International Journal of GEOMATE, Aug., 2025 Vol.29, Issue 132, pp.158-171
ISSN: 2186-2982 (P), 2186-2990 (O), Japan, DOI: https://doi.org/10.21660/2025.132.4883
Geotechnique, Construction Materials and Environment

INSPECTION OF HISTORICAL MONUMENT USING FISH TANK VIRTUAL REALITY: A CASE STUDY OF WAT SI PHICHIT KIRATI KANLAYARAM, SUKHOTHAI, THAILAND

Vishal Jangid[1], *Sirisilp Kongsilp[1], Apichat Buatik[2] and Phromphat Thansirichaisree[2]

[1]Thammasat School of Engineering, Thammasat University, Rangsit, Thailand
[2]Thammasat Research Unit in Infrastructure Inspection and Monitoring, Repair and Strengthening (IIMRS), Faculty of Engineering, Thammasat School of Engineering, Thammasat University, Rangsit, Thailand

*Corresponding Author, Received: 16 Jan. 2025, Revised: 2025, Accepted: 02 April 2025

ABSTRACT: Preserving historical monuments is vital to safeguarding cultural heritage, yet traditional inspection methods often face challenges such as inaccessibility, subjective evaluation, and insufficient documentation of structural defects. Advancements in virtual reality and photogrammetry offer new opportunities to enhance inspection workflows. This study presents a Fish Tank Virtual Reality application to address the limitations of conventional techniques. The system employs UAV-based image acquisition at a 1 mm/pixel ground sampling distance, ensuring the precise capture of structural details. The application integrates a high-resolution photogrammetry model of a stupa, allowing users to navigate the digital environment and annotate structural defects, such as cracks and erosion. Unlike immersive virtual reality, the FTVR approach ensures accessibility and ease of use while maintaining spatial accuracy. The system fills a critical gap by providing an efficient method for visualizing and documenting structural conditions, supporting informed decision-making in monument preservation. A case study on Wat Si Phichit Kirati Kanlayaram, Sukhothai, Thailand, a stupa, demonstrates the practical utility of the system in identifying and annotating defects, offering a scalable and adaptable application for inspecting diverse historical structures. This research highlights the potential of integrating advanced visualization technologies with photogrammetry to streamline monument inspection workflows, enhancing both the efficiency and reliability of heritage conservation efforts. Additionally, it demonstrates how the FTVR inspection method reduces reliance on physical scaffolding, offering safer and more thorough remote assessments of heritage structures, while paving the way for further innovations in digital heritage preservation.

Keywords: Fish Tank Virtual Reality, Structural Inspection, Heritage Preservation, UAV Photogrammetry, 3D Model Reconstruction

1. INTRODUCTION

Preserving historical monuments is essential to maintaining cultural heritage and ensuring their structural integrity for future generations. These structures are subjected to environmental wear, aging materials, and human activity and require regular inspections to identify structural issues. Traditional inspection methods, such as manual surveys and physical documentation, are often limited by accessibility challenges, subjective evaluations, and inadequate record-keeping. For instance, inspectors may find it difficult or unsafe to access fragile or intricate areas and traditional methods may fail to provide detailed documentation that supports long-term monitoring or maintenance planning. These limitations necessitate the development of innovative approaches to improve the efficiency, precision, and safety of inspection processes. Recent advancements in digital technologies, including photogrammetry and Virtual Reality (VR), offer promising solutions to these challenges.

Photogrammetry, a process that converts high-resolution images into detailed 3D models, has gained widespread use in documenting complex structures with remarkable accuracy. These models serve as a digital twin of the structure, capturing intricate architectural details and providing a foundation for thorough analysis. When integrated with VR, these 3D models enable immersive and interactive exploration, enhancing the inspector's ability to identify and annotate structural defects effectively.

Among VR approaches, Fish Tank Virtual Reality (FTVR) stands out as an accessible and efficient option for inspection tasks. Unlike fully immersive VR, which requires specialized headsets, FTVR operates on desktop-based systems, providing a virtual window into the 3D environment. This setup is particularly advantageous for professionals who require spatial interaction and detailed visualization without extensive hardware setups. FTVR bridges the gap between traditional inspection techniques and modern visualization technologies, offering precision and ease of use.

In this study, Unmanned Aerial Vehicles (UAVs) were employed to capture high-resolution images of Wat Si Phichit Kirati Kanlayaram stupa, ensuring comprehensive coverage of its surface, including

Fig. 1 Wat Si Phichit Kirati Kanlayaram stupa used for case study

hard-to-reach areas. These images were processed using photogrammetry software to generate a detailed 3D model, accurately reflecting the stupa's structural condition. The model was then imported into an FTVR application developed using HoloSDK [1]. This application allows inspectors to navigate, identify structural defects, and annotate issues such as cracks, erosion, and material deterioration directly within the 3D space. The case study presented in this research demonstrates the practical application of this system and its potential to revolutionize large-scale structural inspections. By enabling safe, detailed, and interactive virtual inspection workflows, this method addresses the limitations of traditional techniques. It also provides a robust alternative for professionals involved in preserving and maintaining historical structures. Beyond its immediate applications, this approach represents a significant step forward in advancing inspection methodologies, offering insights for both heritage conservation and broader structural analysis.

The subsequent sections of this article are organized to provide a comprehensive overview. Section 3 explores previous research on Fish Tank Virtual Reality (FTVR), the role of photogrammetry in cultural heritage and structural inspection, and how VR/AR is being used for structural inspection. Section 4 presents the methodology, covering frustum calculations, interaction tools, navigation, and annotation features within the FTVR environment. Section 5 presents a case study on inspecting Wat Si Phichit Kirati Kanlayaram, including drone image collection, photogrammetric model creation, and the virtual inspection process. Section 6 discusses the impact of our system, its limitations, and possible future improvements. Finally, Section 7 wraps up with conclusions, summarizing key findings and potential advancements in FTVR-based inspection systems.

2. RESEARCH SIGNIFICANCE

This study introduces a Fish Tank Virtual Reality-based application for structural inspection, offering a novel, immersive, and precise method for evaluating structures. By integrating photogrammetry and FTVR, the research addresses key challenges in traditional inspection practices, such as accessibility, safety, and accuracy. The proposed system demonstrates its potential to transform the state of practice by enabling detailed, interactive evaluations in a virtual environment. This advancement not only enhances inspection efficiency but also sets the foundation for broader applications in remote and automated structural assessments, thereby contributing to the future of engineering inspections.

3. RELATED WORKS

In this section, previous FTVR studies, photogrammetry, and recent applications of VR/AR in structural inspection are discussed.

3.1 Fish Tank Virtual Reality

The concept of Fish Tank Virtual Reality is described as a stereo image of a 3D scene displayed on a monitor with perspective projection aligned to the observer's head position [2]. Their pioneering work demonstrated the effectiveness of FTVR in determining depth ordering within virtual environments. Rekimoto expanded on this by showing that head-coupled perspective techniques significantly improve users' cognitive abilities to interpret complex 3D structures using standard displays, eliminating the need for specialized hardware [3] A user-centered quantitative study confirmed that participants preferred FTVR and stereoscopic setups for presenting depth perception [4].

