Enhancing spatial image datasets for utilisation in a simulator for smart city transport navigation

Lepekola Ignatius Lenkoe[1], Ben Kotze[2]

Department of Electrical, Electronic and Computer Engineering

Faculty of Engineering, Built Environment and Information Technology

Central University of Technology, Free-State

Contact email: bkotze@cut.ac.za

Central University of Technology, Free State

IARIA

Lepekola Ignatius Lenkoe

Mr Lepekola Ignatius Lenkoe holds B-Tech: Engineering: Electrical: (Computer Systems), N.Dip: Computer Systems Engineering and Post Graduate Certificate in Education (PGCE) all from Central University of Technology, Free State. He is currently studying towards obtaining Master of Engineering in Electrical Engineering with Central University of Technology, Free State. Mr Lenkoe is currently employed as a Mathematics Educator at Ntediseng Secondary School in Botshabelo (South Africa). Mr Lenkoe is currently doing research in simulations, Artificial Intelligent, machine vision, mixed reality, spatial datasets and development of image-based rendering techniques.

Introduction & background

- The captured images can provide valuable information about the incident e.g. location. The location has the exact Global Positioning System (GPS) coordinates, which can also be an estimation of the location. Figure 1 presents the graphical representation of inquiring a query image to the reference database to find the match between a stored image and the query image [1].

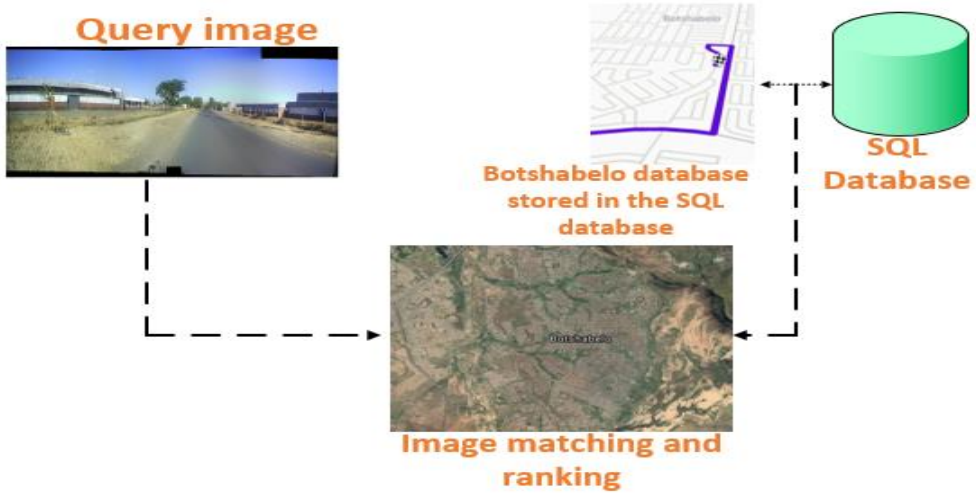

Figure 1: Street-view to overhead view image matching.

Problem statement

- The 3D data acquisition process provides the probe position and orientation that remain in static order to produce accurate datasets.
- A single 3D capacity is not able to support the translational motion of the simulated probes, thus the need to develop a methodology for recording and capturing single 2D images and amalgamating the images into multiple 3D images within a single unit.
- Furthermore, the issue of emulating street-view images for multiple image transitions for application in geolocalisation for utilisation in a simulator needs to be investigated.

Research aim & objectives

This article seeks to develop a 3D rendered model from 2D captured images.

The objective of this study is therefore to:

- To identify datasets with capabilities such as frame position, frame elevation, and frame indexing.
- To incorporate the system into the simulation system in real-time for increasing the reality of the simulation system in different geographical locations.

