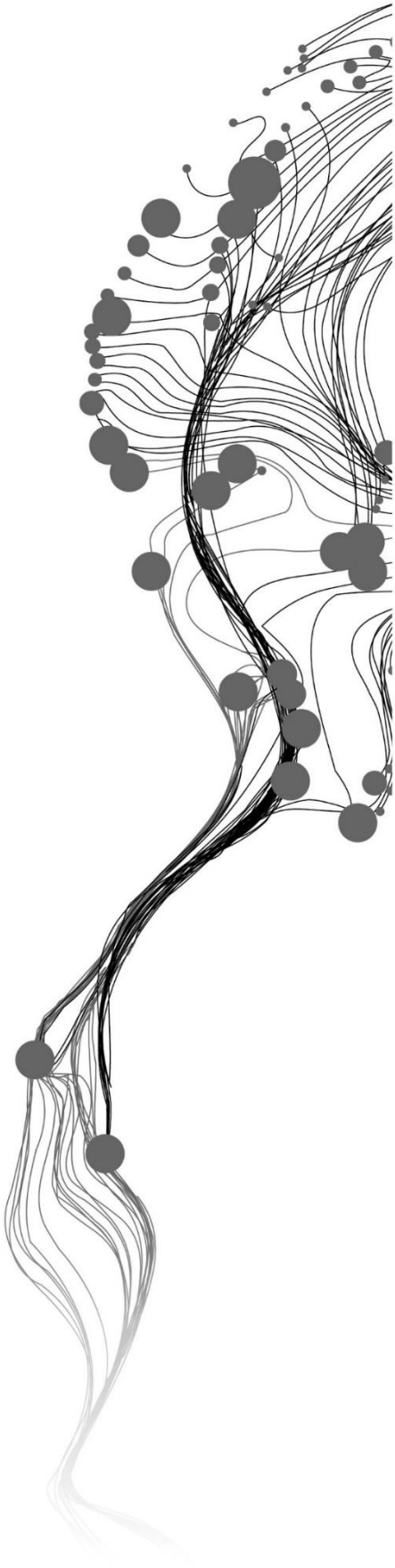
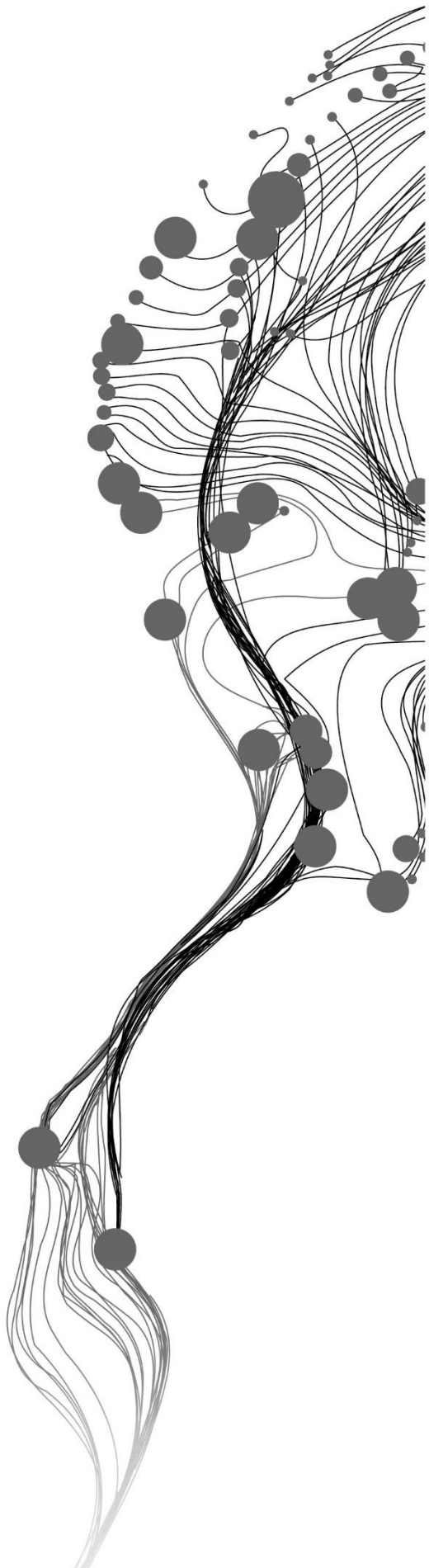


OPTIMISATION OF UAV FLIGHT PLANS FOR RELIABLE 3D BUILDING (LOD3) MODEL GENERATION

HARINISH KUMAR PALANIRAJAN

March, 2019

SUPERVISORS:
Dr. F.C. Nex
Dr. ir. S.J. Oude Elberink



Optimisation of UAV flight plans for reliable 3D building (LOD3) model generation

HARINISH KUMAR PALANIRAJAN
Enschede, The Netherlands, March, 2019

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

SUPERVISORS:
Dr. F.C. Nex
Dr. ir. S.J. Oude Elberink

THESIS ASSESSMENT BOARD:
Prof. Dr. ir. M.G. Vosselman (Chair)
Dr. M.N. Koeva (External Examiner, University of Twente, ITC-PGM)

ABSTRACT

Three-dimensional (3D) building model is gaining more scientific attention in recent times due to its application in various fields such as autonomous vehicle navigation, urban planning, heritage building documentation, gaming visualisation and tourism. The Level of Detail (LoD) of building models determines the number of details that are represented for each building. In this regard, LoD3 models both the roof and the facades structures. The survey undertaken with ALS, TLS and MMS devices for this purpose is often unpractical because of the expensive instruments and the cost involved in a crew. In such a scenario, Unmanned Aerial Vehicles (UAV) are efficient in collecting good quality images and generate reliable LoD3 of buildings at comparatively lower cost and time. This research proposes a novel methodology to generate the flight plan in correspondence of building facades. The Digital Surface Model (DSM) obtained from an initial nadir flight is used as an input to identify the target building and plan the image acquisition around it. The footprint of the building is initially obtained using a region-growing segmentation method. The footprint boundary traced from roof structures is then projected to the ground level, assuming the building facades are vertical. The building volume can, therefore, be defined, and a regular point grid can be defined on the building surface. The initial camera network design is planned along this boundary considering a pre-defined Ground Sampling Distance (GSD), Base/Distance ratio, overlap percentage with estimated theoretical accuracy for the block. This initial acquisition is planned to be redundant on purpose, adopting very short baselines. However, it would require a very long time to be collected, without significantly increasing the accuracy or the completeness of the generated 3D point cloud. Therefore, a filtering procedure is defined to reduce the number of images, assuring a homogenous coverage (i.e. number of visible images) for each point of the building. For this purpose, an iterative visibility testing has been implemented. The main constraint is given by the image overlap, considering a minimum number of images (n) on each façade point and removing the redundant images. Cameras classified as redundant for all the point on the facades are then iteratively removed based on constraints. The coordinates as well as the attitude of each planned image can be finally loaded on a UAV to perform the flight. The algorithm has been implemented in Matlab. Some tests were performed on synthetic buildings with complex building footprint shapes, in order to demonstrate the reliability of the implemented methodology

Keywords: UAV flight plan, camera network design, 3D building model, LOD3

ACKNOWLEDGEMENTS

Thanks to my parents and friends who always believe in me when I am in pursuit of my dreams. Their unconditional faith and love keep me going.

Thanks to my supervisor, Dr. F.C. Nex for the motivation and comments that he provided during this entire journey. His immense support and guidance during crucial phases of this research helped to finish the study in the best way. Thanks for reminding me of my strengths at times when needed.

Thanks to my chair Prof. Dr. ir. M.G. Vosselman and my second supervisor Dr. ir. S.J. Oude Elberink for the critical questions during the research phase that stimulated alternative thinking for solutions.

Thanks to drs. J.P.G. Wan Bakx for a countless number of reasons. Thanks to Mr. John Horn and drs. E.J.M. Dopheide, for the support.

Thanks to all the teaching faculty and staff at ITC for making the entire journey at ITC unforgettable.

And finally, thanks to my friends at Enschede who became an inseparable part of my life in this period. Thanks for all the memories you chose to share with me. Thanks for being there for me when I needed the most.

Thanks to all the smiles and good wishes from people who barely knew me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Motivation

Since the introduction of digital maps, the levels of detail (dimensionality) incorporated in it have consistently increased. 3D city models have become hot topic of research in the last decade as it finds application in autonomous vehicle navigation, urban planning, heritage building documentation, gaming visualisation, civil engineering, disaster management, rooftop solar potential estimation, tourism, augmented reality, etc (Döllner, Buchholz, Nienhaus, & Kirsch, 2005; Engel & Döllner, 2014; Kobayashi, 2006; Noor, Aiman, Bin, Abdullah, & Hashim, 2018; Themistocleous, Cuca, & Hadjimitsis, 2014). This field has been a common area of interest among Computer visions community and Photogrammetric experts for a long time, but a common standard framework was not specified for this field (Forlani, Roncella, & Nardinocchi, 2015). Finally, Open Geospatial Consortium (OGC, 2012) came up with the Level of Detail (LOD) scheme ranging from 0 to 4, which is widely followed in 3D city models. LOD0 represents 2.5D DTM, LOD1 is a block model, LOD2 is the textured roof model, and LOD3 has façade information, LOD4 is the interior of buildings (Biljecki, 2017; Kolbe, 2009). Figure 1 below depicts the Levels of Detail of a building from 0 – 4 according to standards defined by OGC.

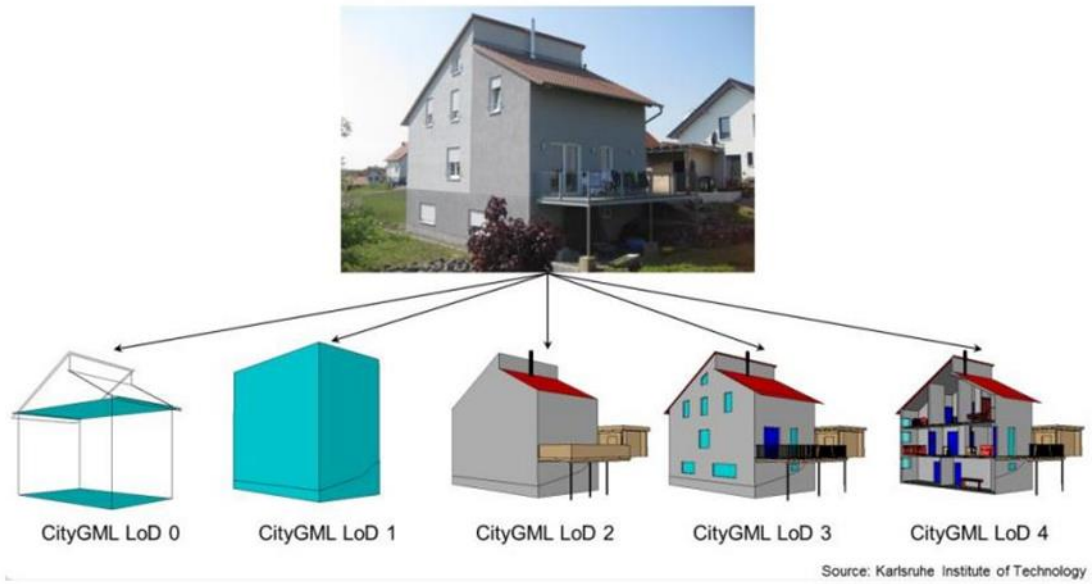


Figure 1.1 The Levels of Detail of building from 0-4 according to OGC standards

In recent times, the reconstruction of building structures with façade information became a fascinating field of study for survey services (Cefalu, Haala, Schmohl, Neumann, & Genz, 2017). 3D reconstruction of building models has always depended on the point cloud data from Airborne laser scanner (ALS) which gives accurate depth information (Emelianov, Bulgakow, & Sayfeddine, 2014a). Using ALS proves to be one of the best solutions for reconstructing planar roof structures (LOD 2). However, ALS cannot be useful to model the façade information (LOD3) of building due to the high altitude at which the data is collected. Besides, the survey undertaken using ALS seems unpractical because of the expensive instruments and staffing a pilot to fly a crewed aeroplane.

On the other hand, to reconstruct facades, Terrestrial and Mobile Laser scanners (TLS & MLS) can be useful. Unlike ALS, these scanners collect data from ground level. Therefore, these scanners cannot fetch details to map the roof structure due to self-occlusion in the line of sight by the building façade itself (Zakaria, Abdul Shukor, Basah, Rushforth, & Wong, 2015). Hence, to construct a 3D building model only using ALS data, predefined roof styles need to be used to fit for roof planes which are not accurate at times (Roo, Bourgeois, & Maeyer, 2017). Not to mention, the initial investment even on TLS could vary from 10.000EUR to 80.000EUR & MLS can be much more expensive. These instruments are heavy and may require more than one person to carry out the survey. TLS survey involves the selection of site location to get better coverage of target from different angles. So, moving around and setting up these heavy devices can be a tedious process. These factors increase the time and cost required for data acquisition (James & Quinton, 2014). Meanwhile, Unmanned Aerial Vehicles (UAVs) become handy in this situation as it can be flown at required proximity to the object of interest such as a building. The images acquired can be used to generate reliable LoD3 of buildings at comparatively lower cost and time.

UAVs are widely becoming popular because of their ability to collect high resolution close-range photogrammetric data at lower costs. The fact that sensors carried by UAV can be customised based on the needs of mission provides a strategic advantage over other platforms of remote sensing. Although initially the developments in this field were primarily focused on military requirements, the gates are open now to many civilian applications such as infrastructure documentation and inspection, building construction, road maintenance, mineral exploration, precision agriculture, monitoring forests, surveillance, disaster mitigation, traffic control, parcel delivery (Darmawan, Walter, Brotopuspito, Subandriyo, & I Gusti Made Agung Nandaka, 2018; Dong, Gary Burnham, Boots, Rains, & Dellaert, 2017; Emelianov, Bulgakow, & Sayfeddine, 2014b; Entrop & Vasenev, 2017; Erdelj, Natalizio, R. Chowdhury, & Akyildiz, 2017; Rakha & Gorodetsky, 2018; Roca, Lagüela, Díaz-Vilariño, Armesto, & Arias, 2013; Torresan et al., 2016; Vladimirovich Kim & Alekseevich Chervonenkis, 2015). The low cost and non-intrusive nature of UAV have made it a viable alternative solution for researchers to monitor and document sensitive archaeological sites also (Themistocleous et al., 2014).

Unfortunately, the payload limitations of UAVs did not support heavier 3D laser scanners in the UAV platform previously (Palazzolo & Stachniss, 2018). At present advances are made on lightweight LiDAR systems to be mounted on UAVs, but they are still expensive in the order of 50,000 to 300,000 USD (including IMU and GNSS) (Buczkowski, 2018). However, by selecting appropriate hardware and flight parameters using photogrammetry, the same quality of data can be obtained using commercial drones that are ten times cheaper.

Modelling of façade details such as windows, doors and other external elements in building are useful in autonomous vehicle navigation, virtual tourism, evacuation route planning, Lighting and energy analysis of the building, crowd management, post-disaster damage assessment, reconstruction of damaged archaeological sites (Díaz-Vilariño, Khoshelham, Martínez-Sánchez, & Arias, 2015; Vetrivel, Gerke, Kerle, Nex, & Vosselman, 2018; Vetrivel, Gerke, Kerle, & Vosselman, 2015).