Jácome et al. introduced the Parallax engine, a system supporting both FTVR and 2D parallax modes using the Unity game engine [5]. Notably, their approach eliminated the need for additional hardware, making FTVR more accessible. Demiralp et al. compared FTVR with CAVE systems, concluding that FTVR displays outperformed CAVE setups for tasks anchored within the user's reference frame [6]. Participants favored FTVR due to its superior resolution, brightness, crispness, and ease of use. Fafard et al. explored volumetric FTVR displays, such as spherical setups, highlighting their potential for mixed-reality applications, including collaborative tasks like multiplayer gaming [7].

In another FTVR-focused study, Teather et al. examined visual and motor co-location using a 3D object manipulation task [8]. Their findings suggested that while co-locating display and input offered limited advantages, movement tasks within the FTVR environment were simpler. Zhou et al. presented an FTVR-based multi-user collaboration system, integrating spherical displays and additional mobile screens [9]. Their research demonstrated that spherical displays encourage the co-location of multiple users around a single FTVR display, facilitating collaborative interactions.

Kongsilp et al. investigated the critical role of motion parallax in enhancing the sensation of presence within FTVR environments, reducing visual fatigue, and augmenting depth perception [10]. Their subsequent work highlighted that in FTVR setups, users relied more on stereopsis than motion parallax for depth perception, particularly for tasks requiring fine depth acuity. Together, these studies underscore the versatility of FTVR in delivering high-fidelity visual experiences and its potential for collaborative and task-oriented applications.

3.2 Photogrammetry in Cultural Heritage and Structural Inspection

Photogrammetry has emerged as a valuable tool for structural inspection and damage detection in various contexts. It offers advantages in assessing large-volume structures and heritage buildings providing accurate 3D reconstructions and damage mapping [11, 12]. Recent advancements in technology, including the use of Unmanned Aerial Vehicles (UAVs) and specialized software, have made photogrammetry more accessible and efficient for capturing and processing data [13, 14, 15]. Photogrammetry has been successfully applied to various cultural heritage sites, including the UNESCO Foguang Monastery in China, the Global Vipassana Pagoda in India, Pegulingan Temple and the Sanggrahan Temple in Indonesia [16, 17, 18, 19]. The method combines aerial imagery from drones with terrestrial photography to generate comprehensive 3D models, allowing for accurate geometric analysis and structural monitoring [17, 18]. Close-range photogrammetry applied to the Sewu Temple in Indonesia demonstrated high precision in creating detailed architectural models, proving it an efficient and low-cost solution for complex architectural surveys [20] These models enable precise measurements, defect identification, and monitoring of structural changes over time[13, 21]

Photogrammetry enables accurate 3D reconstruction of real-world objects, while VR provides immersive and interactive experiences [22]. This combination allows for the creation of detailed virtual environments of cultural heritage sites, offering remote access and exploration with six degrees of freedom [23]. Fully immersive VR applications have revolutionized access to cultural heritage, offering virtual museum experiences and ancient building tours [24]. The integration of photogrammetry, 3D modeling, and VR has enhanced restoration workflows, as demonstrated in the case study of Pashas Bridge in Greece [25]. The workflow typically involves photogrammetric capture, 3D model optimization, and VR application development using game engines [22]. Some systems incorporate augmented reality and VR headsets to enhance user experience and visualization [26]. Advanced applications, like PhotoTwinVR, enable manipulation, inspection, and dimension measurements of 3D photogrammetric models in VR, potentially benefiting engineering and non-engineering professionals in off-line inspection processes [27]. These technologies offer new possibilities for preserving, studying, and experiencing cultural heritage and complex structures.

3.3 Applications of VR/AR in Structural Inspection

Virtual and augmented reality (VR/AR) technologies are emerging as valuable tools for structural inspection and monitoring. Recent research explores the application of VR/AR technologies in infrastructure inspection, offering significant improvements in safety, efficiency, and data visualization. VR applications enable remote inspection of bridges and viaducts, integrating 3D models with structural information for comprehensive analysis [28]. AR technology has been deployed in various civil infrastructure domains, including construction, structural health monitoring, and damage detection [29]. The use of LIDAR scanning combined with VR creates digital twins of bridges, allowing for safer and more accessible inspections [30] UAV-based photogrammetry has shown promise in monitoring concrete bridge cracks with millimeter-level accuracy [31]. The technique allows for non-contact inspections, reducing costs and safety risks associated with manual methods. UAVs combined with Internet of Things (IoT) systems and VR enable remote monitoring of historical buildings, enhancing Structural Health Monitoring (SHM) activities [32].

VR/AR applications can integrate building information modeling (BIM) data, photogrammetry, LiDAR scans, and finite element analysis results to create immersive environments for inspectors [33, 34]. Overall, these technologies show promise in enhancing structural inspection processes, enabling remote collaboration, and improving data visualization and analysis.

4. METHODOLOGY

This study employs photogrammetry and FTVR technologies to develop an interactive virtual inspection system. The methodology involves two key components: creating a photogrammetry-based 3D model and integrating it into an FTVR application. This section outlines the complete workflow, focusing on generating a high-resolution 3D model and the development of FTVR application.

4.1 Photogrammetry Model Creation

Photogrammetry was utilized to create a detailed 3D model of the structure, ensuring accurate visualization and documentation. The process is divided into two main stages: image acquisition and image-based 3D modeling.

4.1.1 Image Acquisition

The first step in photogrammetry involves capturing high-resolution images of the structure’s surface. UAVs equipped with professional-grade cameras are employed to achieve comprehensive coverage, including areas that are difficult to access. UAVs are flown along pre-planned paths to ensure consistent image overlap, essential for accurate 3D reconstruction. A combination of flight paths is used to capture the structure from multiple perspectives.

During image acquisition, camera settings, such as focal length, aperture, and shutter speed, are carefully configured to maximize image clarity and minimize distortions caused by lighting variations or motion. Consistent ground sampling distance (GSD) is maintained to ensure uniform image resolution across the entire structure. A typical workflow includes capturing approximately 500 images with at least 60% overlap between consecutive images and 40% side overlap between adjacent passes, providing sufficient data for effective photogrammetry processing.

4.1.2 Image-Based 3D Modeling

After image acquisition, the collected images are processed using photogrammetry software (such as Agisoft) to create a 3D model of the structure. The process begins with feature extraction, where the software identifies distinct points, such as edges and corners, across overlapping images. These features are matched between images to establish correspondences, forming the basis for reconstructing the spatial geometry of the structure.

Using the matched features, the software generates a sparse point cloud, representing a basic 3D outline of the structure. This point cloud is then densified through interpolation, resulting in a dense point cloud that captures intricate structural details. From the dense point cloud, a triangular mesh is created to form the geometric framework of the 3D model. The mesh is optimized to reduce the polygon count without compromising critical structural details, ensuring efficient rendering in virtual environments.

Finally, high-resolution texture mapping is applied to the mesh, using the original images to project realistic colors and surface details onto the model. This step ensures that the 3D model accurately reflects both the geometry and visual appearance of the structure. The completed model is then exported in a standard format, such as OBJ or FBX, to facilitate its integration into the FTVR application.

This photogrammetry-based approach provides a precise and detailed digital representation of the structure, forming the foundation for the subsequent development of the virtual inspection environment. The next section describes the integration of this model into the FTVR application and its functionality for defect identification and annotation.