Original contributions of this research article

This research article has produced the following contributions in furthering the knowledge contribution in the field of computer vision as follows:

- 6 degree of freedom- where the user can move in any direction as opposed to the use of single slides show that allows for one-directional movement in a street view scene.
- The ability of the application to use multiple cameras between 3 to 6 inputs as opposed to the use of single omnidirectional camera feedback and still obtain the same output rendered panoramic and simulated results.

Methodology

- The simulator model design is developed for a driver or person riding a bicycle inside the simulator following a track in either forward or reverse direction.

- The reason for not utilising the omnidirectional camera is because the rendering construction specifically for this research article requires individual camera feeds as opposed to one 360 degrees feed.

Methodology

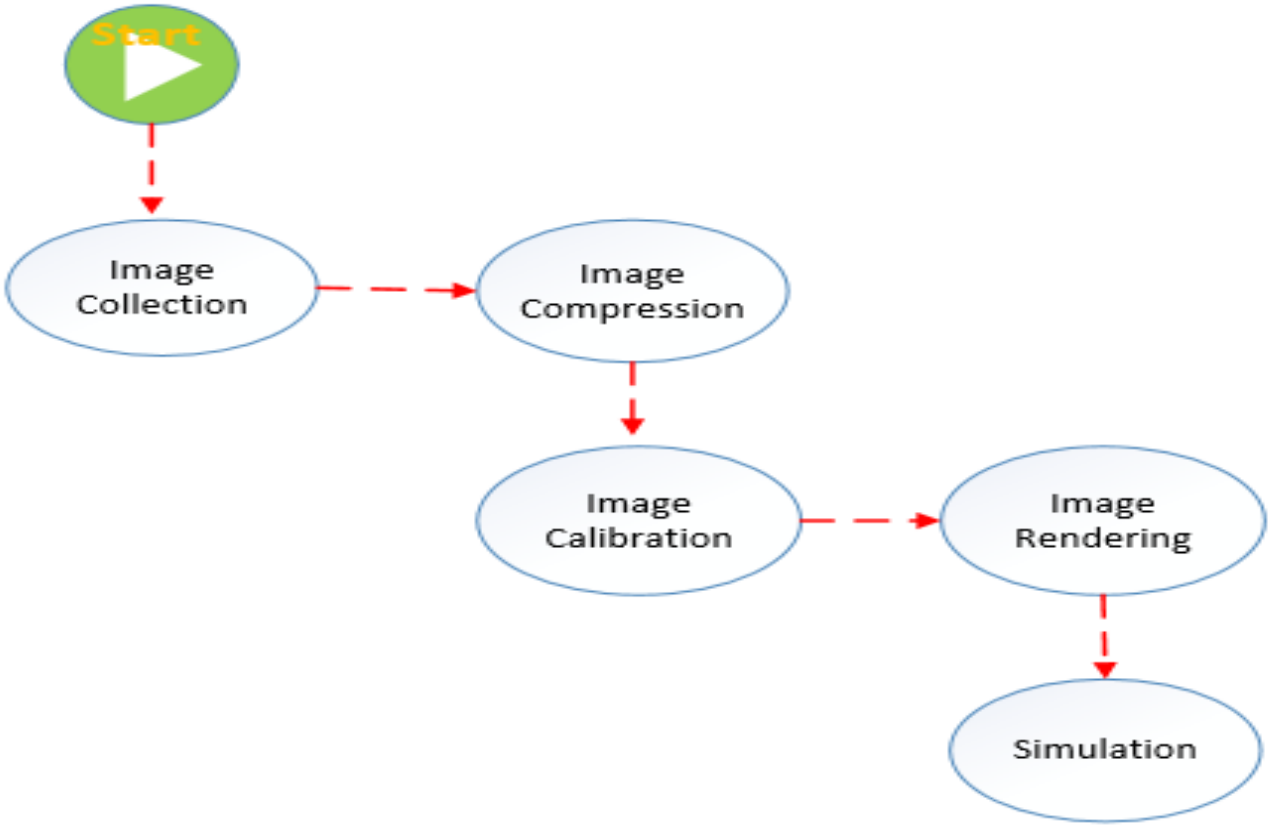

Figure 2: System model layout.