3D reconstruction is the reverse process of photography. In photography, the 3D environment is typically captured in a 2D image. This results in the loss of depth information of the object. To reconstruct an object in 3D space, the camera captures the same object from different angles in 2D. These images after orientation and image matching techniques yield the desired 3D model of the object.

In order to model the building façade, oblique images of the scene gathered from UAV missions can be useful. The orientation of images acquired in a flight mission is possible by state-of-the-art Structure from

Motion (SfM) algorithm using feature descriptors such as SIFT or SURF for tie-point matching. Recent developments in image orientation, Dense Image Matching (DIM) techniques (Cavegn, Haala, Nebiker, Rothermel, & Tutzauer, 2014; Haala, 2013) paved way for generation of 3D models from the captured 2D images using SfM and image matching algorithm such as Multi-View Stereo (MVS) and Semi-Global Matching (Ahmadabadian et al., 2013; Besada et al., 2018a; Hirschm, 2007). Several commercial and open source software are available online for this purpose ("Acute3D," 2019; "Agisoft," 2019; "DroneZon," 2019; "insight3d," 2019; "MeshLab," 2016; "Pix4D," 2019; Furukawa, 2010; Wu, 2011). The dense meshed 3D point clouds generated by this process will serve as an alternative for ALS & TLS point clouds.

The quality of the 3D reconstruction heavily depends on the completeness and accuracy of dense point cloud generated (Mostegel, 2011). The point cloud comparison between Close Range Photogrammetry (CRP) and TLS showed that the difference in accuracy is considerably small (Harwin & Lucieer, 2012; Incekara & Seker, 2018; James & Quinton, 2014; Murtiyoso & Grussenmeyer, 2017). This is true when images are captured with ambient illumination conditions, and the surface is textured. Both CRP and TLS fail considerably in smooth surfaces due to lack of identifiable tie points and high reflectance respectively.

The reliability of 3D models developed from UAV images depends on the manual operations, the expertise of pilot and in some cases grid- based automation. However, this approach can result in gaps in the data coverage leading to inconsistency in the model (Martin, Rojas, Franke & Hedengren, 2016). A reliable 3D model can be created only when images acquired in the flight mission is of good quality and complete. Therefore, to ensure homogenous image coverage with same resolution on each element of façade and to acquire complete image sets without any discontinuity in the surface: an autonomous procedure is required to define the principal façade planes and completeness of the data. On the other hand, collecting redundant data over the same region increases the total flight time as well as the computation time during the image matching process.

"A Flight Plan is a time-ordered list of orders that a drone has to complete to fulfil the designed mission (i.e. take off, go to a waypoint, then hover to take a picture, then reach a second waypoint, then hover again to take a picture, finally land), while a trajectory is defined through the time-sampled dynamic state variation of the drone along time" (Besada et al., 2018b). Flight Plan along with other instructions of the mission form the complete Measurement Plan (Besada et al., 2018b).

As stated earlier, the flight plan will be composed of a number of viewpoints where the images will be captured. One of the classic problems in robotics is to identify optimal viewpoints for reliable 3D modelling, i.e. identifying next-best-view (NBV) in reference to the previous image that will have sufficient overlap. Technically, the number of such views can be infinite (Chen, Li, & Kwok, 2011). To solve this NBV problem, many heuristic approaches relying on dense point clouds are available. However, they are not highly suitable for complex scenes like an urban area where actual 3D structures are unknown (Huang, Zou, Vaughan, & Tan, 2017). Hence, it becomes necessary to determine a minimum number of images that are sufficient to reconstruct the complete 3D model of facades at a predefined resolution to assure maximum coverage and overlap in each viewpoint. These are the optimal viewpoints which constitute the optimal flight plan necessary to realise coherent data.

Addressing the objectives as mentioned above, we try to minimise the risk in autonomous UAV exploration without intervention from the operator for data collection. The motivation of the proposed study will be, to develop an algorithm that can produce an optimal flight plan for acquiring images to generate reliable and complete point cloud for 3D building model (LOD3). For an unknown area, the rough DSM generated from nadir images will be used as the initial input for the algorithm. From the DSM, the target building will

be identified, and the footprint is extracted. Based on the building footprint, the algorithm will compute the location of viewpoints (data capturing) and the camera sensor orientation for the flight mission initially. The initial camera network is planned to be redundant with very short baselines. The planned images are analysed, and redundant cameras which do not significantly improve the accuracy and completeness are filtered. By the iterative process of visibility testing and filtering based on coverage, homogeneity of each point in the scene (building facade) is ensured. This algorithm will help to reduce the number of viewpoints without compromising the quality of dataset thus saving time and cost during the data acquisition phase. The output of the algorithm will be the position and attitude of the camera for all images to be acquired. This algorithm will support the ITC application that is being developed for UAV flight missions.

1.2. Problem Statement

The proposed research will be, to develop an algorithm that can produce an optimal flight plan for UAV to acquire images which will generate reliable and complete point cloud for 3D building model (LOD3). For a given target building, facades need to be identified and modelled with acceptable accuracy and completeness. The collection of images for the building facades is usually performed by manual flights to assure homogenous coverage of each element with the same resolution. Therefore, this procedure is highly dependent on the expertise of the UAV pilot. Procedures like Next-Best-View (NBV) are computationally expensive as they figure out the new viewpoint online during the acquisition phase. Usually, these procedures collect an enormous amount of redundant data which will increase the cost of data acquisition and time of computation afterwards. No autonomous procedure to define the façade planes and acquire complete image sets before flying on the field has been developed yet. The optimal flight planning needs to be done based on user requirements of Ground Sampling Distance (GSD), overlap percentage, Base to Depth ratio to assure theoretical accuracy of the dataset. The output of the algorithm should be the position (X, Y, Z) and attitude of the camera (Ω, Φ, K) for all images to be acquired.

1.3. Research Identification

1.3.1. Research objective

The primary objective of the study is to develop an algorithm for determining the optimal flight path of image position for UAV to cover a 3D object, i.e. building. The optimal viewpoints computed by algorithm takes into consideration the homogenous coverage of the facades with the same resolution. The viewpoints will be sufficient to generate a complete 3D model of the building. Each element of façade when re-projected in images are visible in more than three cameras to assure accuracy and coverage. The algorithm initiates with rough Digital Surface Model (DSM) computed from nadir orientation images of the area as input data. The target building outline is extracted from the DSM. From the building footprint, the facades are determined and the images with short baselines are planned along these facades. The viewpoints are filtered based on coverage and redundant cameras are removed. This algorithm should give good geometric quality to produce point cloud with lower residuals for a given accuracy. The algorithm should consider the optimal distribution of images (i.e. a number of visible images for each point of the building) while minimising the time of acquisition by reducing the total number of images. The flight plan should consider occlusions and obstacles in the scene while defining the optimal viewpoints. The output of the algorithm will be the sequence of camera viewpoints with position (X, Y, Z) and the attitude of the camera (Ω, Φ, K) for all the images to be acquired. Thus, obtaining a reliable building footprint, efficient camera planning, ensuring completeness with a minimal number of images are the key aspects of this research.

1.3.2. Research questions

The research objective is divided into the following sub-objectives with corresponding research questions:

Sub-objective 1: The footprint of the target building should be extracted from the DSM input.

- What is the best method to delineate the boundary of the building?
- How to remove the obstacles like trees from the building footprint?
- How to define the main facades of the building?
- How to identify the interior and exterior corners of building for camera planning?

Sub-objective 2: The algorithm should plan efficient camera network design to capture the geometry which will assure the theoretical accuracy of the 3D model.

- What is the best method for path planning to map a 3D object?
- What are the photogrammetric parameters to be considered to minimise the error in the 3D model?
- How to ensure better coverage in the corners of the building?

Sub-objective 3: The images captured in flight mission should minimise the occlusions and assure the completeness of data.

- What is the best overlap to get façade details with minimum error?
- How to reduce occluded areas by varying different parameters in camera view planning?
- How the completeness of the dataset obtained will be assessed?
- How to filter the redundant images from camera network design without affecting the completeness and quality?

Sub-objective 4: The algorithm should be validated for different geometric shapes of the building to ensure the quality of the flight plan.

- What are the different building shapes to be considered to evaluate the algorithm?

1.4. Innovation aimed at

Previous studies have used TLS and MLS data for 3D reconstruction of façade which use expensive laser scanner hardware and require a lot of manpower for the survey. While the proposed study aims to obtain similar quality data at low-cost commercial UAVs. This study presents a novel approach to automatically determine the flight plan in correspondence to building facades using drones. Similar applications are seen to be developed in the field of robotics using customised drones which are very expensive. On the contrary, this algorithm can be used in any off-the-shelf drones. This makes the solution viable to many non-commercial users. The approach is specifically conceived for the LOD3 model, which is not seen to have been done earlier. From the literature that was reviewed, the visual odometry techniques in robotics are not implemented in the 3D modelling of buildings. The proposed study is trying to leverage the benefits of this method to reconstruct a 3D building model. At the moment, solutions like this one are not available for the intended field of study.

2. LITERATURE REVIEW

The overall process of UAV flight planning for image-based modelling of building involves the following sub-domains:

- 3D building reconstruction
- Voxelization for occlusion detection
- UAV flight plan

This chapter presents the existing state-of-the-art available in these fields of research.

2.1. 3D building reconstruction

The 3D reconstruction of an object involves two main steps: image orientation using Structure from Motion (SfM) and Dense Image Matching (DIM). SfM is the process of determining the camera orientation and extracting 3D tie points in the given set of images. The SfM typically relies on the feature descriptors ability to detect tie points between the images. The operators such as local Scale Invariant Feature (SIFT) (Lowe, 1999) and Speeded Up Robust Features (SURF) (Bay, Ess, Tuytelaars, & Van Gool, 2008) are useful to extract features from the corresponding images. However, SfM has various constraints such as overlap, texture, lighting condition, viewing angle which affects the final results. In smooth surfaces and repetitive patterns, it is difficult to find exact matches between the images. Therefore, outliers are filtered using RANSAC (Random Sample Consensus) (Fischler & Bolles, 1981). This is followed by the computation of the projection matrix for each image and bundle block adjustment of the entire set of images. This process tries to minimise the error between computed and observed image coordinates (Alsadik, 2015).

The reconstruction of 3D models of building with façade details (Level of detail 3) including doors and windows from high resolution images has become possible with combination of SfM methods to determine camera orientation followed by one of the dense matching algorithms such as Semi-Global Matching (SGM) (Hirschmüller, 2005), Patch-based Multi-View Stereo matching (PMVS) (Furukawa & Ponce, 2009), SURE (Mathias Rothermel, Wenzel, Fritsch, & Haala, 2012), Micmac, Agisoft Photoscan. The dense matching algorithms try to find the correspondence of each pixel in the overlapping region of the neighbouring image. In this way, the disparity map is computed for every pairwise image. Thus, the dense point cloud is obtained at the end of dense image matching. The potential to use DIM point clouds as an alternative to ALS is evaluated and documented in (Zhang, Gerke, Vosselman, & Yang, 2018). This study also found that the impacts of oblique images, illumination conditions, surface roughness, the contrast in the accuracy and precision of DIM.

In photogrammetric based 3D reconstruction, determining the interior and exterior orientation parameters is a primary step. One of the approach to reconstruct building from uncalibrated photographs was studied by Farinella and Mattiolo(2006). This involved the creation of a raw model of the building by the user using primitive blocks and defining a parent-child relationship to each block. The association between the photographs obtained at ground level to that of the raw model and intrinsic parameters of the camera are given as input. These data are used to infer the extrinsic camera parameters at the end of this semi-automatic reconstruction method.

A thorough understanding of various photogrammetric parameters before determining the location and pose of cameras to satisfy the geometric demands for the reconstruction of images is mandatory. The camera planning is a preliminary step which needs to be carefully taken up for the reliable and accurate 3D model. The accuracy of the model increases with optimal B/D ratio (Hasegawa, Matsuo, Koarai, Watanabe, & Masaharu, 2000; Mouget & Lucet, 2014), using convergent images, a number of images in which a point is

visible. However measuring the same point in more than four images provides the least improvement (Remondino & El-hakim, 2006).

Existing reconstruction pipelines work well in datasets which have numerous images with high redundancy. For 3D reconstruction purposes, it is not likely to acquire data with large redundancy as it will increase the cost and time during acquisition and computation phases. To overcome these problems, Daftry, Hoppe and Bischof (2015) proposed a closed-loop interactive process which integrates the acquisition process with online SfM feedback. In the same paper, the author also uses a multi-scale approach to tackle the scene drift problem which is the result of high-resolution vs accuracy trade-off. This method claims to improve the reliability of image-based 3D reconstruction as it acquires data in an incremental approach.

For the 3D reconstruction purposes, static images (still cameras) or moving video sequence can be adopted. In the latter case, it is easy to find correspondences frame-to-frame considering the sequence of images occurring consecutively. However, the problem arises due to short baselines to depth ratio (B/D ratio) creating intersections at angles that produce an error in depth. Besides the low radiometric quality of images obtained in motion are likely to infuse more errors (Alsadik, 2015).

This dense matching involves checking of all the images against all, which involves a considerable amount of computations. However, capturing images at a wide baseline to overcome the problem mentioned above will have further consequences in matching the repetitive features wrongly. Hence a reasonable number of images with appropriate B/D ratio must be obtained for reconstruction (Alsadik, 2015; Remondino & El-hakim, 2006).

Overall an optimal camera network design considering the factors such as the least number of images required, visibility of each element in 'n' images, the orientation & position of images, overlap percentage, B/D ratio, intersecting angles between images is crucial for 3D reconstruction. These factors will profoundly improve the coverage and accuracy while bringing down the cost of the project. The importance of camera network design is well documented by the works of Alsadik, Gerke and Vosselman(2012), (2013); Hoppe et al. (2012).

2.2. Voxelization for occlusion detection

Solving the visibility problem is one of the critical issues to be addressed in a photogrammetric project. In a point cloud data, it is necessary to identify the points that are visible from a particular camera viewpoint. To solve the visibility problem, several methods based on surface triangulation, voxelization are discussed in (Alsadik, 2015). Some methods involve identifying the surface normal of the object and camera optic axis to compute the visibility of a point based on a certain threshold on angle difference. Other methods involve the geometric intersection of a ray traced from camera viewpoint to the surface vertex. The closest vertex intersecting the ray in particular direction is considered visible from the camera. In figure 2, voxel-ray intersection method to detect the visible points on façade for a particular viewpoint can be seen.

In an image-based 3D model of the building, continuity and completeness can be disturbed due to the presence of gaps. The gaps may occur because of occlusions by nearby objects (trees, cars, neighbouring building), protrusion of part of the building (self-occlusion), lack of texture on the facade (due to glass windows), real openings (doors), insufficient coverage by camera network. These gaps are to be identified and corrected to get an accurate and complete 3D model. Nothing can be done for the problem due to texture. However, in camera planning phase without an initial sparse point cloud of the building, we can

estimate only the possible occlusion due to facades. The camera network can be improved in such occluded regions to get better coverage and avoid self-occlusion.