4.2 Fish Tank Virtual Reality Application

FTVR combines head-coupled perspective projection with stereoscopic imaging on a fixed display, creating an immersive illusion of depth. In this study, an FTVR application was developed using Unity game engine and HoloSDK to enable an interactive virtual inspection environment. HoloSDK provides the functionality for head tracking, allowing the system to dynamically adjust the rendered perspective based on the user's head position. This configuration suits FTVR for tasks that demand high precision and detail in a desktop-based virtual environment.

The FTVR setup (Fig. 2) requires a built-in or external web camera and a pair of red-blue anaglyph glasses. The web camera tracks the user's head position in real-time, enabling the application to adjust the displayed view dynamically. This head-coupled perspective ensures the user experiences a realistic depth illusion when interacting with the virtual environment. The anaglyph glasses allow users to view stereoscopic content rendered on a standard monitor, creating depth perception.

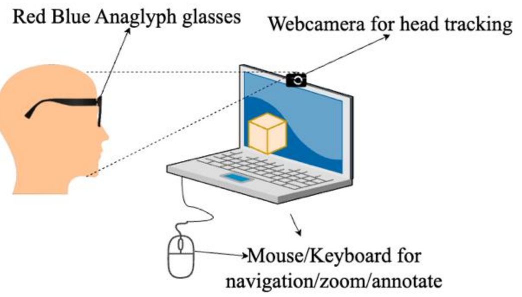


Fig. 2 FTVR setup

The stereoscopic 3D effect in FTVR is achieved by generating two slightly offset perspectives, one for each eye, based on the user's inter-pupillary distance (IPD). These perspectives are rendered using a half-anaglyph shader compatible with red-blue glasses. The brain combines the two perspectives into a cohesive 3D image, producing a convincing sense of depth. The rendering process involves computing the frustum for each eye, which defines the visible volume and determines the proper disparity between the two views.

Frustum (Fig. 3) calculations are based on parameters such as IPD, screen width, and height, near and far clipping planes, and the projection matrix. These calculations ensure that the left and right eye views are correctly aligned to replicate how the eyes perceive depth in the real world. For instance, the frustum defines the spatial boundaries within which objects appear in 3D while maintaining accurate proportions and relative distances. These computations are crucial for creating a realistic stereoscopic experience, as any misalignment can disrupt the perception of depth and immersion.

Frustum calculation is central to creating a realistic depth perception in the FTVR environment. This process involves defining the visible volume from the perspective of each eye, ensuring accurate stereoscopic rendering. Several parameters are used in this calculation, including the camera's position, the near clip plane distance, the field of view (FoV), and the aspect ratio of the display screen.

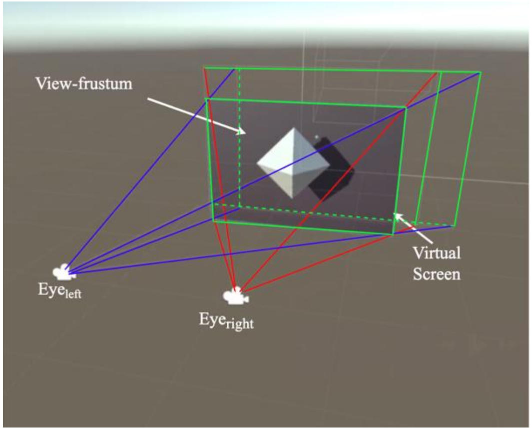


Fig. 3 FTVR frustum

The aspect ratio of the display screen is calculated by dividing its width by its height, ensuring that the rendered scene aligns with the proportions of the physical screen:

$$\text{aspectRatio} = \frac{\text{screenWidth}}{\text{screenHeight}} \tag{1}$$

Using the aspect ratio and the field of view, the virtual screen dimensions at the near clipping plane are calculated as follows:

$$\text{virtualScreenHeight} = 2 \times \text{nearClipPlane} \times \tan\left(\frac{\text{FOV}}{2}\right) \tag{2}$$

$$\text{virtualScreenWidth} = \text{aspectRatio} \times \text{virtualScreenHeight} \tag{3}$$

For stereoscopic rendering, the inter-pupillary distance (IPD) introduces the parallax required for depth perception. The x-coordinates of the frustums for the left and right eyes are offset by half the IPD, scaled by the ratio of the virtual screen width to the actual screen width. These offsets are calculated as follows:

$$x_{\text{left}} = x - \left(\frac{\text{IPD} \times \left(\frac{\text{virtualScreenWidth}}{\text{screenWidth}}\right)}{2}\right) \tag{4}$$

For the right eye:

$$\boldsymbol{x}_{\text{right}} = \boldsymbol{x} + \left(\frac{\text{IPD} \times \left(\frac{\text{virtualScreenWidth}}{\text{screenWidth}}\right)}{\mathbf{2}}\right) \tag{5}$$

Next, the edges of the frustum are determined by calculating the angles that define the left, right, top, and bottom boundaries. These angles depend on the relative positions of the virtual screen and the camera:

$$\text{leftAngle} = \frac{\left(\frac{\text{virtualScreenWidth}}{2} + x_{\text{left}}\right)}{-z} \tag{6}$$

$$\text{rightAngle} = \frac{\left(\frac{\text{virtualScreenWidth}}{\mathbf{2}} - \boldsymbol{x}_{\text{right}}\right)}{-\boldsymbol{z}} \tag{7}$$

$$\text{topAngle} = \frac{\left(\frac{\text{virtualScreenHeight}}{\mathbf{2}} - \boldsymbol{y}\right)}{-\boldsymbol{z}} \tag{8}$$

$$\text{bottomAngle} = \frac{\left(\frac{\text{virtualScreenHeight}}{\mathbf{2}} + \boldsymbol{y}\right)}{-\boldsymbol{z}} \tag{9}$$

Using these angles, the left, right, top, and bottom boundaries of the frustum are defined at the near clipping plane:

$$\text{left} = -\text{nearClipPlane} \times \text{leftAngle} \tag{10}$$

$$\text{right} = \text{nearClipPlane} \times \text{rightAngle}. \tag{11}$$

$$\text{top} = \text{nearClipPlane} \times \text{topAngle}\,. \tag{12}$$

$$\text{bottom} = -\text{nearClipPlane} \times \text{bottomAngle} \tag{13}$$

Finally, the perspective projection matrix is constructed using these frustum boundaries along with the far clipping plane. This matrix transforms 3D world coordinates into 2D screen coordinates, simulating the viewer's perspective:

$$P = \begin{bmatrix} \frac{2 \times \text{nearClipPlane}}{r-l} & 0 & \frac{r+l}{r-l} & 0 \\ 0 & \frac{2 \times \text{nearClipPlane}}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & -\frac{z_{\text{far}} + z_{\text{near}}}{z_{\text{far}} - z_{\text{near}}} & -\frac{2 \times z_{\text{far}} \times z_{\text{near}}}{z_{\text{far}} - z_{\text{near}}} \\ 0 & 0 & -1 & 0 \end{bmatrix} \tag{14}$$

Where:

- l, r, t, and b represent the left, right, top, and bottom frustum planes.
- z_{near} and z_{far} represent the near and far clipping planes.