Figure 2 depicts the technical approach for the design and development of the simulator utilising the hexagon camera configuration model.

Image compression

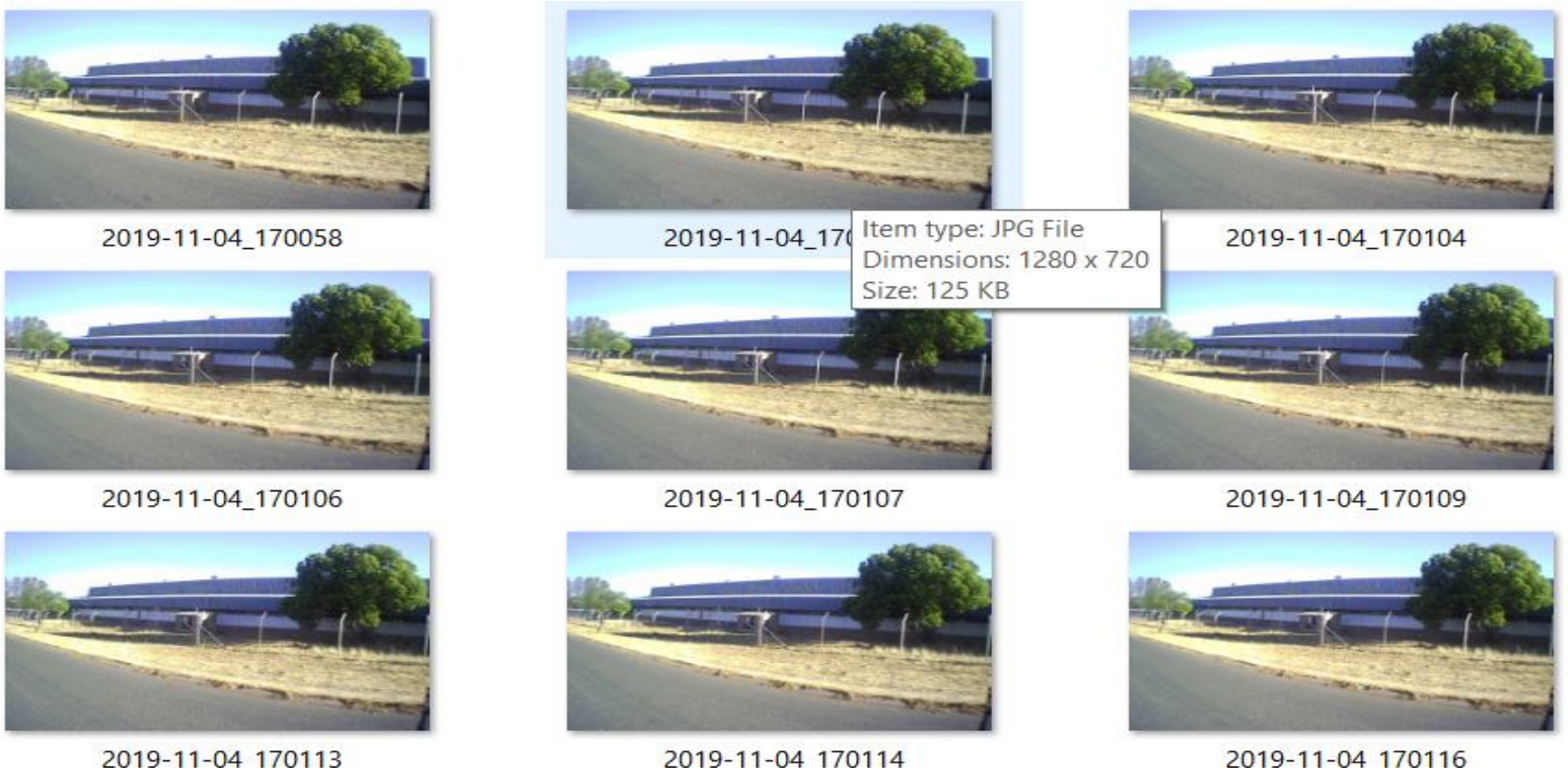

Figure 3: Image compression on a JPEG file.