Figure 2.1 Visibility check using Voxel-ray intersection method by (Alsadik, 2015)

Volumetric representation (voxels) of the target building is a useful technique to accomplish the gap detection in 3D space. For a dense point cloud, voxels are created in 3D space for the entire area and filtered based on the presence of points as occupied or empty. Box-line intersection from the planned viewpoints helps to classify the voxels based on their visibility from each camera in the network (Alsadik, 2015).

In a novel approach, weighted voxels are used for 3D reconstruction of objects by Xie(2018). Instead of binary values, each voxel represents a weighted sum of neighbouring voxels as it helps to interpret the voxels contextually. This approach helps to preserve the geometric shapes of the object and helps to identify real openings.

Voxel segmentation is used in building detection from airborne LIDAR. In this method, building footprint and the approximated façade of the building is obtained by the author, Wang, Xu, Li and Zhao(2018). This method seems to be a promising alternative compared to TIN, raster grid and point clouds and octree because of its explicit 3D nature.

2.3. UAV flight plan

In general, Flight-path planning is aimed at lowering the cost and time required for the mission by identifying the shortest path, as we can see from the work of (Gao, Hou, Zhu, Zhang, & Chen, 2013). However, every flight mission has a specific primary objective based on the application field. E.g., 3D city modelling, traffic movement tracking, Cultural Heritage building documentation, parcel delivery has different goals to achieve with their constraints. Therefore, multiple flight planning algorithms exist catering to different needs of the mission. In this section, the flight plans which are developed by computer vision and robotics community that can be useful in data collection for 3D modelling is discussed.

In the 3D modelling applications, we need to consider many aspects like optimising the viewpoints to capture complete data and the waypoints to define trajectories for avoiding collisions. Previously, in robotics, vision-controlled path planners considered only the geometric information of the scene to find the viewpoints with minimal pose uncertainties (Costante, Forster, Delmerico, Valigi, & Scaramuzza, 2016; Wolf & Dewitt, 2013). This method often required 3D maps of the environment in forehand. Whereas in Perception-aware path planning, UAVs explore the unknown area and new trajectory is computed from the texture and geometric structures in collected images on-the-fly (Costante, Delmerico, Werlberger, & Valigi, 2017). This can be a key to achieve high-quality datasets as it considers the photometric information (texture) in generating the flight path.

To reconstruct a 3D structure, the Next Best View (NBV) algorithm was suggested by Vasquez-Gomez, Sucar, Murrieta-Cid and Lopez-Damian (2014). This method identifies several candidate views based on search paradigm and evaluates based on the predefined constraints to determine the best view. This research claims to have resolved inaccuracies in positioning while the search strategy is 20 times faster than traditional ones. In the latest developments (Bircher, Kamel, Alexis, Oleynikova and Siegwart (2018) propose another receding horizon NBV scheme. The best branch is determined from a random tree based on the unmapped search space. In each iterative step, only the first edge of this branch is explored. In this way, the entire area is covered by repeating the process (Kriegel, Rink, Bodenmüller, & Suppa, 2013). NBV schemes for 3D volumetric reconstruction of an object using an active vision approach is presented in (Isler, Sabzevari, Delmerico, & Scaramuzza, 2016). In this method, the probabilistic volumetric map is used to optimise the information gain, likelihood of visibility, robotic movements while maximising the likelihood of viewing new parts of the object for each set of candidate views. This is performed based on online feedback from the partially constructed model at each step.

Fast Marching Square (FM2) method is a path planning algorithm that generates a flight plan for Terrain Following Flight (TFF) while avoiding obstacles. This method can be useful in non-homogenous terrain such as built up environments to have smooth and safe trajectory as it is evident from the experiments carried out by González, Monje, Moreno and Balaguer(2017). The two adjustment parameters help to control the flight level, smooth and safe path.

Sampling-based path planning algorithms such as Rapidly-exploring Random Trees (RRT) and Probabilistic RoadMap (PRM) have shown high theoretical consistency in probabilistic completeness. However, the extensive research by (Karaman & Frazzoli, 2011) on the asymptotic behaviour of these methods conclude that solutions obtained from these widely used algorithms are not asymptotically optimal while they are probabilistically complete. This means that the solution to the path planning problem will not converge to optimal cost when the number of samples increases. The optimisation problem falls in Kino-dynamic planning category with constraints of completeness and obstacles in the trajectory.

In the Urban environment, flying at low altitudes for collecting high-resolution imagery comes with its price when the GPS signal is lost. In such cases, localising the pose of flying UAV and the camera orientation is difficult. To find the position, Air-ground matching technique as suggested by (Majdik, Verda, Albers-Schoenberg and Scaramuzza(2015) uses the approach of matching airborne images acquired by drones with ground-level geotagged images database by back-projecting in the 3D model. A similar approach is seen to be documented in (Scaramuzza et al., 2014).

UAVs allows carrying multiple imaging sensors to capture data at different angles. This was demonstrated by the work of Feifei, Zongjian, Dezhu and Hua(2012), in which the author combines four combined cameras with a particular design to improve overlap and self-calibration by avoiding the errors due to mechanical deformation and time lags in each exposure. By combining nadir and oblique images, buildings were mapped at 1:1000 and 1:500 scale with high precision in this case.

From the literature, it is clear that UAV path planning is performed online in most of the cases. This is a computationally intensive process, and data acquisition is redundant in many cases due to the lack of confidence in information gain of the views planned further. Hence, these methods consume more time to collect data which may be redundant at times. Therefore, offline flight planning from a rough DSM considering the various photogrammetric parameters which ensure completeness and accuracy of 3D reconstruction is a novel and practically-feasible approach.

3. DATA AND SOFTWARE

In this study, the DSM prepared from the UAV images acquired over the city of Dortmund, Germany on May 19th, 2014 was used (Nex et al., 2015). Two different UAV platforms were used in that data collection. Multi-rotor UAV DIJ S800 and Mavinci fixed-wing plane. The images were captured using Panasonic GX-1 on Mavinci platform and Sony Nex-7 on DIJ S800 with GSD ranging from 1 to 3cm. The DSM prepared with this dataset had a GSD of 2.25cm. Figure 3.1 shows the orthophoto of the region and the DSM of the same. Two buildings were able to be identified from the DSM. Other buildings that are seen in the image are not fully covered in the images. Therefore, they were not used in camera planning. In order to evaluate the algorithm with different shapes of the building, synthetic buildings were used. These shapes were designed to check and improve the algorithm for view planning, visibility analysis, camera filtering to check various conditions such as self-occlusion, the effectiveness of convergent cameras on the corners, etc. Some of the shapes of the simulated buildings used in the study are shown in figure 3.2.

Figure 3.1 Ortho photo (left) and DSM (right) of the study area in Dortmund, Germany

The code for the entire process was done in MATLAB R2018. This was preferred since the study involved many matrix data to process. This also had the advantage of having numerous tools for image processing and in-built functions which became handy at times. Visualization of the data and final results were easier to perform using these in-built functions and tools.

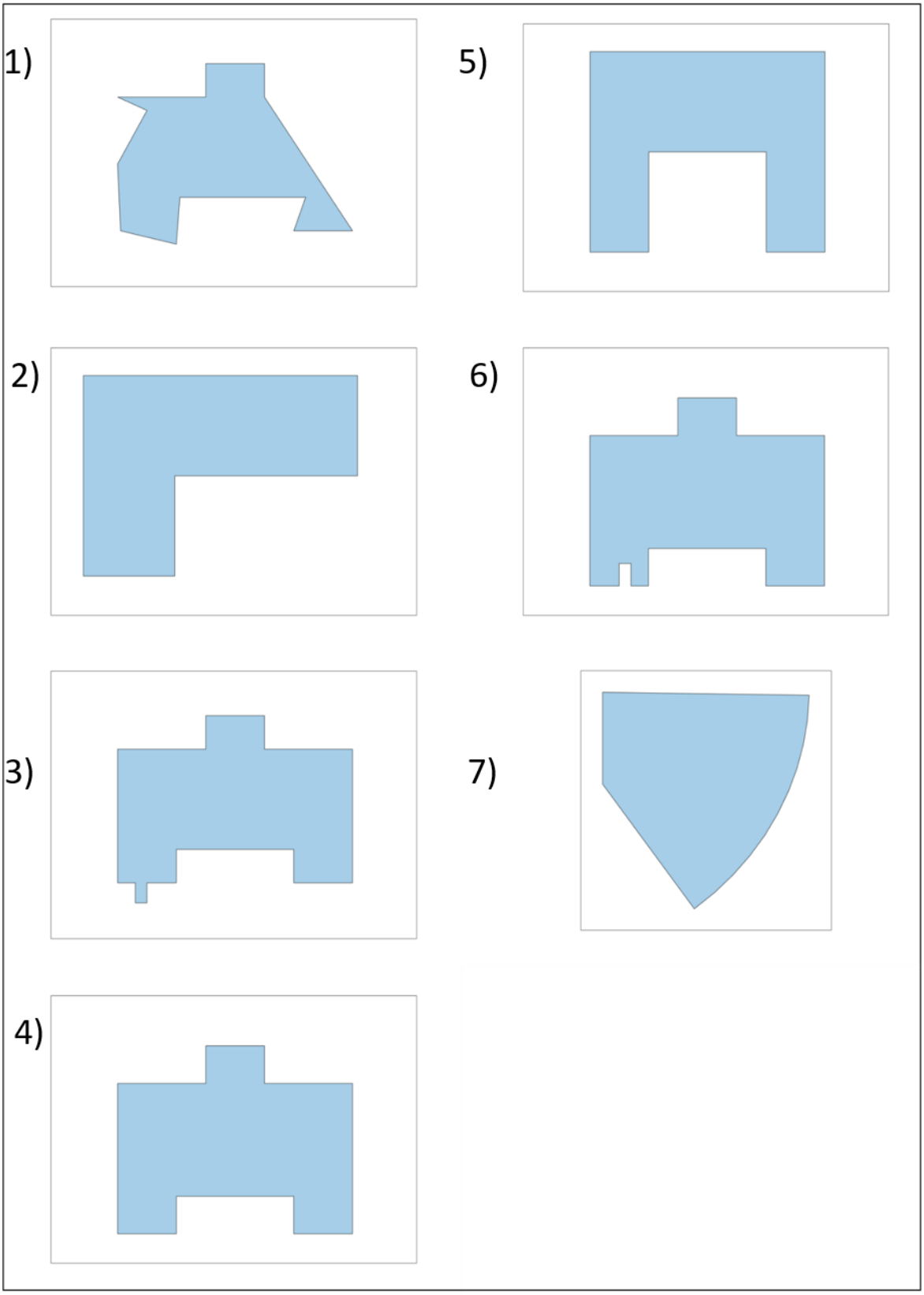


Figure 3.2 Different shapes of simulated building used for development and testing of algorithm

4. METHODOLOGY

The key processes involved in this study are detailed in this chapter. The chapter is sub-divided into the following five sections:

4.1. Building extraction from DSM
4.2. Façade definition
4.3. Initial dense camera network
4.4. Visibility analysis
4.5. Filtering based on coverage

Each section describes the methods used in the research for solving a particular problem. The key idea was to establish the building facades from the building footprint using DSM and plan image acquisition along the facades for 3D reconstruction. Figure 4.1 briefly outlines the overall steps involved in the process.

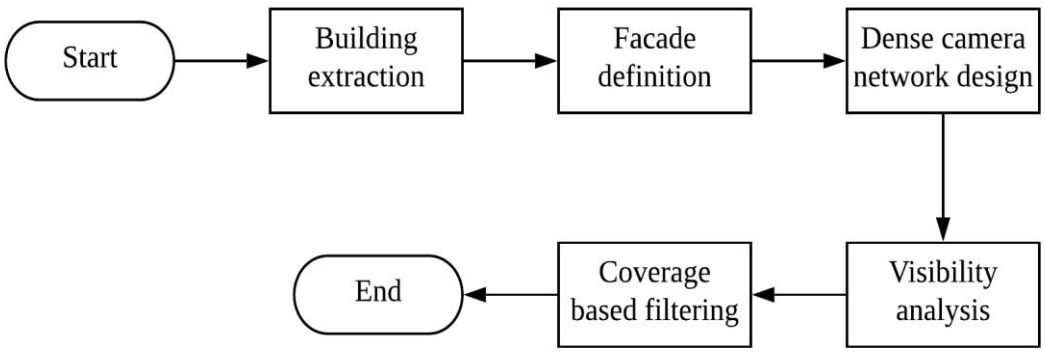


Figure 4.1 The general overview of steps involved in the methodology

4.1. Building extraction from DSM

The process aimed to identify the building footprint to be used in the camera network design for traversing along the boundary of the building. The initial step was to identify the target building in the DSM. Region growing segmentation method was used to extract the segment of the building area. Green Red Vegetation Index (GRVI) computed from the orthophoto of the region was used to remove the trees if present in the building segment. The step was followed by morphological operations on the building segment to close the openings. The boundary pixels of this final segment was traced to extract the outline of the building. The building boundary was further regularised to remove the outliers. Figure 4.2 explains the process involved in the extraction of the building footprint. A detailed explanation of these steps is reported in the following sub-sections.