By applying the projection matrix to each eye's view, slightly different perspectives of the scene are created, which are then fused by the brain into a cohesive 3D image. In the FTVR application, this projection matrix is combined with a half-anaglyph shader to achieve the depth illusion. Anaglyph rendering is employed to encode two distinct images, one for each eye, into different color channels: red for the left eye and blue for the right eye. The half-color anaglyph method enhances image quality by keeping one eye’s view in grayscale while applying the red-blue filter to the other. The shader processes the frustum views for both eyes, applying the appropriate red-blue filters to each. This process creates the necessary visual separation, which is then decoded by the red-blue glasses worn by the user. The glasses allow the left eye to perceive only the red channel and the right eye to perceive only the blue channel. When the combined image is displayed, the brain interprets it as a 3D scene, providing an immersive perception of depth.

Effective inspection in the FTVR environment is facilitated by an integrated suite of user interaction tools, tailored to enhance navigation, close examination, and defect documentation.

- **Navigation Tools:** The navigation tools allow users to seamlessly explore the virtual space. This includes controls for panning, rotating, and zooming the virtual camera, enabling users to view the 3D model from any angle. Designed with user intuitiveness in mind, these tools ensure that navigation is smooth, allowing users to focus entirely on the inspection process.
- **Zoom Functionality:** Zoom functionality is a vital component for detailed inspections, allowing users to closely examine specific areas of the 3D model. The zoom tools ensure high levels of clarity and precision, even at significant magnification, enabling inspectors to identify subtle defects and intricate features that may not be visible from a standard view.
- **Annotation Tools:** An annotation system is integrated to enable inspectors to document findings directly on the 3D model. This includes marking areas of interest, classifying defects, and adding descriptive notes. The annotations are designed to be user-friendly, allowing inspectors to efficiently record observations with minimal effort. These annotations are stored within the system, creating a digital record that can be revisited for further analysis or reporting purposes. purposes.

5. CASE STUDY: INSPECTION OF WAT SI PHICHIT KIRATI KANLAYARAM, SUKHOTHAI, THAILAND

Wat Si Phichit Kirati Kanlayaram (Fig. 1), a historical stupa located in Sukhothai, Thailand, embodies the cultural and religious heritage of its era. As a traditional Buddhist monument, it holds immense value as both a sacred site and an architectural artifact, representing the intricate craftsmanship and spiritual significance of ancient stupa design. However, like many historical structures, the stupa has faced challenges due to prolonged exposure to environmental factors, leading to issues such as surface erosion, cracks, and material degradation. Regular and detailed inspections are essential to preserve its structural integrity and ensure the longevity of its cultural legacy. In this study, photogrammetry techniques and FTVR are utilized to inspect the stupa, offering a detailed and interactive remote inspection method. Fig. 4 shows the complete overview of the case study. By combining high-resolution 3D modeling with virtual reality, the application enabled a comprehensive examination of the stupa's surface while minimizing the risks and logistical challenges associated with on-site inspections.

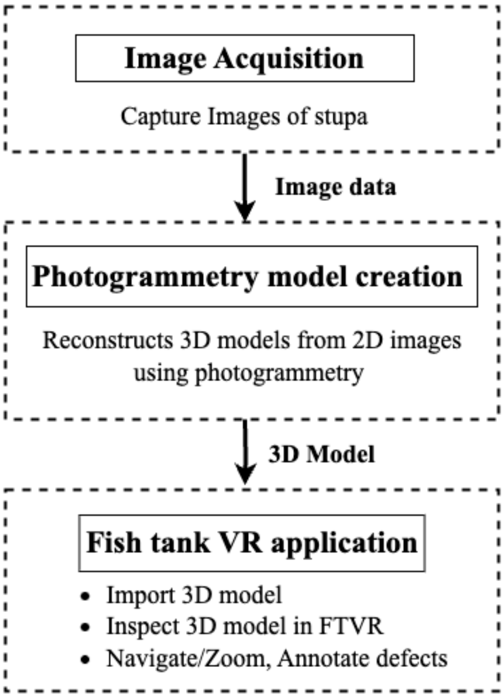


Fig. 4 Overview of the case study

5.1 Drone Image Collection

The inspection process begins with the collection of high-resolution images (Fig. 5) of the stupa using the DJI Phantom 4 RTK drone, a high-precision UAV optimized for photogrammetric applications (Table 1). The drone is programmed to fly around the target, capturing images from multiple angles and elevations to ensure comprehensive coverage of both the overall structure and fine details such as surface cracks and erosion. For this case, a Ground Sampling Distance (GSD) of 1 mm/pixel was selected, ensuring highly detailed imagery suitable for precise 3D modeling and structural inspection.

To achieve this GSD, the required flight height was calculated using the formula:

$$\mathrm{GSD} = \frac{\text{Sensor width} \times \text{Flight height}}{\text{Image width} \times \text{Focal length}} \tag{15}$$

Using the camera specifications of the Phantom 4 RTK, including a sensor width of 13.2 mm, a focal length of 8.8 mm, and an image resolution of 5472 pixels, the optimal flight height was determined to be approximately 3.64 meters. This height ensures that the images captured provide the necessary resolution for detailed modeling while adhering to operational safety standards. While this study maintained a fixed altitude for consistency, variations in flight height can influence model accuracy. Higher altitudes would increase GSD, potentially reducing defect detection capability, whereas lower altitudes could enhance detail but might introduce operational constraints such as increased image overlap requirements and flight duration. Fig. 6 shows drone flight path used for image collection. This approach allows for precise data acquisition and forms the foundation for accurate photogrammetric reconstruction in subsequent stages of the inspection process.

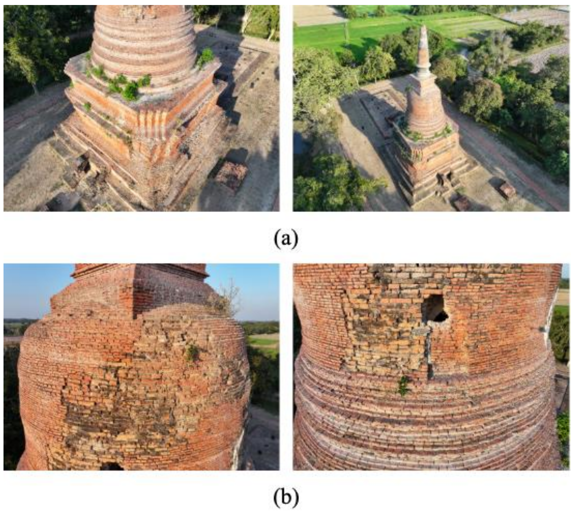


Fig. 5 Captured Images from drone (a) Overview (b) Close-up

Table 1. Drone specification

Parameter	Details
Drone Model	DJI Phantom 4 RTK
Camera Sensor	1-inch CMOS
Sensor Resolution	20 MP (5472 x 3648 pixels)
Focal Length	8.8 mm
Sensor Width	13.2 mm
GSD	1 mm/pixel
Flight Altitude for GSD	3.64 meters
Flight Method	Point of Interest (POI)
Software Used	DJI GS Pro for flight planning

5.2 Photogrammetry Model Creation

After collecting high-resolution images, the next step involves processing the data using photogrammetry software to generate a precise 3D model.