Image calibration

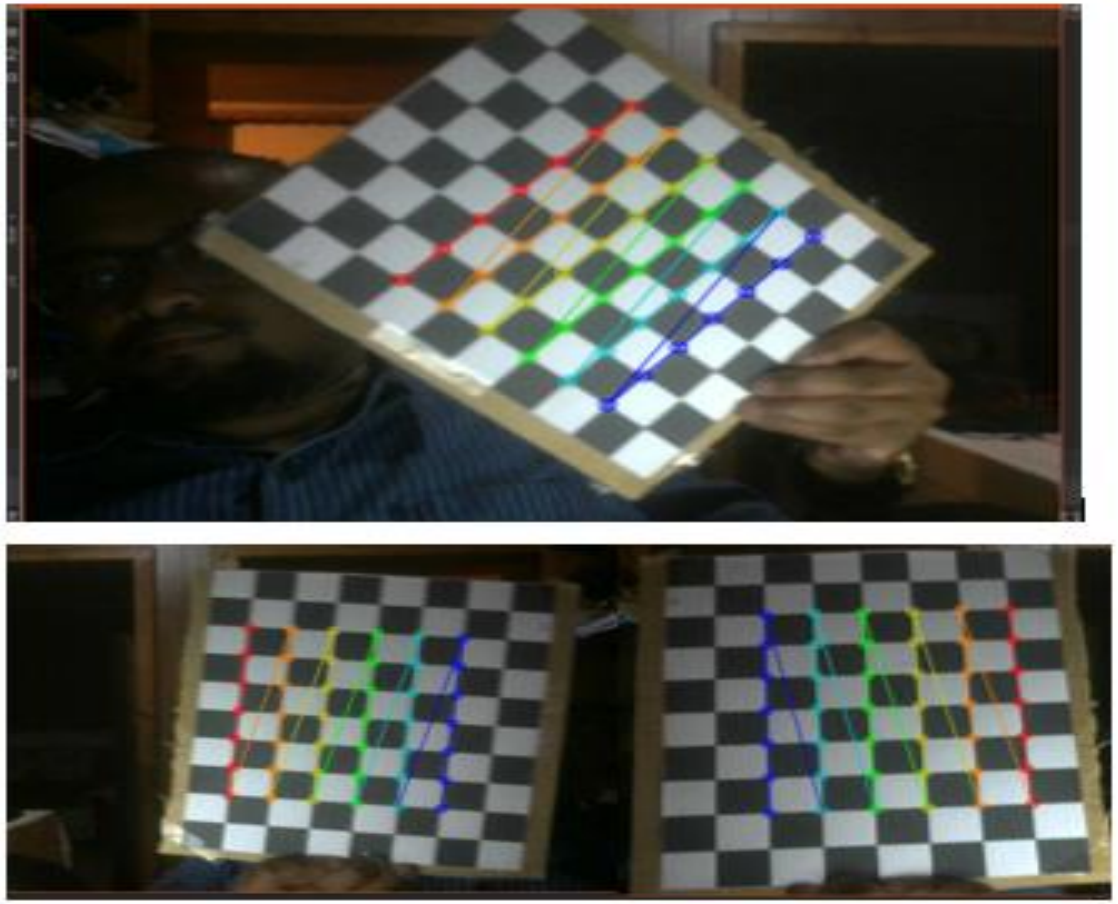

Figure 4: Black and white test match on a chessboard.

