©s, C., Lerma, J.L., PÃ©rez, C.: Photogrammetry and augmented reality for cultural heritage applications. The Photogrammetric Record. 24, 316–331 (2009). https://doi.org/10.1111/j.1477-9730.2009.00549.x.

49. Hilfert, T., König, M.: Low-cost virtual reality environment for engineering and construction. Vis. in Eng. 4, 2 (2016). https://doi.org/10.1186/s40327-015-0031-5.
50. Placitelli, A.P., Gallo, L.: Low-Cost Augmented Reality Systems via 3D Point Cloud Sensors. In: 2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems. pp. 188–192. IEEE, Dijon, France (2011). https://doi.org/10.1109/SITIS.2011.43.
51. Vu, H., Keriven, R., Labatut, P., Pons, J.-P.: Towards high-resolution large-scale multi-view stereo. Proc. IEEE Conf. CVPR'09, 1430–1437 (2009).
52. Jakob, W., Tarini, M., Panozzo, D., Sorkine-Hornung, O.: Instant field-aligned meshes. ACM Trans. Graph. 34, 1–15 (2015). https://doi.org/10.1145/2816795.2818078.