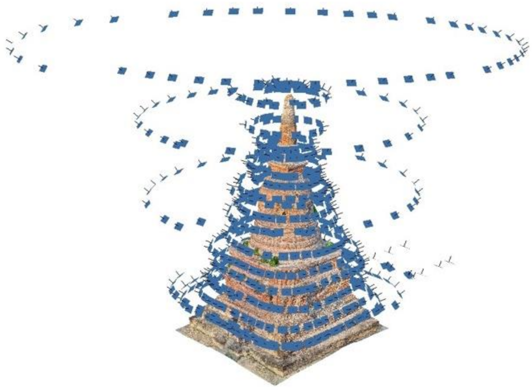


Fig. 6 Image collection for 3D model drone flight path (Camera position in blue rectangle)

The software analyzes overlapping images, identifying shared key points across multiple photographs to establish spatial relationships and reconstruct the stupa's geometry in three dimensions. This process employs advanced algorithms to align the images, estimate camera positions, and create a dense point cloud representing the stupa's surface.

The dense point cloud is further refined to create a detailed mesh model that captures the structural features, textures, and visible imperfections of the dam. The final output is a photorealistic model (Fig. 7) that provides a comprehensive and accurate representation of the structure.

To prepare the 3D model for use in the FTVR application, it is optimized to balance detail and performance. The optimization process involves reducing the polygon count and compressing textures while retaining essential features for inspection. This ensures the 3D model can be rendered smoothly in the FTVR environment, facilitating real-time interaction and analysis without compromising visual fidelity.

Table 2. Summary of 3D modelling results for stupa

Parameter	Details
Number of Images	828 images
Projection	1,371,118 features
RMS Re-projection error	0.303 pixels
Sparse Point Cloud	530,872 points
Dense Point Cloud	38,185,298 points
Surface Mesh	6,527,385 faces
Vertex	3,286,955 vertices
GSD (Ground Sampling Distance)	1 mm / pixel

5.3 Integration of the 3D Model of stupa into the FTVR Application

Once the optimized 3D model of the stupa is prepared, it is imported into the FTVR application to facilitate detailed inspection in an interactive virtual environment. The integration process begins by setting up the virtual scene within the Unity engine, where the model is positioned and scaled to match real-world dimensions. Lighting conditions are configured to enhance visibility and highlight surface features, ensuring inspectors can identify structural details such as cracks, erosion, or other defects. The FTVR environment utilizes head-tracked perspective rendering combined with stereoscopic imaging to provide a realistic 3D visualization. This involves configuring the projection parameters and aligning

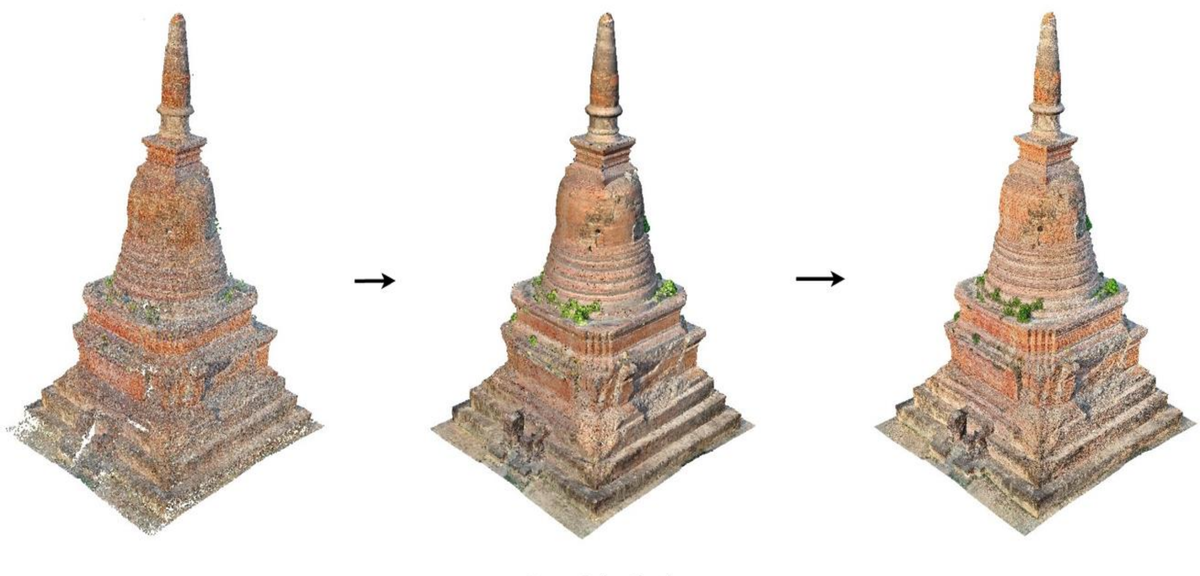


Fig. 7 Photogrammetry to generate realistic 3D model

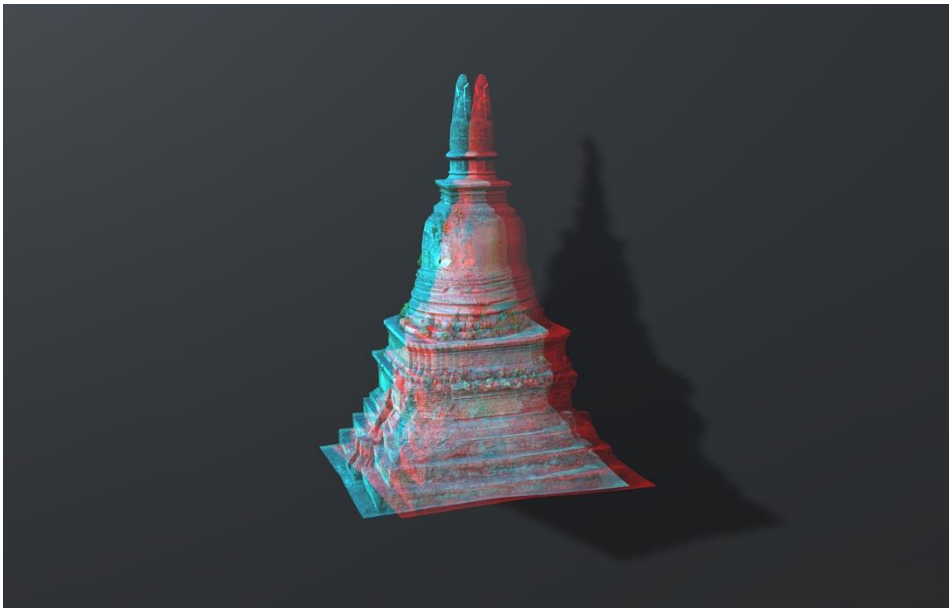


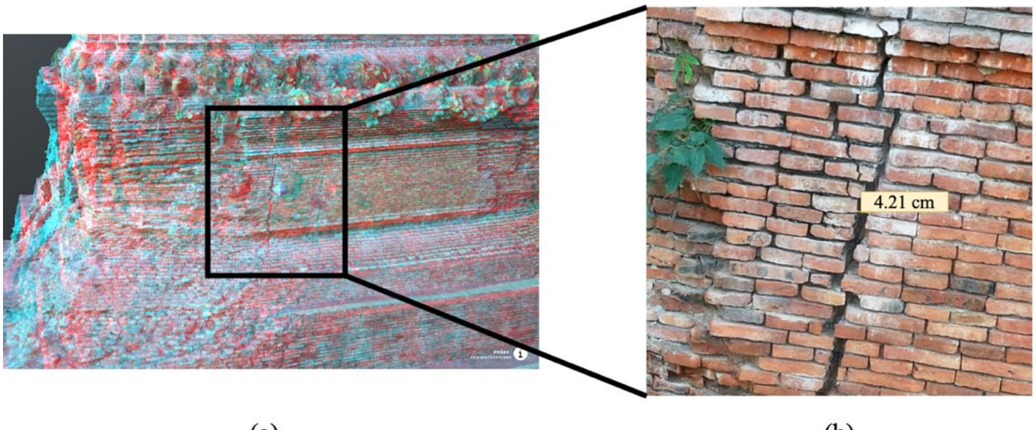
Fig. 8 Stupa model combined with half-anaglyph shader in FTVR

the 3D model within the frustum defined for each eye. By combining these elements, the application creates a depth-perception effect that allows inspectors to examine the structure as if they were physically present (Fig. 8).