Structure for motion

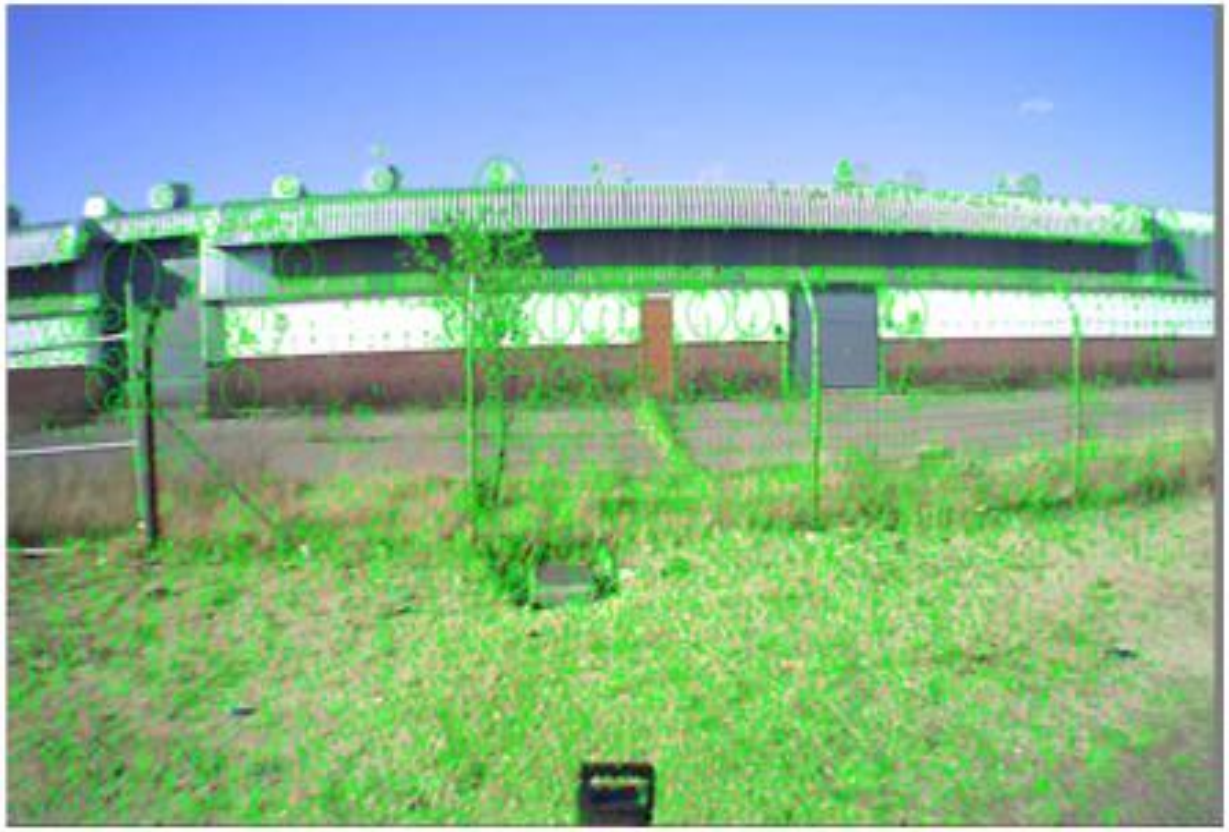

Figure 5: Detecting the image feature using SIFT.

- Figure 5 depicts the application of the Scale-Invariant Feature Transform (SIFT) on the original captured image.