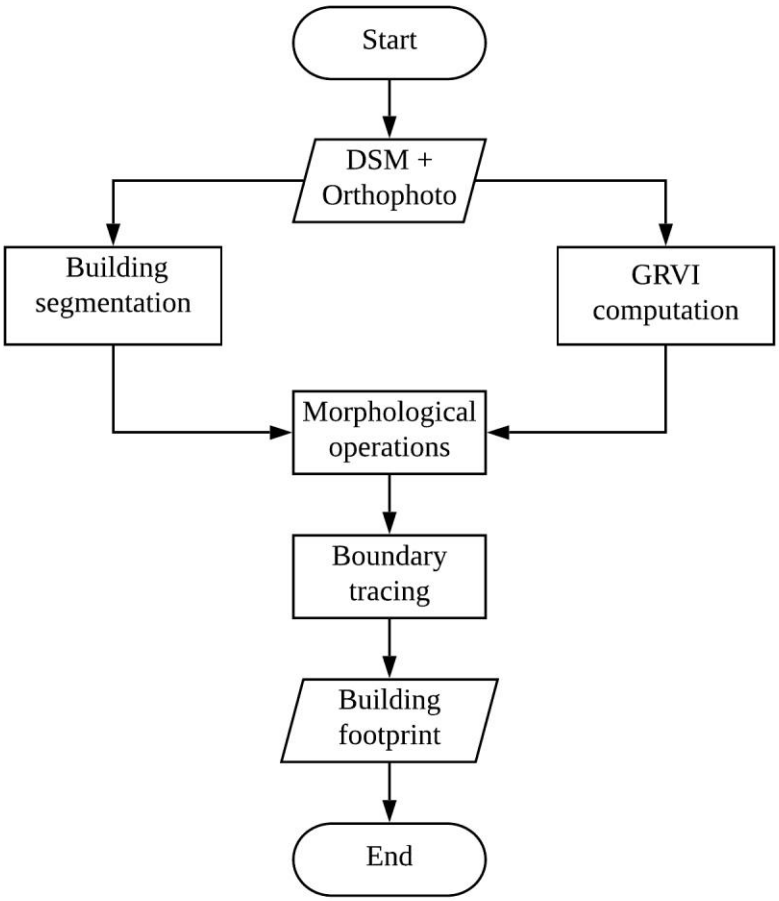
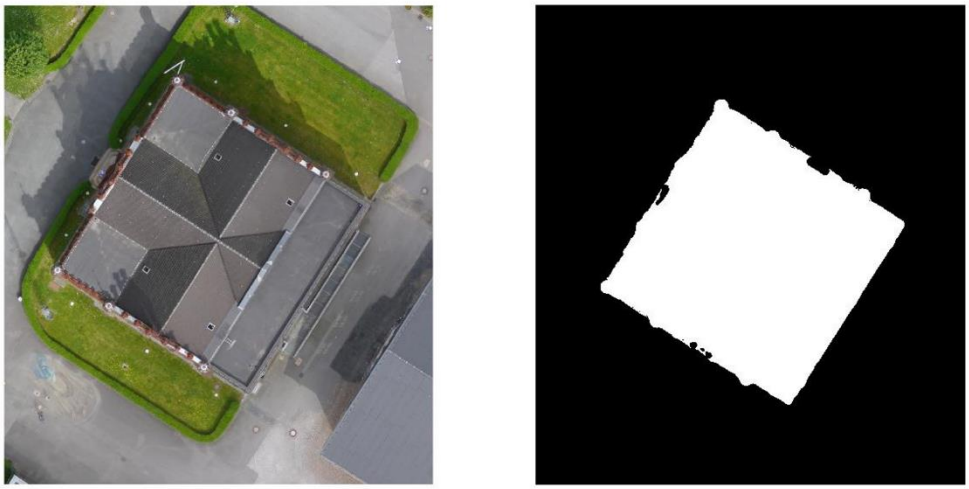


Figure 4.2 Steps involved in building footprint extraction

4.1.1. Building segmentation

To segment the target building from DSM, a seed pixel on the roof surface of the building was selected by the user in an interactive approach. With the height data of seed pixel as the mean value of the segment, the neighbouring pixels were checked to find the closest value to the mean. The pixel was added to the region only if it satisfied the height threshold criteria. Initially, this threshold was set to 5m which is the height of two floors of the building approximately. This was done to identify the different parts of the building with varying heights such as ITC building, Church towers as a single segment. This threshold can be altered if the user finds the jump in height to be greater. In this way, the entire building surface pixel was added to the region until there was no pixel found to be closer to the mean of the region i.e. due to a steep fall in height due to the ground next to a building pixel. This region-growing segmentation was performed with the code developed by (Kroon, 2008). The result of this process was a binary image with the building segment as value 1 and the rest of the pixels as 0. The building segment is shown in figure 4.3 (b)



(a) (b)

Figure 4.3 (a) The building in orthophoto (b) Segmentation of the building

4.1.2. GRVI computation

From the ortho-mosaic of the region, Green-Red Vegetation Index was computed to exclude trees close to the building in the segmentation output. GRVI is one of the spectral indices used for identifying vegetation in RGB images. It is the ratio of the difference between Green and Red bands to the sum of the same bands. It is given by the equation:

$$GRVI = (\beta_{Green} - \beta_{Red}) \,/\, (\beta_{Green} + \beta_{Red}) \qquad (1)$$

The reflectance of the canopy is greater in β_{Green} than β_{Red}. The reflectance of soil particles is lesser in β_{Green}. Elements like water have almost equal reflectance in both these bands. Therefore, by this equation, GRVI > 0 can be identified as vegetation and GRVI <= 0 can be interpreted as non-vegetation pixel (Motohka, Nasahara, Oguma, & Tsuchida, 2010). Based on this threshold, a binary image of the region with vegetation pixels as value 1 and other pixels as value 0 was obtained. Figure 4.4(b) shows the detection of trees in the building neighbourhood.

(a) (b)

Figure 4.4 (a) Vegetation near the building (b) Vegetation identified using GRVI

4.1.3. Morphological operation

Image subtraction was carried out between the binary image from the building segmentation and the GRVI output. This step resulted in the elimination of trees if present in the building segment. The segment was binarized to remove values less than 0 because of the subtraction performed earlier. At this stage, there can be jagged edges due to the removal of some vegetation pixel. This was corrected by performing morphological closing.

Morphological closing is the sequence of dilation and erosion process to improve the continuity of boundary pixels by filling in small gaps and removing the outliers in the edges. As the values in the image are binary, the operations are performed with a flat structuring element of the 3x3 window. Only the true pixels (value =1) are included in the morphological computations. The closing of the building segment helped to get rid of cavities and jagged edges due to missing and hanging pixels respectively. In this way, the building segment was fine-tuned to produce a regular geometric shape of the building footprint.

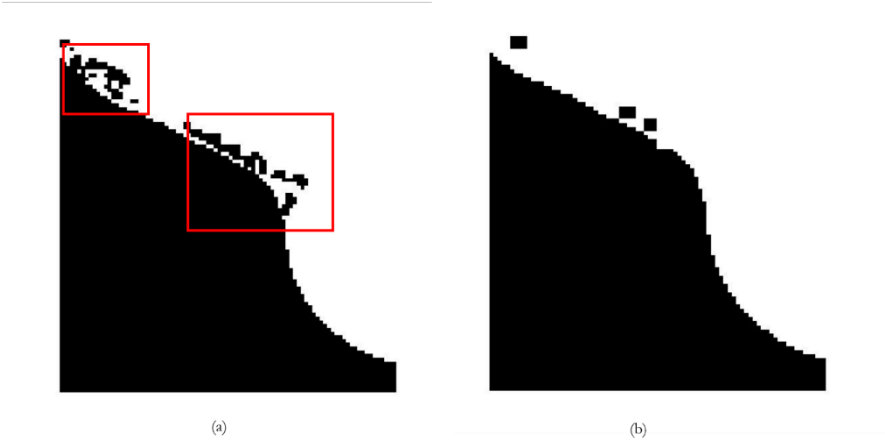


Figure 4.5 (a) Boundary before morphological closing (b) Boundary after filling the gaps by morphological closing

4.1.4. Boundary tracing

The contour of the building segment was traced using Moore-Neighbor tracing algorithm with Jacob's stopping criterion (C Gonzalez, E Woods, & R Masters, 2009; Ghuneim, 2000). In a tessellation containing a connected segment, the Moore-Neighbor tracing algorithm starts checking from left to right column from the bottom to top row to identify the start pixel. The start pixel is labelled; saved as in boundary pixels and the contour tracing begins. The neighbouring 8 pixels are checked in the clockwise direction until it finds the next pixel in the boundary. The next boundary pixel is saved in the list and it backtracks to the pixel from where it entered the current boundary pixel. The tracing process is repeated until the start pixel is visited again. The stop criterion for the algorithm is to terminate when the start pixel is visited for the second time. But this will lead to termination before tracing the entire pattern in a large complex segment. This can be seen in figure 4.6 (b) The stopping criteria was modified by Jacob Eliosof known as Jacob's stopping criterion. The algorithm terminates only when the start pixel is visited for the second time in the same manner as it visited the first time. Otherwise, the contour tracing continues to look for the neighbouring pixel. This can be seen in figure 4.6 (c). By checking all the 8 connected pixels, the algorithm ensures that the entire outer boundary is identified. The end result of this process is a sequence of coordinates (x, y) of boundary pixels.

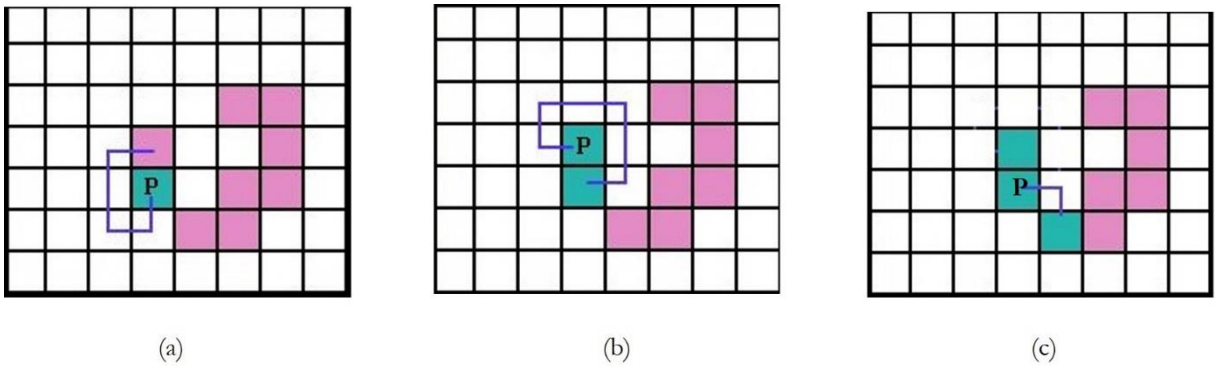


Figure 4.6 (a) Tracing begins from Start pixel (b) Tracing terminates without Jacob's stop criterion (c) Tracing continues to detect full pattern due to modified Jacob's criterion (Ghuneim, 2000)

4.1.5. Regularise the boundaries

The boundary pixels needed further processing to retrieve a regular polygonal building shape. Therefore, the outlier vertices were removed by relative distances of the vertices in the boundary. A tolerance threshold of 3 was used to remove the outliers. This process resulted in straight line boundaries ensuring the building shape. Figure 4.7(a) shows a section of boundary after removal of outlier vertices. However, there was numerous vertices for a single straight façade of the building which was not very useful to define an approximate boundary that can be used for flight planning. To generalise the boundary, vertices that were within 2-meters apart were filtered. However, this limit can be decided by the user, on inspecting the segment outline. For example, if the removal of trees along the boundaries resulted in irregular boundaries. Reducing this threshold to a lower value for filtering can preserve the shape along the façade. Regularising the boundaries helped to overcome the difficulties in the segmentation process. Tall towers that were twice as high as the main building itself resulted in cavities in the boundaries in the building segment. Figure4.7 (b) shows a section of boundary after filtering vertices <2m that retrieved close to real boundary including the cavities due to tall towers in segmentation.

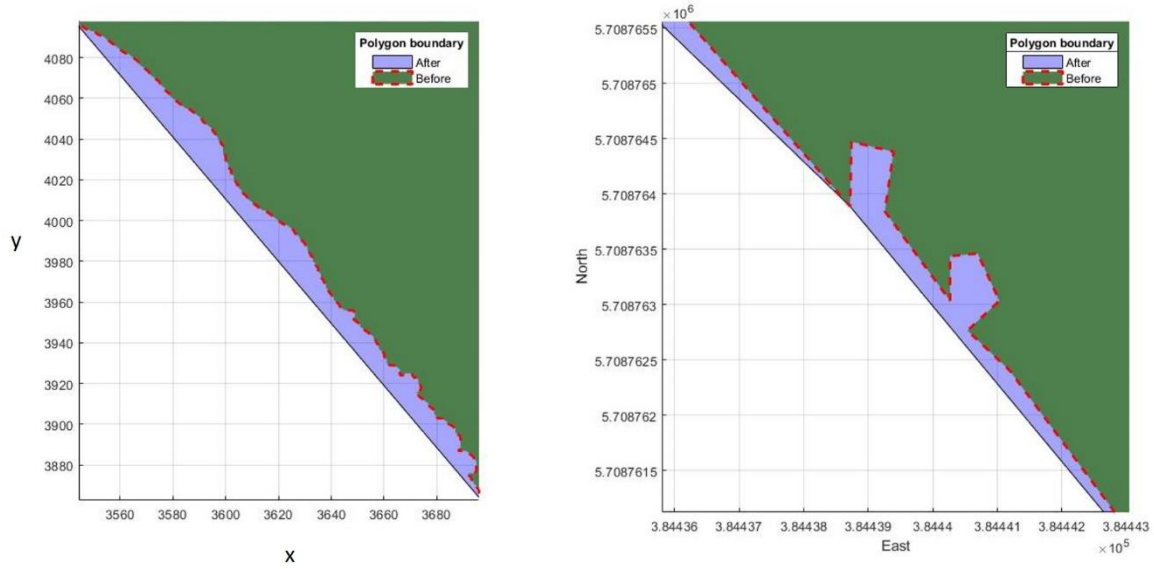


Figure 4.7 (a) Boundary after removal of outlier vertices (b) Improvised boundary after vertices <2m apart are filtered

4.2. Façade definition

The facades of the building must be defined in terms of length and orientation for interpretation. This helps to plan the exterior orientation of the camera network to capture sufficient images by traversing along these facades. The corners of the façade need to be clearly defined in order to plan a convergent camera network for connecting the facades. Hence the important steps in this process were to define facades, distinguish the building corners, generate rough 3D point grid to represent each façade.

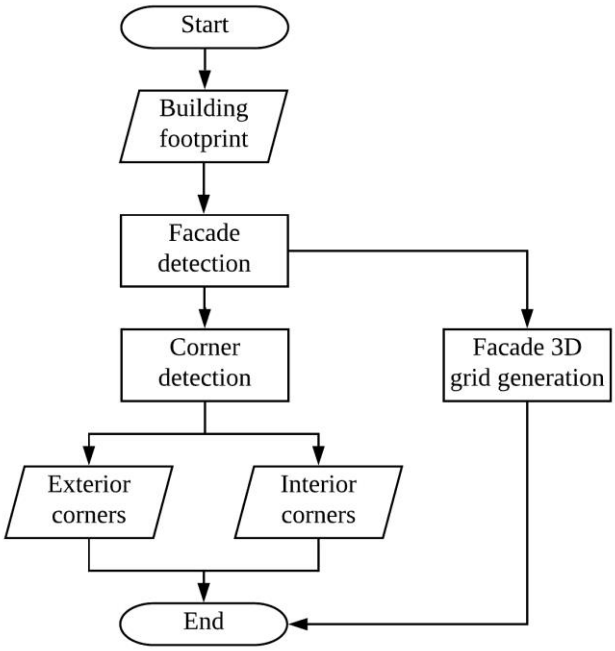


Figure 4.8 The workflow of building facade definition

4.2.1. Façade detection

The boundary vertices obtained from the previous step represents the façade outline. Since the boundaries were filtered with a threshold of <2m, a line segment greater than or equal to 2m can be distinguished from these vertices. Therefore, a façade of a minimum length of 2m were able to be identified for capturing in the camera network design. For a typical 'U'-shaped building, the walls running inside from main façade plane can be differentiated as a separate façade if it is greater than 2m and the image acquisition for that surface can be planned accordingly. This can be seen in figure 4.9 (a). The orientation of each façade with respect to North (0º) was computed in a clockwise direction, i.e. East = 90º; South = 180º; West = 270º as shown in figure 4.9 (b)