To further enhance usability, the application includes navigation and interaction tools tailored for inspection tasks. These tools allow users to move around the virtual stupa model, zoom in on specific areas of interest, and annotate defects directly on the stupa model. This seamless integration of the 3D model into the FTVR application provides a practical and efficient platform for conducting virtual inspections, enabling detailed analysis and documentation of structural conditions.

5.4 Conducting the Virtual Inspection

The virtual inspection process leverages the FTVR application to facilitate a detailed and interactive examination of the 3D model of the stupa. Once the model is integrated into the FTVR environment, inspectors can navigate the virtual space to identify and document potential structural issues. The intuitive interface and interactive tools ensure a seamless inspection experience, allowing users to focus on analyzing the structure. Navigation within the VR environment is achieved through head-tracked perspective rendering, enabling users to view the model from multiple angles and perspectives. Inspectors can rotate, pan, and zoom into specific areas to closely examine fine details such as surface cracks, erosion, or other visible defects. The stereoscopic imaging creates a realistic depth-perception effect, enhancing the ability to detect subtle irregularities on the surface. The annotation functionality further augments the inspection process. Inspectors can mark areas (Fig. 9) of concern directly on the 3D model. These annotations are stored within the system, creating a detailed record of findings that



(a) (b)

Fig. 9 Annotation marked (a) In FTVR application (b) Captured image showing crack on stupa - crack width (4.21 cm) can be reviewed later. This capability not only aids in documenting the current state of the structure but also provides valuable data for comparison during future inspections. The reported annotation accuracy of crack dimensions (4.21 cm) demonstrates the system's precision in documenting structural defects. This measurement was confirmed through repeated annotations by the same operator, with minimal variation observed—within a few millimeters—around the recorded value. While this suggests reliable performance for a single user, we acknowledge that a formal repeatability and reproducibility study across multiple users was not conducted in this work.

The virtual inspection process demonstrates the potential of FTVR to enhance traditional structural assessment methods by offering a safe, detailed, and accessible platform for remote inspection. This approach minimizes the risks associated with on-site inspections and provides inspectors with tools to conduct a thorough analysis with high precision and efficiency.

6. DISCUSSION

6.1 Impact of the Proposed System on Inspection Practices

FTVR and photogrammetry in structural inspection introduce a transformative approach to evaluating large-scale structures. By offering a high-resolution 3D model in a virtual environment, the system enhances the ability to detect and document defects such as cracks, erosion, and other structural anomalies with remarkable precision. This shift toward virtual inspection not only reduces the need for physical site visits but also ensures a safer and more efficient workflow, particularly for inaccessible or hazardous locations. One of the most significant impacts of the proposed system is its capacity to facilitate a comprehensive inspection without the constraints of time or weather. Inspectors can repeatedly analyze the virtual model, focusing on specific areas and revisiting annotations to ensure thorough documentation. This iterative capability fosters a deeper understanding of structural conditions and allows for cross-verification, minimizing the likelihood of oversights. While this study did not conduct a direct quantitative comparison of error rates between traditional and virtual inspection methods, existing literature supports the premise that virtual methods can substantially reduce human error in inspection tasks. While this study did not conduct a direct quantitative comparison of error rates between traditional and virtual inspection methods, existing literature supports the premise that virtual methods can substantially reduce human error in inspection tasks. A comparison study of conventional visual inspection techniques with a VR approach using Lidar for concrete bridge inspection. Their findings indicate that the VR inspection improved accuracy, safety, and efficiency over traditional methods[35]. Another study explores the use of 3D modeling for the conservation and monitoring of heritage buildings, focusing on photogrammetric, digital twin techniques, and control points for accuracy. It reports an 83% of the cracks were detected with modern techniques[36]. Moreover, the system's interactive tools, such as annotation and zoom functionalities, streamline the inspection process by enabling users to document findings directly on the model. Navigation and documentation tools into a single platform provide a holistic inspection workflow, bridging the gap between visual assessment and recordkeeping. The system's scalability and adaptability further enhance its impact. It can be applied to a variety of structures beyond historical monuments, such as dams, bridges, and buildings, making it a versatile solution for diverse inspection needs. By leveraging advanced visualization techniques, the proposed system sets a new standard for precision and accessibility in structural inspection.

6.2 Limitations of the Proposed System

While the proposed FTVR-based inspection system offers significant advancements in structural assessment, certain limitations must be acknowledged. These constraints primarily stem from technical dependencies, operational considerations, and user interaction challenges. One limitation lies in the dependency on high-quality image data and the precision of photogrammetry software. The accuracy of the 3D model is directly influenced by the quality of the images collected, the GSD, and the processing algorithms. Variations in lighting, obstructions during image capture, or incomplete coverage can result in artifacts or gaps in the final model, potentially affecting the reliability of the inspection. The system also relies on hardware components such as head-tracking cameras and stereoscopic displays, which may introduce calibration challenges or compatibility issues. For instance, maintaining consistent head-tracking performance is critical for preserving the depth perception and interactivity essential to FTVR. Any hardware failure or suboptimal setup can compromise the overall user experience. User proficiency is another consideration. While the interface and tools are designed to be intuitive, inspectors may require training to navigate and interact effectively with the virtual environment, particularly those accustomed to traditional inspection methods. Additionally, the red-blue anaglyph system can cause visual fatigue over extended use, potentially limiting the duration of inspection sessions. Lastly, the current system does not incorporate real-time data updates or integrate multi-sensor input, such as thermal or ultrasonic data, which could provide a more comprehensive inspection framework. Addressing these limitations will be essential to optimize the system for broader adoption and use in various contexts.

6.3 Future Direction

The proposed FTVR-based inspection system represents a significant step forward in structural inspection, but its potential can be further enhanced through targeted advancements and integration of emerging technologies. Future developments could focus on addressing current limitations and expanding the system's capabilities to make it more robust and versatile.