Multi-view stereo

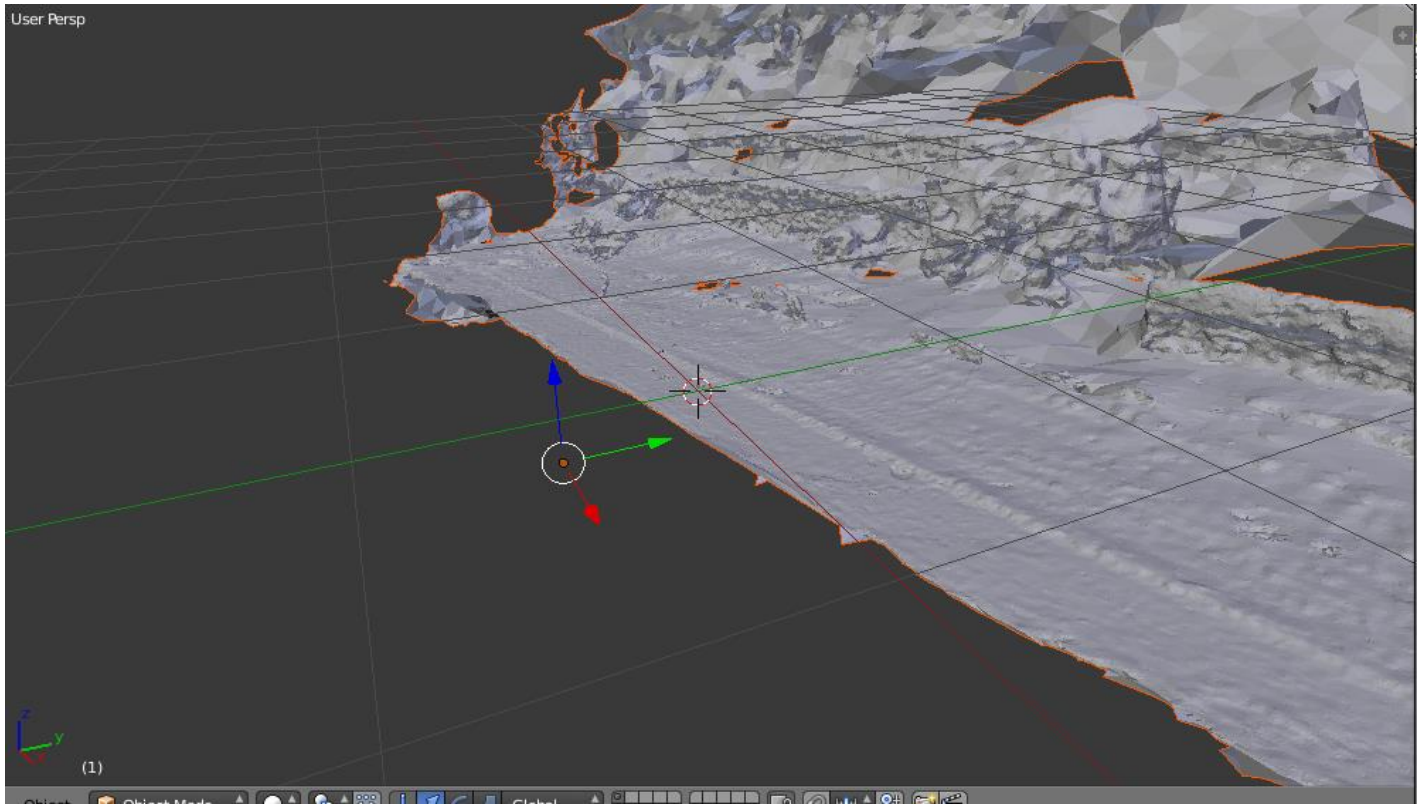

Figure 6 depicts the integration of patch-based matching that are used to generate point tracks and depth maps to generate the final reconstruction.

Figure 6: Mesh output from the multiview stereo.

Results

- The results outlined in this section depict the image rendering framework for the 364 .JPEG images that were captured on each camera at a total dataset worth 2184 JPEG images at a high resolution of 1280X720 pixels at a total size of 39.8MB.
- The system required a dynamic scene with 6 cameras arranged a 2D arc at a spanning of about 600cm apart from each other.
- Each camera frame comprised of 364 JPEG images were captured from the real scene, and only then the process of image matching and texturing is applied using “depth maps resulting with the output textured PNG image of 25.2 MB of 8192X9192 pixels resolution.