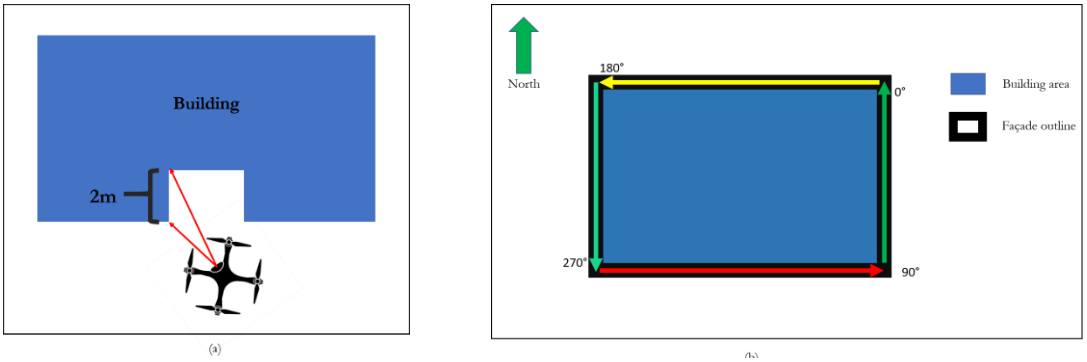


Figure 4.9 (a) U-shaped building with facade >2m (b) Orientation of facades of building

4.2.2. Corner detection

The building corners need to be identified to plan the connecting cameras, so that transition from one façade to the other is smooth. In the real world, not all the facades make right angles to each other as shown in figure 4.9. Therefore, from the boundary vertices, the true corner locations need to be detected. The corners are further classified into Exterior and Interior based on the angles made between the two facades. This classification helps to plan the camera network according to the type of corners. The facades making angles between 30º and 160º are classified as interior corners. The facades making angles between 190º and 350º are classified as exterior corners. Thresholding the angles in this way, the vertices lying on the same façade were eliminated from being misunderstood as corners. Figure 4.10 shows the synthetic building with interior and exterior corners as defined

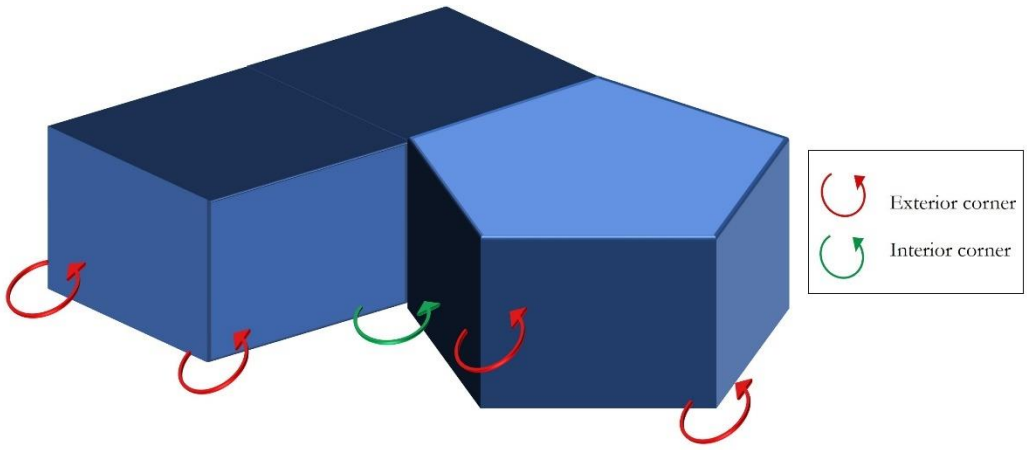


Figure 4.10 Exterior and Interior corners of a synthetic building

4.2.3. Façade 3D grid generation

In order to ensure that every feature on the surface of façade is captured in the 3D model, the 3D grid representation of each façade was necessary. These 3D points were used for checking the visibility of façade points from the cameras. For 3D grid generation, apart from the boundary vertices that were traced, approximate height of the building was necessary to represent the entire volume of the building. The height of the building is extracted from the DSM by getting the maximum Z value within the building segment. The lowest point of the building was obtained from the neighbouring pixels outside of the building segment corresponding to the ground surface. The minimum Z value from these pixels was assumed to be the ground surface, over which the building stood. Assuming the walls to be vertical, the boundary vertices were projected to the ground level. The points were computed for the surface of facades in a regular grid interval. Depending on the required point cloud density for the façade, the parameter s_grid can be modified to produce the desired number of points per sq.m. Figure 4.11 shows a section of building façade represented as a 3D grid. The regularised building outline obtained from the DSM gives the X, Y positions of the building footprint. The highest point within the building segment is taken as h2. The lowest point of the building is at the foot where it meets the ground surface. Therefore, some of the neighbouring pixel to the boundary is selected and from that, the minimum value was found which is h1. From this, h1-h2 gives the approximate height of the building. From the boundary outline, equally spaced points are generated based on the value of s_grid according to the required density of point cloud. This is extended to Z axis from

bottom to the top of building in a sequence, from the height determined by h1-h2. In this way, the façade of the building is represented as a grid of points.

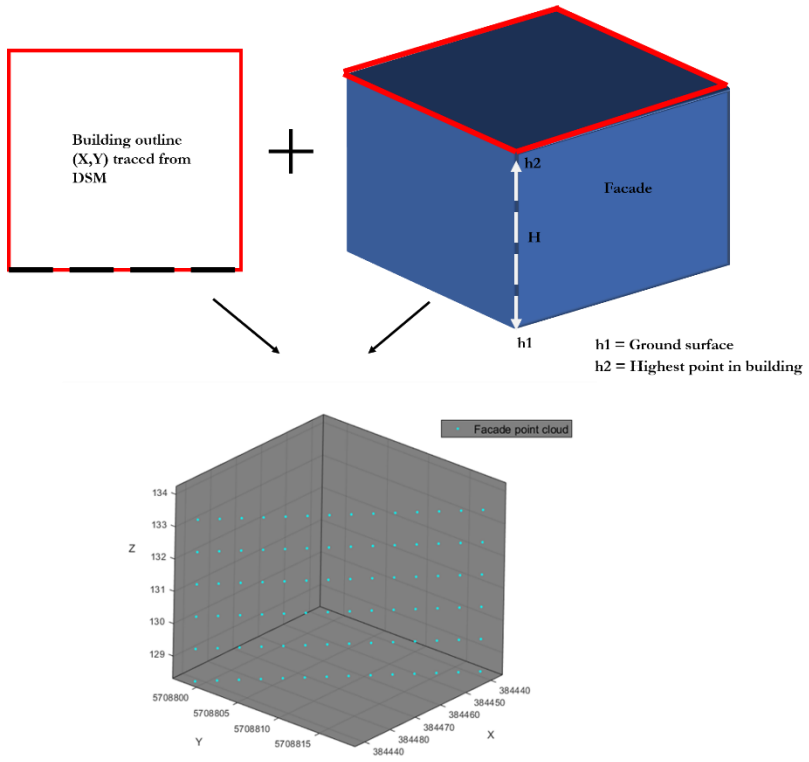


Figure 4.11 Facade grid generation from boundary vertices and height of building

4.3. Initial Dense camera network design

In this section, the planning of the initial placement of the cameras along the facades is described. The camera network is designed to be redundant on purpose, not to lose coverage in any region on the façade. However, the collection of such redundant images can be time-consuming without improving the accuracy or completeness of the final 3D model. Therefore, the cameras were filtered based on coverage in the next step. This kind of reduction from the maximum observation scheme in photogrammetry is suitable to arrive at the optimal number of images to reconstruct the object (Agarwal et al., 2011; Alsadik, 2015). The overview of key steps involved in initial dense camera network design is described in figure 4.12. The input for the camera planning is the user's choice of the resolution, camera interior parameters and façade definition of the building obtained from the DSM. From this information, the dense camera network along the façades was computed. If there was any camera, too close to the façade, they were removed for the safety of UAV. This dense network plans special connecting cameras between the facades of the building. To ensure coverage of every point in the surface, to produce a 3D model the basic photogrammetric parameters need to be considered. The algorithm allows the user to set the required criteria such as Ground Sampling Distance (GSD) and overlap percentage for the 3D model. GSD is an important parameter which describes the theoretical accuracy by taking into account the pixel size. Based on these criteria, other photogrammetric parameters for the camera placement are determined from the traditional aerial photogrammetric block described by (Wolf & Dewitt, 2013) in equations(2) to (6). The baseline is the distance between two images that need to be captured to have a required overlap to produce a stereo pair. In multiple strips, lateral advance (W) is the area covered in side for a required sidelap. The number of photos can be determined by

the perimeter of the building and the baseline distance computed. The number of strips to cover lateral side is computed from the height of the building and lateral advance (W). This will ensure the sufficient coverage required to produce the 3D model.

$$B = Scale * (1 - endlap) * Fw \quad (2)$$

$$W = Scale * (1 - sidelap) * Fh \quad (3)$$

$$Photos = (Length / B) \quad (4)$$

$$Strips = ((h2 - h1) / W) \quad (5)$$

$$GSD = Pixel * D / f \quad (6)$$

Where,
B = The imaging base
W = The lateral advance (along the Z axis in this case)
Photos = The number of photos in a strip
Strips = The number of strips in the block
Length = Perimeter of the building in this case
h2-h1 = The height of the building
Scale = Image scale number
Fw = Frame width of camera
Fh = Frame height of camera
endlap = Percentage overlap between consecutive images in a strip
sidelap = Percentage overlap between the strips
f = focal length

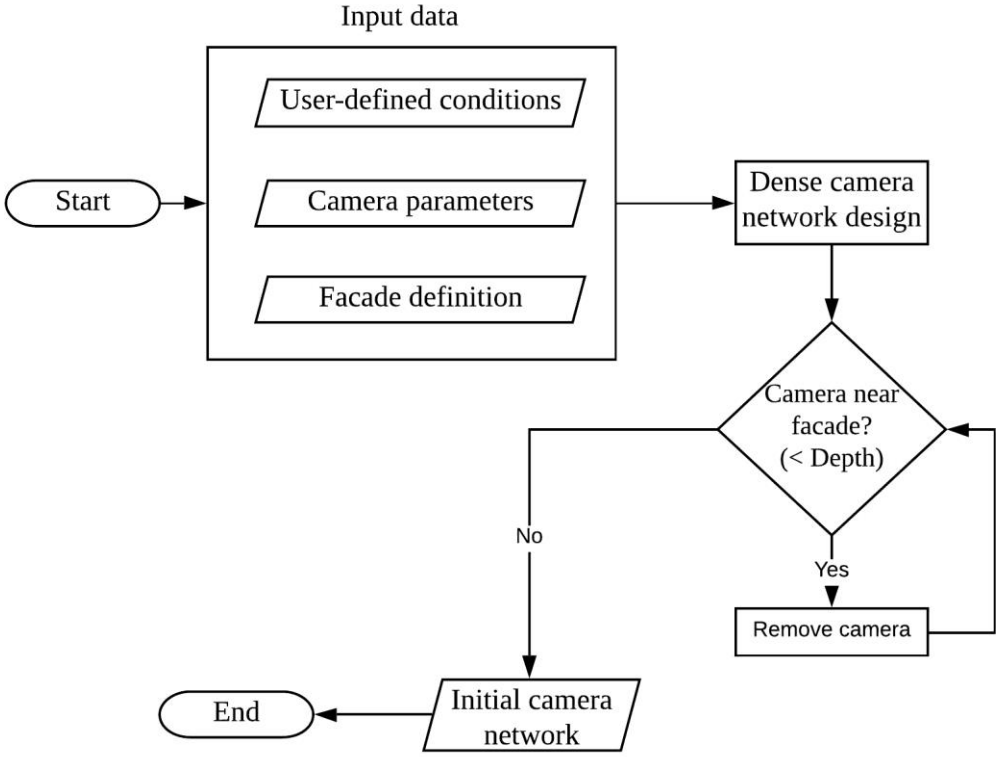


Figure 4.12 Steps involved in initial dense camera network planning

Pixel = pixel size of sensor
D = Depth (Distance from the object to the camera)

From the user-defined criteria and the sensor parameters, the theoretical accuracy of the camera network can be estimated using the equation (7) and (8) from (Forstner, 1998)

The theoretical horizontal accuracy for the image pair is given by

$$\sigma_H = \frac{H^2 \cdot P_x \cdot S_{px}}{B \cdot f^2} \quad (7)$$

The theoretical accuracy of point measurements in Z direction is

$$\sigma_Z = \frac{H^2 \cdot S_{px}}{B \cdot f} \quad (8)$$

Where,
σ_Z : the standard deviation in height
σ_H : the standard deviation in XY direction
H : the distance from the object to the sensor
Spx : is the expected collimation accuracy
Px : is the maximum parallax between homologous points in the stereo-pair.

The exterior orientation parameters for camera network were defined as follows. For the quadcopter, Omega (Ω) value ranges from 0º (nadir to the ground) to 90º (forward-looking). When Ω = 90º, the camera is directly looking at the horizon. Some drones with gimbal allow the camera to tilt upwards up to 30º additionally (i.e. 30º above the horizon). In camera planning, a standard 0º-90º range was adopted. In most cases, the Ω value is 90º, to acquire nadir images of the façade for uniform resolution. However, to ensure the coverage of the foot of the building, the first (lowest) strip of the camera is tilted down to 80º not to miss the lowest point of the building as the height of the building is a rough estimate. Phi (Φ) is defined based on the orientation of the facades to look in that direction. Kappa (K) for a quadcopter that uses gimbal is 0. The omega and phi angles for UAV are represented in figure 4.13

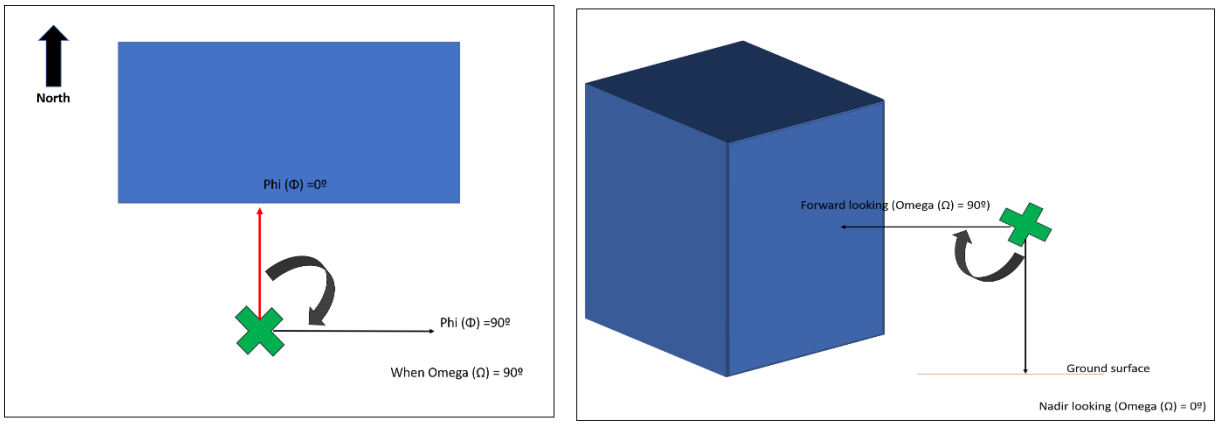


Figure 4.13 The orientation of UAV Phi (left) and Omega (right)

The flying height Z for each strip was calculated from the lateral advance W and the frame height Fh as follows

$$Z = h1 + ((Fh / 2) * Scale) + W * (Nstrip - 1) \quad (9)$$

Where,
$Nstrip$: the nth strip in the block

Considering the accuracy of GPS positioning in UAVs, a minimum depth of 10m from the object was used. Given the minimum depth from equation (9) the minimum flying height for the UAV was 4.6m. This was to ensure that the foot of the building is covered. In order to ensure complete coverage, different approaches were adopted for various conditions while traversing along the facades. These are discussed in following subsections 4.3.1 and 4.3.2.