One promising direction is the integration of multi-sensor data to enrich the inspection process. By incorporating thermal imaging, LiDAR, or ultrasonic data alongside photogrammetry, the system could provide a multi-dimensional analysis of structural integrity, detecting subsurface defects and thermal anomalies that are not visible on the surface. This comprehensive approach would significantly enhance the system's diagnostic capabilities. Improving the scalability and efficiency of 3D model generation is another area for development. Automating image capture with pre-programmed drone flight paths and advanced AI-based photogrammetry algorithms could streamline the creation of high-resolution models, reducing processing times and minimizing human intervention. Moreover, the selection of an optimal flight altitude plays a critical role in determining the accuracy and level of detail in photogrammetric reconstructions. In this study, a flight height of approximately 3.64 meters was chosen to achieve a 1 mm/pixel GSD, ensuring high-resolution capture of surface details. While this approach provided clear documentation of defects such as cracks and material erosion. Different flight altitudes could impact model accuracy and defect detectability. Higher flight altitudes would increase GSD, potentially reducing the resolution necessary for identifying fine structural details, while lower altitudes could enhance precision but introduce challenges such as increased data volume, processing time, and flight duration. Future research should explore a comparative analysis of multiple flight heights to determine the optimal balance between resolution, efficiency, and operational feasibility in UAV-based inspections. Integrating adaptive flight planning techniques based on structural complexity and defect density may further refine data acquisition strategies for enhanced inspection outcomes.

The photogrammetry workflow successfully reconstructed a detailed 3D model, ensuring an accurate geometric representation of the inspected structure. While terrestrial laser scanning (TLS) is often used as a benchmark for validating photogrammetric accuracy, our study did not incorporate TLS data due to resource constraints. Instead, we relied on built-in photogrammetry accuracy metrics, such as reprojection errors, and conducted manual ground control checks to verify scale accuracy. TLS cross-validation could further enhance credibility and provide a more rigorous comparison. However, our study's primary objective was to demonstrate the feasibility of an FTVR-based inspection workflow rather than performing a multi-instrument accuracy assessment. Future research could explore TLS integration to further refine validation methods and assess comparative performance.

Another direction involves enhancing the user interface and interaction tools. Implementing advanced navigation controls, voice-command functionalities, or haptic feedback could improve the inspection experience. The inclusion of real-time collaboration features is also a valuable enhancement. Enabling multiple inspectors to collaborate within the virtual environment, share annotations, and discuss findings in real time could increase efficiency and facilitate team-based inspections. The proposed FTVR system provides an intuitive environment for remote inspection, but its usability may vary across different professional domains. Engineers, who are trained to identify structural defects, may find the system immediately useful for damage assessment. In contrast, archaeologists or heritage conservation professionals may focus more on material composition, artistic details, or historical significance. While the interface is designed to be user-friendly, some users with less experience in 3D navigation might require short training sessions to maximize their effectiveness. Future research should include a comparative usability study, evaluating how different professionals interact with the system, the challenges they encounter, and any adjustments that could enhance accessibility across disciplines. This would provide valuable insights for further refining the system to accommodate a broader range of experts involved in heritage and structural inspections. These advancements, when realized, could significantly broaden the system's impact and establish it as a leading tool in the field of structural inspection.

7. CONCLUSION

This study presents a novel approach to structural inspection by integrating Fish Tank Virtual Reality with photogrammetry-generated 3D models. The proposed system leverages advanced imaging and VR technologies to create a detailed, interactive, and immersive inspection platform. By utilizing high-resolution UAV imagery and photogrammetry software, accurate 3D models of structures can be developed, enabling inspectors to identify and annotate defects with precision. The application of

FTVR enhances the inspection process by providing an intuitive, depth-perceptive virtual environment, accessible with relatively low-cost hardware.

The case study of Wat Si Phichit Kirati Kanlayaram illustrates the practical utility and adaptability of the system. The detailed 3D reconstruction of the stupa and its subsequent virtual inspection demonstrate the system's potential to improve traditional inspection workflows. The ability to annotate defects directly on the model, coupled with advanced navigation tools, ensures a thorough and efficient evaluation of structural integrity.

While the system has limitations, such as dependency on high-quality imagery and potential hardware constraints, it lays the foundation for future advancements. Integration of multi-sensor data, real-time collaboration tools, and predictive maintenance capabilities could further enhance its functionality and impact. Additionally, the cost implications of adopting this technology are considered. While the initial investment in UAV platforms, photogrammetry software licenses, and FTVR setups may seem high, it is offset by reduced long-term costs associated with traditional inspection methods that require scaffolding, extensive site visits, and prolonged labor hours. The ability to conduct repeated virtual inspections without additional field deployment offers significant economic advantages. However, a more detailed cost-benefit analysis comparing this approach to conventional methods remains an important area for future research.

In conclusion, the FTVR-based inspection system represents a significant advancement in the field of structural assessment, providing a safe, detailed, and interactive alternative to traditional methods. Its adaptability to various structures and potential for continued innovation make it a promising tool for engineers and inspectors seeking to enhance the efficiency and accuracy of structural evaluations.

8. ACKNOWLEDGMENTS

We appreciate the financial support that the Thammasat School of Engineering (Thammasat University, Rangsit Campus) offers to students in the form of scholarships for the Master’s degree program.

9. REFERENCES

[1] Codes P., “Holo-SDK,” 2021, [Online]. Available: https://www.holo-sdk.com/

[2] Ware C., Arthur K., and Booth K.S., Fish Tank Virtual Reality, in Proc. INTERACT'93 and CHI'93 Conf. on Human Factors in Computing Systems, 1993, pp. 37-42.

[3] Rekimoto J., A Vision-Based Head Tracker for Fish Tank Virtual Reality—VR Without Head Gear, in Proc. Virtual Reality Annual International Symposium'95, IEEE, 1995, pp. 94-100.

[4] Wright E., Connolly P.E., Sackley M., McCollom J., Malek S., Fan K., A Comparative Analysis of Fish Tank Virtual Reality to Stereoscopic 3D Imagery, in Proc. 67th Midyear Meeting Proceedings, 2012, pp. 37-45.

[5] Jácome J., Gomes A., Costa W. de L., Silva Figueiredo L., Abreu J., Porciuncula L., Brant P.K., Alves L.E.M., Correia W.F.M., Teichrieb V., Parallax Engine: Head Controlled Motion Parallax Using Notebooks’ RGB Camera, in Proc. 23rd Symposium on Virtual and Augmented Reality, 2021, pp. 137-146.

[6] Demiralp C., Jackson C.D., Karelitz D.B., Zhang S., and Laidlaw D.H., Cave and Fishtank Virtual-Reality Displays: A Qualitative and Quantitative Comparison, IEEE Trans. on Visualization and Computer Graphics, Vol. 12, Issue 3, 2006, pp. 323-330.

[7] Fafard D.B., Zhou Q., Chamberlain C., Hagemann G., Fels S., and Stavness I., Design and Implementation of a Multi-Person Fish-Tank Virtual Reality Display, in Proc. 24th ACM Symposium on Virtual Reality Software and Technology, 2018, pp. 1-9.

[8] Teather R.J., Allison R.S., and Stuerzlinger W., Evaluating Visual/Motor Co-Location in Fish-Tank Virtual Reality, in Proc. 2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH), IEEE, 2009, pp. 624-629.

[9] Zhou Q., Hagemann G., Fels S., Fafard D., Wagemakers A., Chamberlain C., and Stavness I., Coglobe: A Co-Located Multi-Person FTVR Experience, in ACM SIGGRAPH 2018 Emerging Technologies, 2018, pp. 1-2.

[10] Kongsilp S. and Dailey M.N., User Behavior and the Importance of Stereo for Depth Perception in Fish Tank Virtual Reality, PRESENCE: Virtual and Augmented Reality, Vol. 27, Issue 2, 2018, pp. 206-225.