Rendering simulation outcomes

Figure 7(a): Denser point of cloud with Multiview stereo.

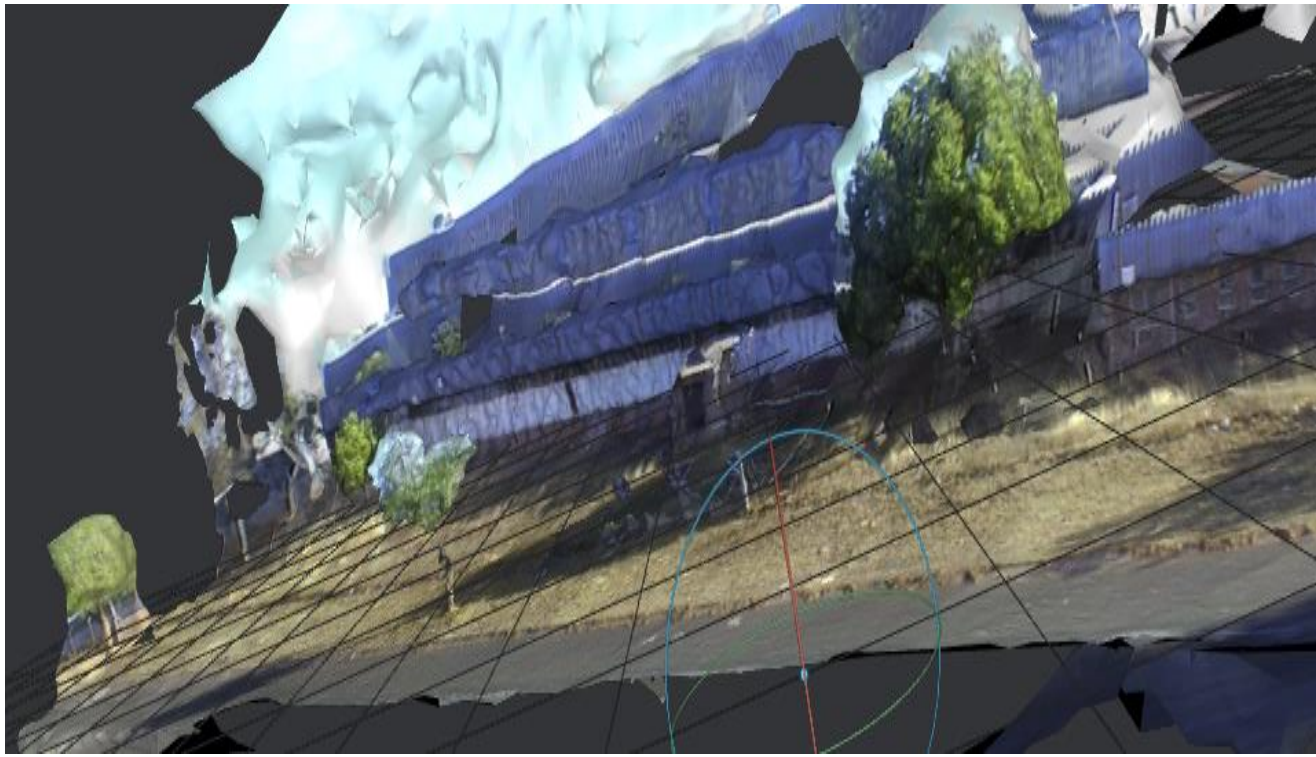

Figure 7(b): Generated texture.