4.3.1. Nadir cameras for building façade

All the cameras in the network are planned along the façade in a sequence in an anticlockwise direction, using the orientation of the facades. This help to reduce the time while flying the drone on the field. The ordered set of images reduces the computation time on the desk, during the bundle block adjustment as it finds the correspondences in successive images rather than checking the entire block for matches.

The cameras were placed in a line parallel to the surface of façade at intervals determined by the Base to Depth (B/D) ratio. Many research on stereo pair images had shown the importance of B/D ratio in reducing the error in the measurement of the model (Hasegawa et al., 2000; Hu, Gao, Li, & Li, 2016; Jing, Yong-she, Jian-ronga, & Jun-fenga, 2008). For better coverage and reliable accuracy B/D ratio for initial camera placement was considered from 0.2 – 0.4. Determining the viewpoints parallel to the façade surface was solved by right-angled triangle geometric problem from the length of the façade vertices and the Depth.

4.3.2. Converging cameras at corner

The nadir cameras of each façade were connected by converging cameras in the corners to avoid wide baselines. These cameras make an arc along the corner of the building, providing oblique images observing points on both the facades. For initial dense camera network, cameras are placed at every 10º transition. Depending on the steep angles made by the facades, more converging cameras were added to the corners. However, for the interior corners defined in section 4.2.2., the convergent transition was not possible as the facades make it difficult to fly close by. In such cases, the last and first cameras of the corresponding facades were tilted to look towards the interior corners. This is shown in figure 4.14(a) where 3 cameras of the corresponding facades are tilted to capture the inner corner. If that did not ensure better coverage, multiple images were taken at different directions capturing the corner from the same location. Figure 4.14 (b) shows the dense camera network with nadir cameras and convergent cameras for the building.

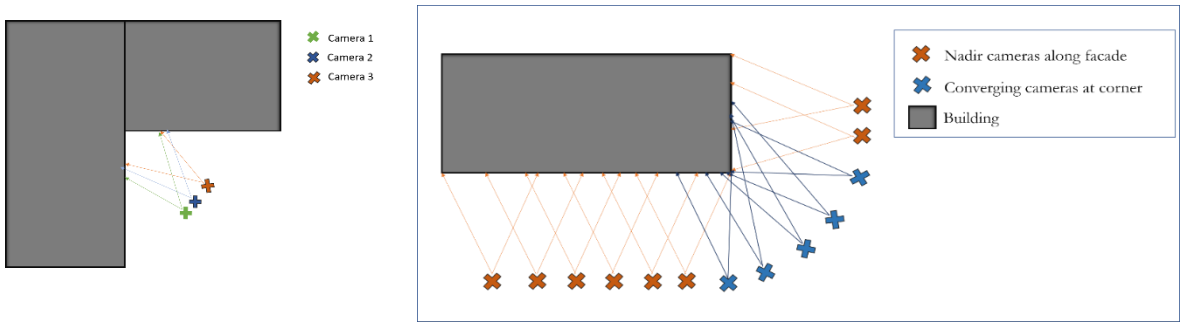


Figure 4.14 (a)Cameras tilted to look at inner corner (left) and (b) exterior corner covered with converging camera (right)

Some of the nadir cameras that are planned based on the façade segment length were close to the next façade surface along the inner corners of the building. These cameras were removed based on a threshold of distance close to the building to ensure a safe flight. However, removing these cameras did not affect the B/D ratio to a larger extent.

Further, the cameras were indexed with façade id and the values 0, 1, -1 to distinguish them as nadir looking, converging cameras at exterior corners and cameras at interior corners respectively. This was helpful during visibility testing process.

4.4. Visibility analysis

The visibility check was performed to confirm if all the points in building facades were covered in at least 3 images. However, points measured in more than 4 images yielded less improvement in terms of accuracy as the previous research shows (Remondino et al., 2006). This was done by ray intersection method. Each façade point was tested with the angle made by the ray from the camera centre to the point. From the Frame width and Frame height of the camera, Horizontal field of view (HFOV) and Vertical field of view (VFOV) were determined using the focal length. A matrix was created with camera id along the rows and façade point id along the columns. The cells were updated with value '1', if the point was visible in the particular camera. In this way, the number of images viewing a specific point was computed by adding along each column. This data was used for filtering the redundant camera in the next stage. A sample representation of the data storage in this visibility matrix is shown in figure 4.14

Fac. id / Cam. id	1	2	3	4	5	6	7	8	Points covered in each camera
Cam 1	1	1	1	0	0	0	0	0	3
Cam 2	1	1	1	1	0	0	0	0	4
Cam 3	1	1	1	1	1	0	0	0	5
Cam 4	0	0	1	1	1	1	1	1	6
Cam 5	0	0	0	1	1	1	1	1	5
Cam 6	0	0	0	0	1	1	1	1	4
No. of images visible	3	3	4	4	4	3	3	3	

Table 4.1 Visibility matrix between cameras and facade points

Each camera based on the façade to which it was mapped, starts checking the points in that particular façade. Sometimes the points in neighbouring facades might be visible if the facades make angles less than 180º. In

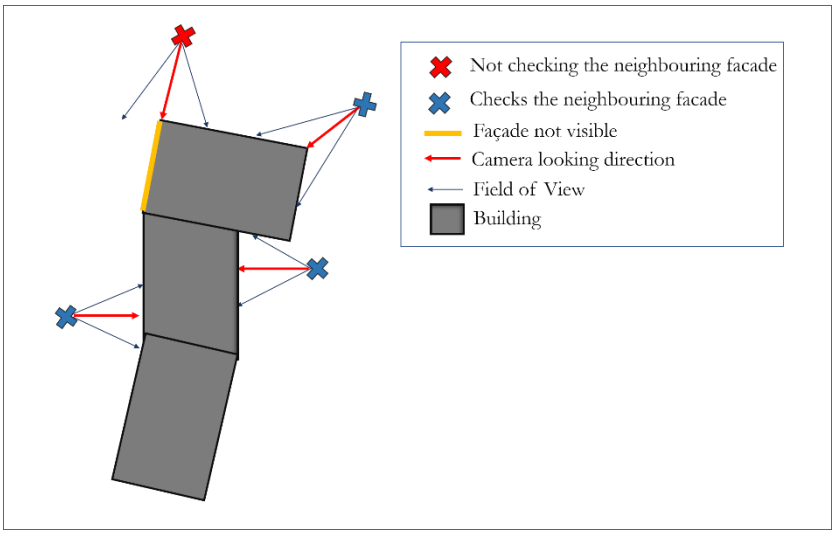


Figure 4.15 Visibility check for points of neighbouring facade

such cases, the angle between the camera looking direction and the orientation of the façade was determined. This angle should be greater than 60º and less than 120º. This helps to avoid taking into account the points that are viewed at steep oblique angles which can be obsolete. In this way, an object in the line of sight was detected as an obstacle blocking the view of other points in the neighbouring façade. In figure 4.16 the red colour camera does not check the points of neighbouring façade since the orientation of that neighbouring façade and orientation of the camera are in the same direction. Therefore, no points beyond the point of the facade to which camera was mapped will be observable. This method helped to check an actual number of cameras viewing a point in the entire façade surface of the building.

4.5. Filter based on coverage

The dense camera network planned has redundant cameras without improving accuracy or coverage. These images result in additional computation time in dense image matching. Also, it increases flight time for data acquisition. Hence, these cameras need to be identified and removed without compromising the completeness of the 3D model. Therefore, an ideal number of images sufficient to produce the same quality of 3D model as that of the dense network need to be acquired. This kind of minimal camera network

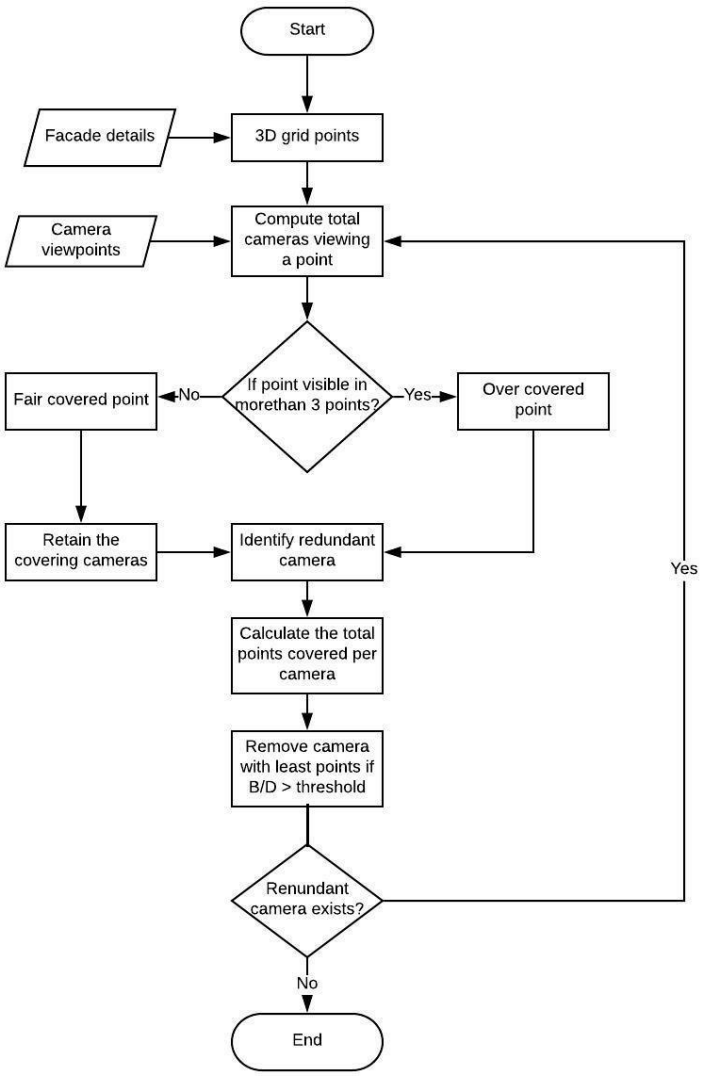


Figure 4.16 The workflow of filtering cameras for coverage based on the research by (Alsadik, 2015)

problem was already studied by (Alsadik, 2015). In that research, the author proposes two methods of filtering the redundant cameras: Filtering for coverage, filtering for accuracy. In this study, the filtering of cameras based on coverage was adapted to the UAV case. A detailed overview of the steps involved in the filtering is shown in figure 4.16

In section 4.4. it was described as how the number of cameras viewing a point was determined. From table 4.2, the number of images in which each point is visible is calculated along the last row of the matrix. The points visible in more than 3 images were identified as overcovered. The corresponding redundant cameras looking at the same point were back-traced from the matrix. Based on the number of points covered by each camera, the redundant cameras are sorted in ascending order. The camera covering least number of points was identified for removing from the camera network. The algorithm checks whether all the points were still visible in at least 3 images if that particular camera was removed from the network. It also checks the effect on B/D ratio before deleting. In case, the removal of the camera creates a huge baseline beyond the threshold specified, the camera was retained and marked as significant. For example, 3 cameras with a baseline(B) of 3m apart from each other are covering an object from a distance D = 15m. If the middle camera was removed, the new baseline between cameras would be 6m. The corresponding B/D ratio would increase from earlier 0.2 to 0.4. Say if the B/D threshold is 0.5, which is greater than the new value (i.e. 0.4) then the camera in middle would be removed. If the new B/D ratio is beyond the threshold, the camera would not be removed. The filtering was repeated until no cameras can be further removed.

5. RESULTS

In this chapter, the results of each step involved in the study are detailed. The subsections of this chapter are divided based on Chapter 4. Methodology into:

5.1 Building extraction
5.2 Façade definition
5.3 Dense camera network
5.4 Visibility check
5.5 Coverage based camera filtering

All the key process described in the previous chapter were carried out step by step for different building forms. Two real buildings from the DSM and a simulated building with complex shape are discussed in this chapter and the tests performed in other shapes are shown in appendix 1.

5.1. Building extraction

In this subsection, the intermediate and final results of building outline extraction using the DSM and orthoimage are shown. To study the difference between, the presence of nearby trees occluding the building and posing difficulty to recover the true shape of building, two different buildings were chosen. Building 1 (left) with no trees nearby and the Building 2 (right) with trees in its neighbourhood are shown in figure 5.1.

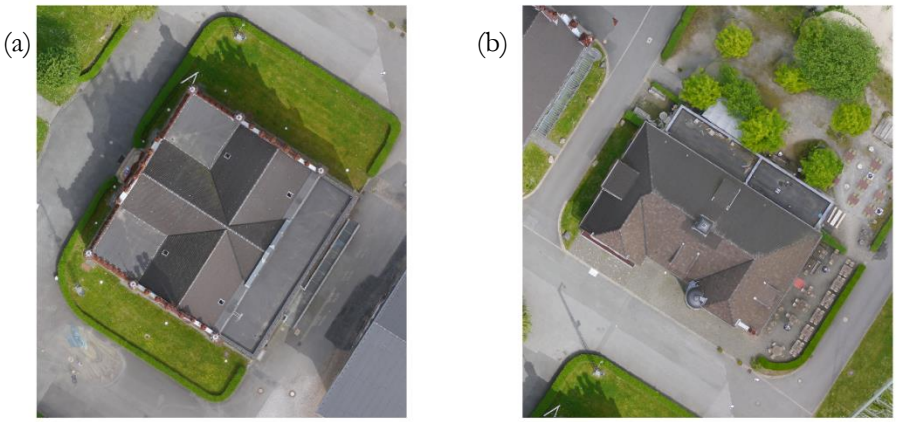


Figure 5.1 (a) Building 1 with no trees nearby (b) Building 2 with trees in neighbourhood

The seed point on these buildings were selected for the region growing segmentation. The results of the segmentation can be seen in Figure 5.2 (a) where building 1 with no trees in the vicinity was retrieved almost close to its original shape. Whereas, building 2 was segmented along with the trees as they overlap some portions of the building as seen in Figure 5.2 (b)

From the orthoimage, GRVI map was prepared by the equation (1) described in chapter 4.1.2. The GRVI image obtained at the end of this process was binarized based on threshold > 0, yielded the following binary map of trees in the neighbourhood. In Figure 5.3 it is seen that all the green vegetation including grass are identified as vegetation clearly. However, in figure 5.3 (a) since only low-level grass are present, it does not overlap with the boundary of the building. But in figure 5.3 (b) it is visibly identifiable that trees overshadow the building boundary.