[11] Trombini M., Maisano D.A., and Pagani A., Photogrammetric Analysis for Inspection and Damage Detection: Preliminary Assessment and Future Extension to Large-Volume Structures, Aerospace Science and Engineering: III Aerospace PhD-Days, Vol. 33, 2023, pp. 355.

[12] De Fino M., Galantucci R.A., and Fatiguso F., Condition Assessment of Heritage Buildings via Photogrammetry: A Scoping Review from the Perspective of Decision Makers, Heritage, Vol. 6, Issue 11, 2023, pp. 7031-7066.

[13] Markova M. and Kravchenko D., 3D Photogrammetry Application for Building Inspection of Cultural Heritage Objects, Bulletin of Prydniprovs’ka State Academy of Civil Engineering and Architecture, No. 1 (237-238), 2018, pp. 90-95.

[14] Themistocleous K., Ioannides M., Agapiou A., and Hadjimitsis D.G., The Methodology of Documenting Cultural Heritage Sites Using Photogrammetry, UAV, and 3D Printing Techniques: The Case Study of Asinou Church in Cyprus, in Proc. Third International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2015), Vol. 9535, SPIE, 2015, pp. 312-318.

[15] Buatik A. and Chaiyasarn K., Deformable Object Removal via Exemplar-Based Algorithm for 3D Model of Heritage Structures, GEOMATE Journal, Vol. 21, Issue 84, 2021, pp. 151-158.

[16] Wang X., Photogrammetry Enables the Critical Reinterpretation and Regeneration of Architectural Heritage, in Proc. 41st Conference on Education and Research in Computer Aided Architectural Design in Europe (41st eCAADe 2023), Vol. 2, 2023, pp. 661-670.

[17] Salunke S., Ramsankaran R., Ghosh S., Milani G., Halani B., Cundari G.A., Varma M., Delhi V.S.K., and Gangurde N., Global Vipassana Pagoda: Exterior Geometry Envelope Extraction Using UAV Photogrammetry, in Proc. 2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv), IEEE, 2023, pp. 225-229.

[18] Bioresita F., Hidayat H., Kurniawan J.R., Purnomo S., and Grussenmeyer P., 3D Modelling of Sanggrahan Temple Using UAV Imagery and Terrestrial Photogrammetry Method, in Proc. IOP Conference Series: Earth and Environmental Science, Vol. 1276, No. 1, IOP Publishing, 2023, pp. 012043.

[19] Partama I.G.D. Y., Yastika P.E., and Wijaya I.M.W., 3D Modeling Using UAV-Photogrammetry Technique for Digital Documentation of Cultural Heritage Buildings, GEOMATE Journal, Vol. 28, Issue 126, 2025, pp. 61-70.

[20] Firzal Y., Architectural Photogrammetry: A Low-Cost Image Acquisition Method in Documenting Built Environment, GEOMATE Journal, Vol. 20, Issue 81, 2021, pp. 100-105.

[21] Buatik A., Ali L., and Chaiyasarn K., 3D Model-Based Image Registration for Change Detection in Historical Structures via Unmanned Aerial Vehicle, GEOMATE Journal, Vol. 16, Issue 58, 2019, pp. 132-138.

[22] Obradović M., Vasiljević I., Đurić I., Kičanović J., Stojaković V., and 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, Vol. 10, Issue 8, 2020, pp. 2743.

[23] Dhanda A., Reina Ortiz M., Weigert A., Paladini A., Min A., Gyi M., Su S., Fai S., and Santana Quintero M., Recreating Cultural Heritage Environments for VR Using Photogrammetry, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 42, 2019, pp. 305-310.

[24] Cecotti H., Cultural Heritage in Fully Immersive Virtual Reality, in Proc. Virtual Worlds, Vol. 1, No. 1, Multidisciplinary Digital Publishing Institute, 2022, pp. 82-102.

[25] Tavantzis I., Tsifodimou Z.E., Skalidi I., Karachaliou E., Giannaris I., Bakousi A., Skondras A., Katakalos K., and Stylianidis E., Restoration Workflows by Means of Photogrammetry: The Case Study of Pashas Bridge, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 48, 2023, pp. 1543-1548.

[26] Putra E.Y., Wahyudi A.K., and Dumingan C., A Proposed Combination of Photogrammetry, Augmented Reality, and Virtual Reality Headset for Heritage Visualization, in Proc. 2016 International Conference on Informatics and Computing (ICIC), IEEE, 2016, pp. 43-48.

[27] Tadeja S.K., Rydlewicz W., Lu Y., Kristensson P.O., Bubas T., and Rydlewicz M., PhotoTwinVR: An Immersive System for Manipulation, Inspection, and Dimension Measurements of the 3D Photogrammetric Models of Real-Life Structures in Virtual Reality, CoRR, Vol. abs/1911.09958, 2019.

[28] Fotia A. and Barrile V., Viaduct and Bridge Structural Analysis and Inspection through an App for Immersive Remote Learning, Electronics, Vol. 12, Issue 5, MDPI, 2023, pp. 1220.

[29] Xu J. and Moreu F., A Review of Augmented Reality Applications in Civil Infrastructure During the 4th Industrial Revolution, Frontiers in Built Environment, Vol. 7, 2021, pp. 640732.

[30] Omer M., Margetts L., Hadi Mosleh M., Hewitt S., and Parwaiz M., Use of Gaming Technology to Bring Bridge Inspection to the Office, Structure and Infrastructure Engineering, Vol. 15, Issue 10, Taylor & Francis, 2019, pp. 1292-1307.

[31] Ioli F., Pinto A., and Pinto L., UAV Photogrammetry for Metric Evaluation of Concrete Bridge Cracks, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 43, 2022, pp. 1025-1032.

[32] Bacco M., Barsocchi P., Cassarà P., Germanese D., Gotta A., Leone G.R., Moroni D., Pascali M.A., and Tampucci M., Monitoring Ancient Buildings: Real Deployment of an IoT System Enhanced by UAVs and Virtual Reality, IEEE Access, Vol. 8, IEEE, 2020, pp. 50131-50148.

[33] Luleci F., Li L., Chi J., Reiners D., Cruz-Neira C., and Catbas F.N., Structural Health Monitoring of a Foot Bridge in Virtual Reality Environment, Procedia Structural Integrity, Vol. 37, Elsevier, 2022, pp. 65-72.

[34] Riedlinger U., Klein F., Hill M., Neumann S., Holst R., Oppermann L., and Bahlau S., Digital Support for Bridge Inspectors through Mixed Reality BIM Data Visualizations, in Proc. Computing in Civil Engineering 2021, 2021, pp. 1359-1366.

[35] Omer M., Margetts L., Mosleh M.H., and Cunningham L.S., Inspection of Concrete Bridge Structures: Case Study Comparing Conventional Techniques with a Virtual Reality Approach, Journal of Bridge Engineering, Vol. 26, Issue 10, American Society of Civil Engineers, 2021, pp. 05021010.

[36] Yiğit A.Y. and Uysal M., Automatic Crack Detection and Structural Inspection of Cultural Heritage Buildings Using UAV Photogrammetry and Digital Twin Technology, Journal of Building Engineering, Vol. 94, Elsevier, 2024, pp. 109952.