Rendering simulation outcomes

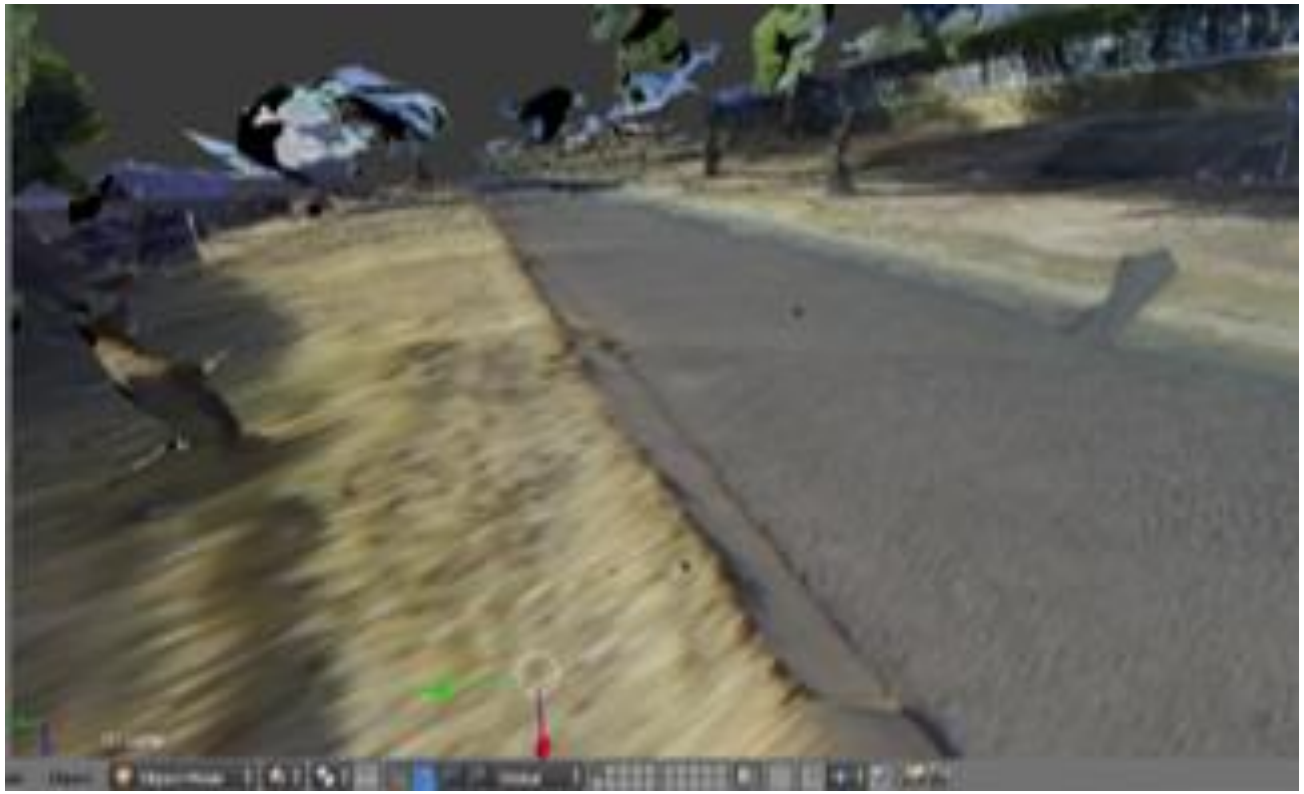

Figure 8: Rendered image horizontal view.

Figure 8 depicts the bidirectional panoramic street view achieved by projecting the horizontal street view in an omnidirectional manner.

These results are outlined to provide the freedom of movement and the views which were not captured by the camera but through rendering the uncaptured scene can be viewed.

Conclusion

The feature detection and matching technique was observed as the best technique in detecting and matching the images from multiple image datasets. As a result, The use of the image-based rendering technique utilising hexagon camera configuration was proposed as an ideal method in this study.

The objectives looking into the integration of IBR and simulator were achieved as indicated:

- The incorporation of the system into the simulation system in real-time for increasing the reality of the simulation system in different geographical locations;
- To simulate a rendering technique for improvement of visual, spatial, and quality of the panoramic images for location identification.

Questions