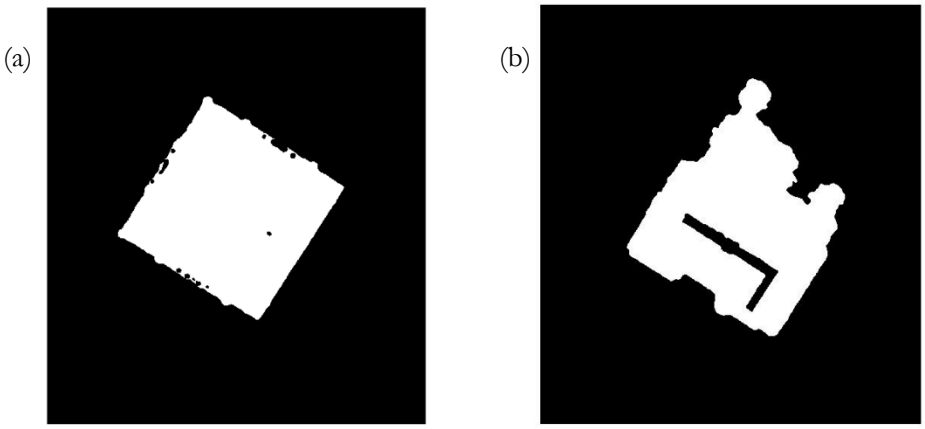


Figure 5.2(a) Building segment without trees in neighbourhood (b) Building segment including the trees in neighbourhood

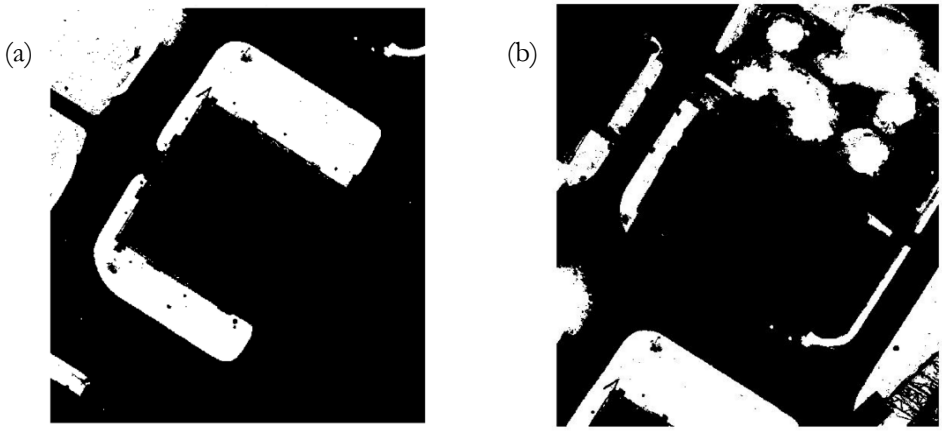


Figure 5.3 (a) Low lying grass in building 1 neighbourhood (b) tall trees overshadowing building 2

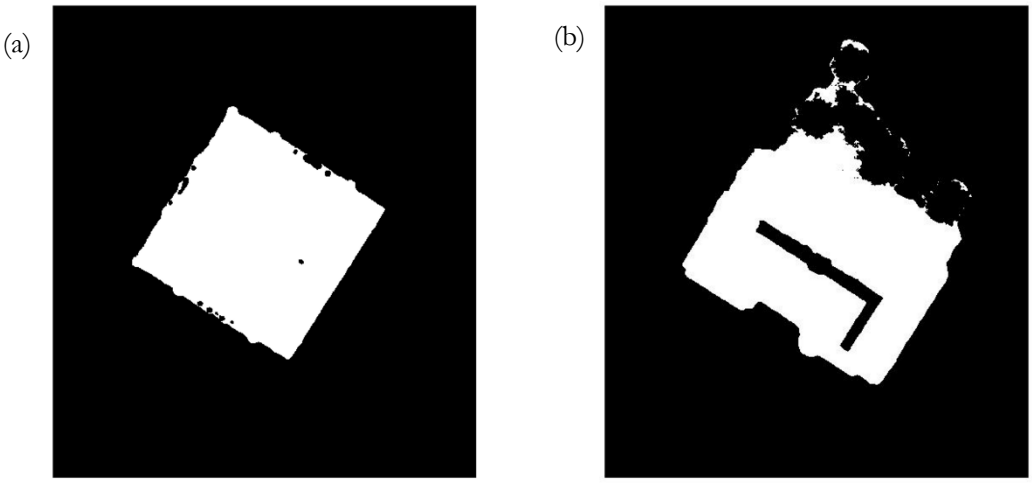


Figure 5.4 (a) Building 1 segment without changes after removing the vegetation (b) Building 2 segment with boundary partially recovered by removing the trees using GRVI map

The difference between the building segment and the GRVI image resulted in partially recovering the boundary of building 2 as can be seen from Figure 5.4 (b), But there was no considerable change in the building 1 segment figure 5.4 (a) as there were no trees near this building.

These building segments were traced by Moore's neighbour tracing algorithm, which resulted in obtaining the boundary pixels of the building segment as can be seen in figure 5.5. Due to the presence of trees, building 2 segment traced do not represent the true shape of the building figure 5.5 (b). Likewise, in building 1 due to the presence of tall tower-like structure on the edges did not result in proper segmentation due to large variations in the values. This resulted in cavities which were not repairable by morphological closing.

(a) (b)

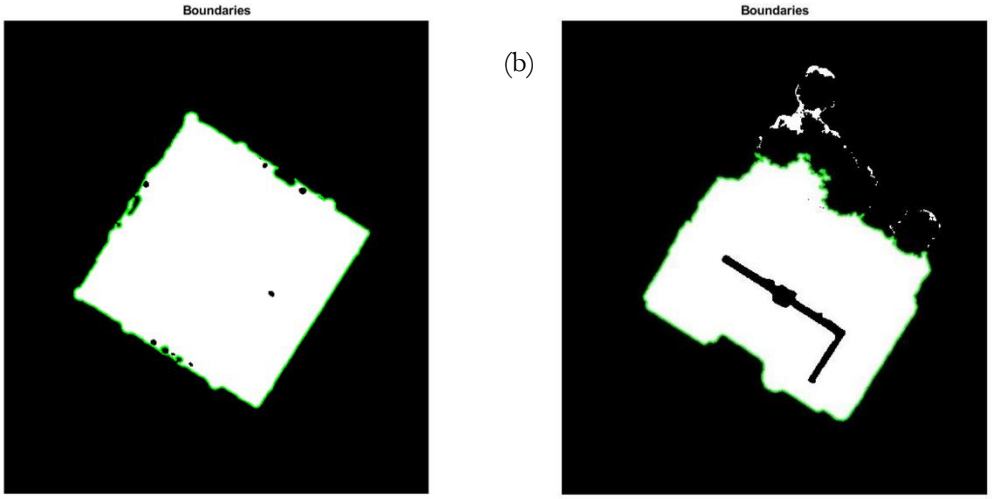


Figure 5.5 (a) Boundary of building segment 1 (b) Boundary of building segment 2

When the traced boundary (x, y) coordinates were used to create a polygon using predefined function polyshape in Matlab, some of the unwanted vertices which do not add any definitive value to the shape were

(a) (b)

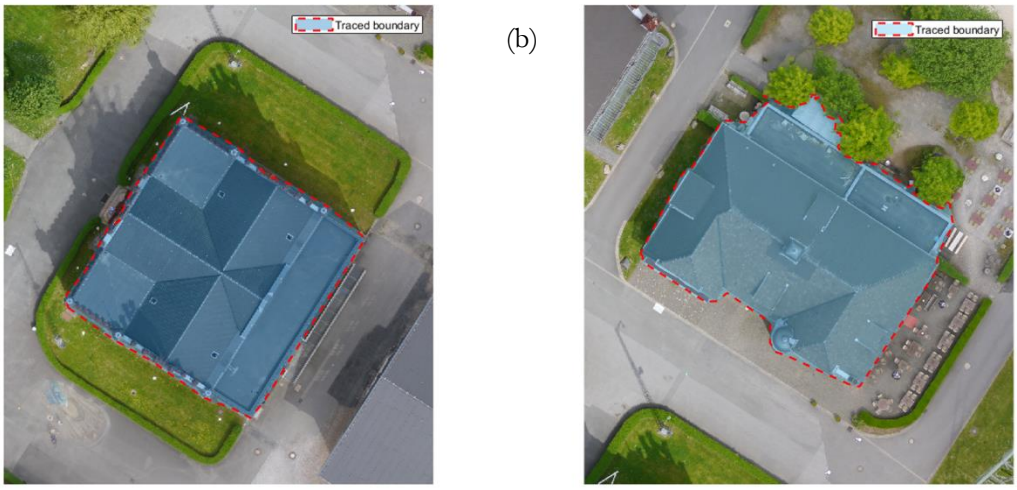
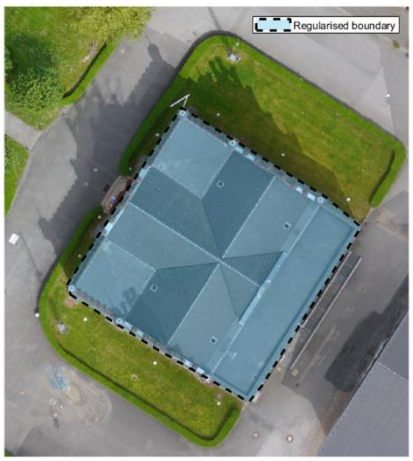


Figure 5.6 (a) Building 1 overlaid with the traced boundary (b) Building 2 overlaid with traced boundary

automatically removed. This helped reduce the vertices of the traced boundary. However, non-linear irregular boundaries due to trees in building 2 and the cavities in building 1 need to be rectified to get a good approximation of the building shape. The overlay of this polygon on the building is shown in figure 5.6

To regularise the boundary, vertices that were less than 2m apart were removed in building 1 and the vertices that were less than 1.5m apart were filtered in building 2. This filtering gave a better approximation of the building shape. However, any boundary segment less than the threshold distance was to be compromised as it is evident in building 2 figure 5.7 (b) where a small segment in the façade running inwards is generalised along with its neighbouring façade. Also, in building2 it was observed that a small portion of the building near the trees, was not able to be recovered as part of façade. When this region was closely examined, it looked like a temporary shed which may actually be not part of the main building itself. However, this needs to be considered for further research as how to retrieve the true boundary outline without losing the sections of the building when overlapping trees were removed from the segment.

(a)



(b)

Figure 5.7 (a) Regularised boundary of building 1 overlaid (b) Regularised boundary of building 2 overlaid

5.2. Façade definition

From the regularised boundary vertices, the orientation of the facades with respect to North was computed. Based on the threshold angles between the facades, the corners of the building were identified as exterior and interior. Building 1 did not have any interior corners. Building 2 had 3 interior corners of which 2 were identified. However, the third interior corner was not identified because it was less than 1.5 m from the main building façade itself. The height of the building is computed by finding the difference between the maximum value in the building segment and the values of the neighbouring pixels to the boundary which is assumed to be the ground surface. Therefore, the entire façade surface was reconstructed assuming the walls to be flat and vertical. This gave the rough surface of the building facades. Points were generated on the

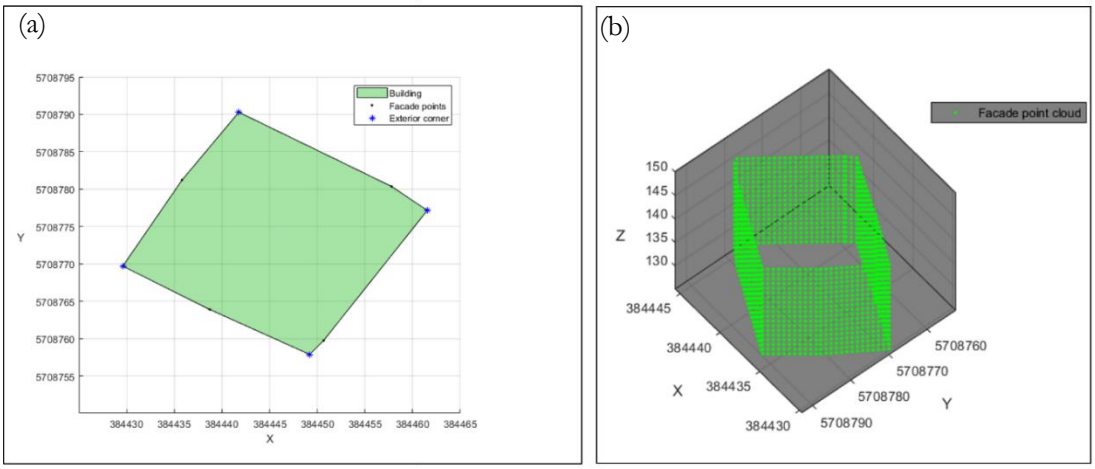


Figure 5.8 (a) Building1 boundary with exterior corner (b) Facade point cloud of the building1

surface at equal intervals of 1m for faster computations. However, the point density can be increased to achieve the required level of coverage. Figure 5.8 (a) shows the Building1 boundary with exterior corners detected from the boundary vertices along with other boundary vertices and (b) shows the facade point cloud of building1 generated using the boundary vertices and height of the building. Figure 5.9 (a) shows the Building2 boundary with exterior and interior corners detected from the boundary vertices along with other boundary vertices and (b) shows the facade point cloud of building2 generated using the boundary vertices and height of the building.

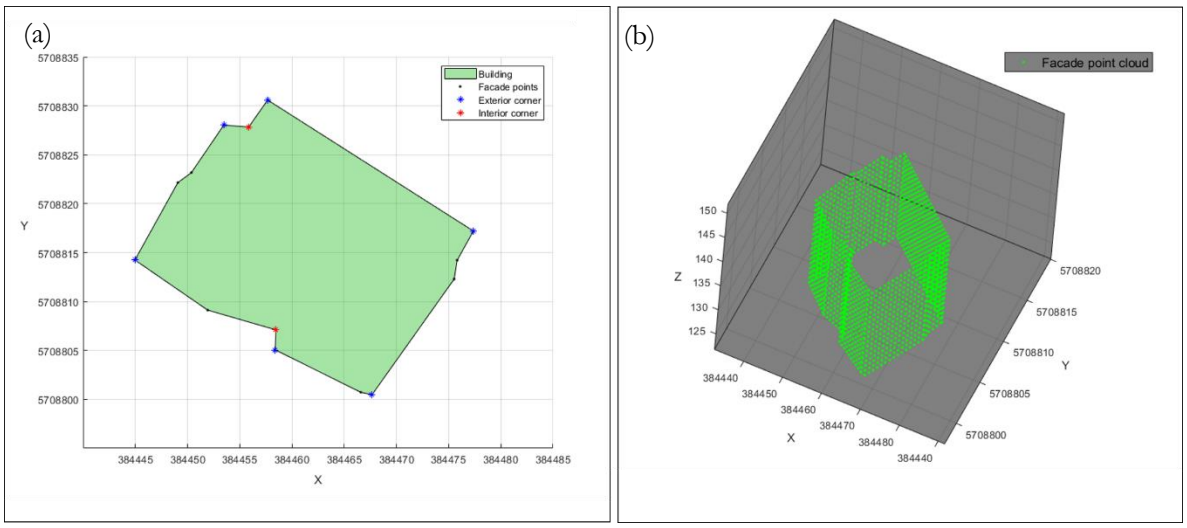


Figure 5.9 (a) Building2 boundary with exterior and interior corners (b) Facade point cloud of the building2

5.3. Dense camera network

For camera view planning, the parameters of Panasonic DMC- GX1 model was used as shown in Table 1. These interior orientation parameters can be changed according to the cameras used. The initial end-lap and sidelap were chosen as 80% and 40% with a GSD of 2cm for the dense camera network. This determines how close the UAV must fly near the object to achieve this resolution and adds a constraint to limit the maximum distance away from the object it can fly.

Sensor Parameter	Value
Focal length	14 mm
Frame width x height	17.3 x13 mm
Image format	4592 x 3448
Pixel	3.75 μm

Table 5.1 Interior orientation parameters of the camera

The number of cameras planned depended on the surface area of the façade to be covered, i.e., perimeter * height of the building. The two buildings mentioned earlier were used to plan the dense camera network. The perimeter and height of these buildings were almost similar although the shape varied considerably. Building 2 had 2 interior corners and many intermediate points along the facades. Therefore, to understand how the algorithm works for different shapes having almost similar properties in terms of area and volume, these 2 buildings were useful. The Perimeter, height and the number of façade points generated for the two buildings are reported in table 5.2

Properties	Building 1	Building 2	Building 3
Total Façade points generated	2134	2178	7018
Perimeter	94.34m	93.09m	310.73
Height	20.42m	20m	20m

Table 5.2 Properties of the buildings for which dense camera network was planned

The depth at which images are to be captured was chosen as 20m away from the facades for building1 and 2. The depth of capturing images for simulated building3 was 40m. This was calculated from equation (6) based on GSD requirements. In this case, the planned GSD for building1 &2 would be 0.5cm and for building3 would be 1cm to ensure a sub-centimetre resolution. These distances are uniform throughout the image capturing to get a homogenous resolution of building facades. Based on this, GSD, scale, baseline distances were computed. From equation (6) and (7), assuming the expected collimation accuracy to be 1 pixel and maximum parallax as half of the frame width, the theoretical accuracy for the camera networks of both the building were calculated. From equation (2), for the dense camera network with 80% forward overlap at a defined predefined scale given by the ratio between focal length and depth, the baseline distances were computed to be 3.69m, 4.11m and 7.15m for building1, building2 and building3 respectively. The horizontal and vertical theoretical accuracy computed from the parameters discussed is shown in table 5.3. It was observed that the estimated accuracy for the computed dense camera network is within the acceptable range of few centimetres. However, it has to be noted cautiously; these values may vary in real field conditions.

Theoretical accuracy	Building 1	Building 2	Building 3
Horizontal accuracy, σ_H	0.026m	0.029m	0.04m
Vertical accuracy, σ_Z	0.018m	0.016m	0.06m

Table 5.3 Theoretical accuracy of building 1 and building 2

The total cameras in the dense camera network configuration for building1 is 118, for building2 is 139 and for building3 is 85. Even though the perimeter and height of buildings were similar for building 1 and 2, the variation in the number of cameras can be attributed to the corners of the building. For building2, in total 6 exterior and 2 interior corners were detected from the building outline. Whereas, for building1, the shape is more like a square with 4 exterior corners. These differences caused the increase in a number of cameras for building2. The initial dense camera network of building 1 and building 2 are shown in figure 5.10 where the façade point cloud of building and the cameras looking at these facades are shown. Even by visual interpretation, it was obvious that these imaging viewpoints can generate a good dataset for modelling the building.

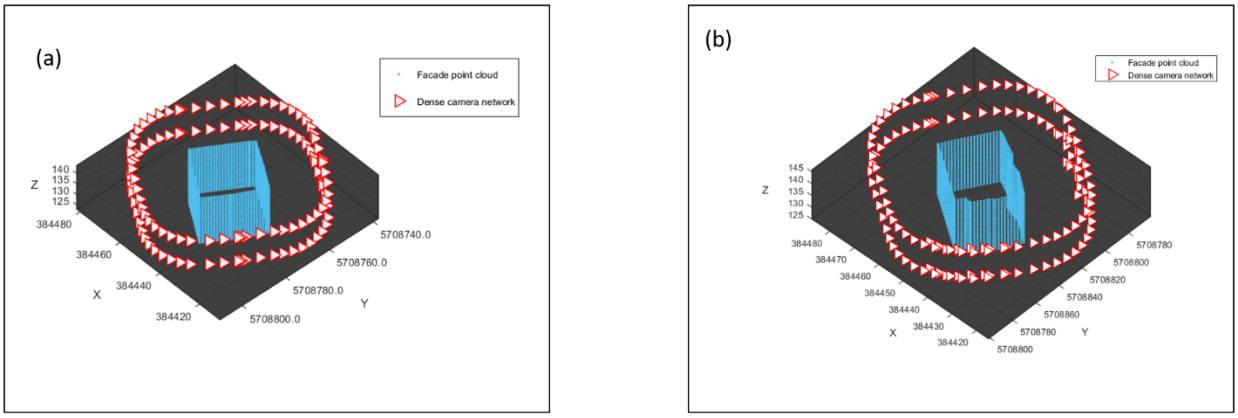


Figure 5.10 Dense camera network for (a) building1 (b) building2

The initial dense camera network for building 3 is shown in figure 5.11. The number of cameras for data capture was lower for this building, even though the perimeter of the building is larger as much as 3 times that of building 1 and 2. This was due to the chosen depth at which images were to be acquired. It is worth noting that the accuracy of the building3 is lower than the other 2 buildings for the same reason.

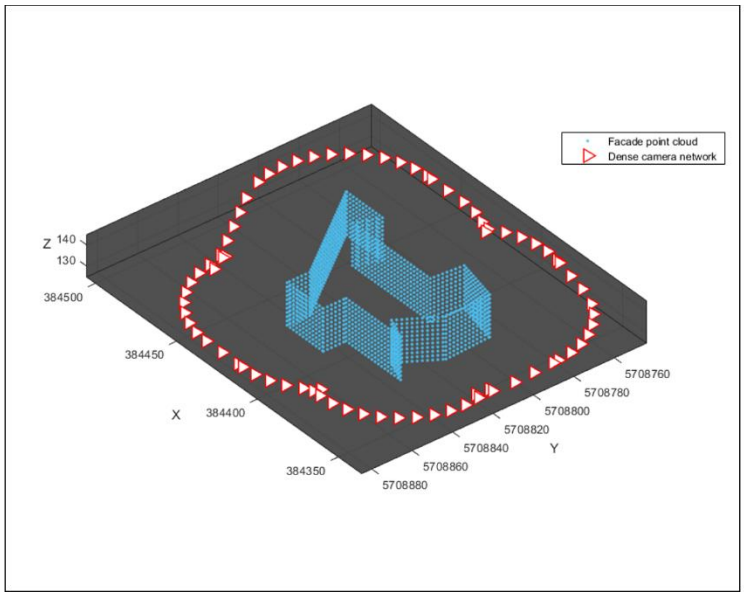


Figure 5.11 Dense camera network for simulated building3

5.4. Visibility analysis

The visibility of each point of the building façade in the dense camera network was verified using the method described in section 4.4. If the tangents made from the façade point to camera were within Horizontal Field of view and Vertical field of view, the point was marked as visible from that camera. This type of visibility check was quick for a sparse point cloud of approximately 10,000 points. The algorithm was able to compute the results for visibility analysis in less than a minute in this case. Figure 5.12 shows the visibility of each point in a number of cameras for building 1 (left) and building2(right). The dense camera network has a large number of cameras to capture the corners. Since the façade points are covered in these oblique images apart from the nadir images that were planned for each section of façade, visibility is as high as 30 images for some points in building1 and 44 images for some points in building2.

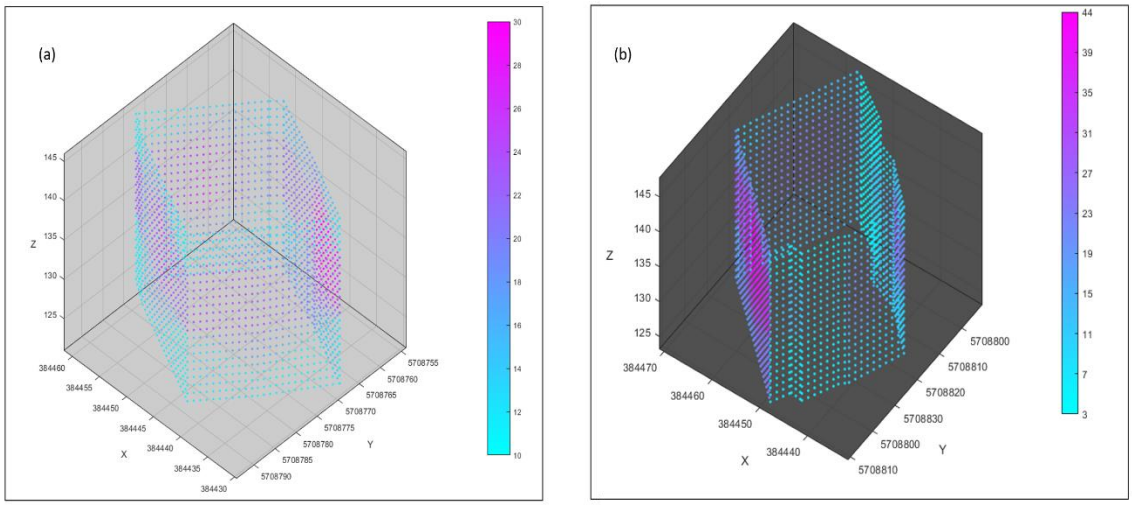


Figure 5.12 (a) Number of images in which each facade point is observed for building 1 (b) for building 2

The visibility analysis for simulated building3 showed a maximum of 22 images at a few points which can be seen in figure 5.13. The results of the visibility analysis were used to detect over-covered facade points, and the corresponding cameras were removed in the next phase.

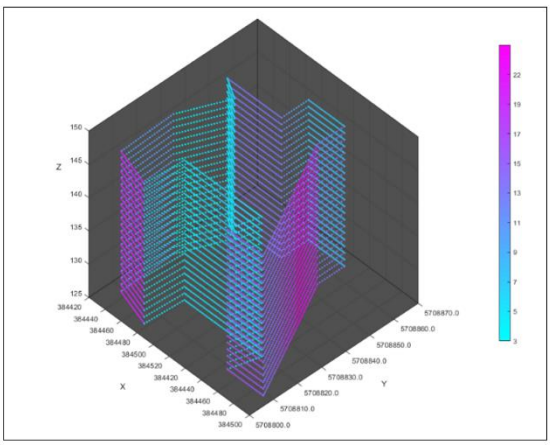


Figure 5.13 Number of images in which each facade point is observed for simulated building3

5.5. Coverage based Camera Filtering

From the visibility check, the algorithm was able to ascertain the parts of the building which were over-covered by cameras. Then, the dense camera network was filtered based on the approach discussed in section 4.5 Filter based on coverage. The façade points that were covered in more than 3 cameras were identified and their corresponding cameras were listed as redundant cameras. From this list, the camera that was covering least number of points was selected and checked for B/D threshold of 0.7 for nadir cameras and 0.6 for the cameras at the exterior corner. If the threshold is satisfied even after removing the camera, it was removed. In this way, redundant cameras are filtered iteratively until no further cameras can be removed. In most of the cases, the cameras were removed in one iteration successfully. For the demonstrative purpose, the filtering for building1, building2 and building3 are shown in this section. For building1, the initial dense camera network configured had 118 cameras. This was reduced to 46 cameras in the first iteration. However, only one more camera was removed during the second iteration. Further, no cameras could be removed from the network. For building2, the initial dense camera network consisted of 139 cameras. This was reduced to 96 cameras after filtering. For building3, the initial dense camera network consisted of 85 cameras. This was reduced to 61 cameras after filtering. On further iterations, no cameras were able to be removed. This can be seen from figure 5.14. which shows the initial camera network getting reduced after filtering based on coverage for building1 figure 5.14 (a) and for the building2 figure 5.14 (b).

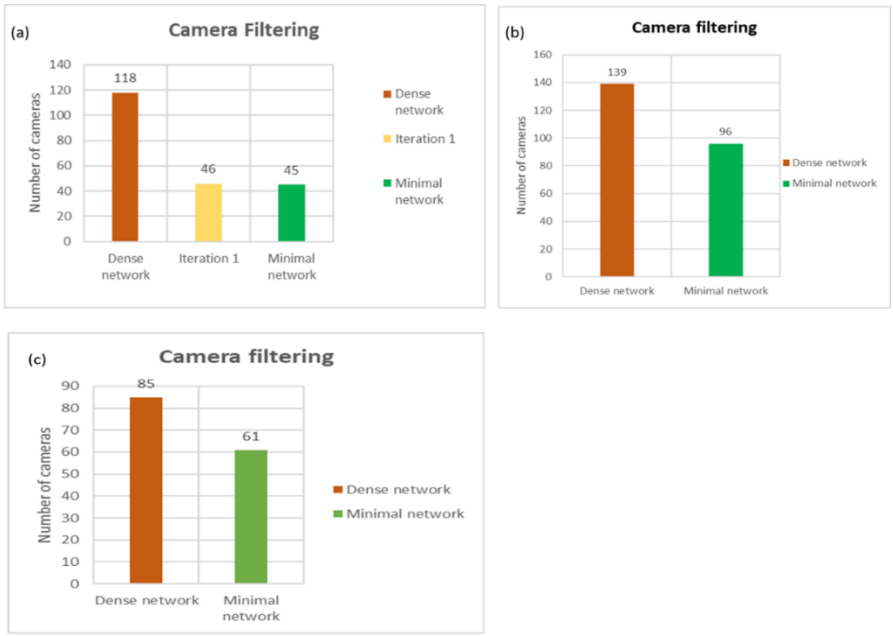


Figure 5.14 Coverage based camera filtering for (a) building1 (b) building2 (c) building3

The filtering of cameras for coverage reduced the number of points captured in total by a camera network. However, the average number of cameras observing a point is as high as 10 for building2 and 7 for building1 which are more than the required 4 images per point for 3D reconstruction. The table 5.11 shows the total number of façade points used to test the visibility in cameras and the total points viewed by the initial dense camera network and Minimal camera network. i.e. the total number of points captured by all the cameras together in a network. The last column of the table shows the average number of cameras looking at a facade point in the minimal camera network after filtering.

In the minimal camera network, the highest number of images viewing at a point was 14 for building1, 32 for building2 and 18 for building3. This is comparatively lesser than that of a dense camera network which was 30, 44 and 22 images viewing at some points for building1, 2 and 3 respectively. From the figure 5.15, it can be noted that the highest coverage of points was along the overlap of 2 strips in building 1&2, especially in the flat façade without any inner corners in all the three buildings. This was particularly because of the additional converging cameras that were placed on both the corners of a façade, looking at it in

oblique angles which allows to cover more points than the nadir cameras of the facades itself. As seen from the 3 building cases, the algorithm ensured the reduction of dense camera network by 62%, 31%, 28% for building 1, 2 and 3 respectively. This significantly saves a lot of processing time in the later stage.

Building id	**Total number of façade points**	**Total points viewed by Dense network**	**Total points viewed by Minimal network**	**Average visibility of a point after coverage filtering**
Building 1	2134	38229	15199	7.1
Building 2	2178	35224	22778	10.4
Building 2	7018	74844	53394	7.6

Table 5.4 Comparison of the visibility of points in Dense and minimal camera network for building1 and building2

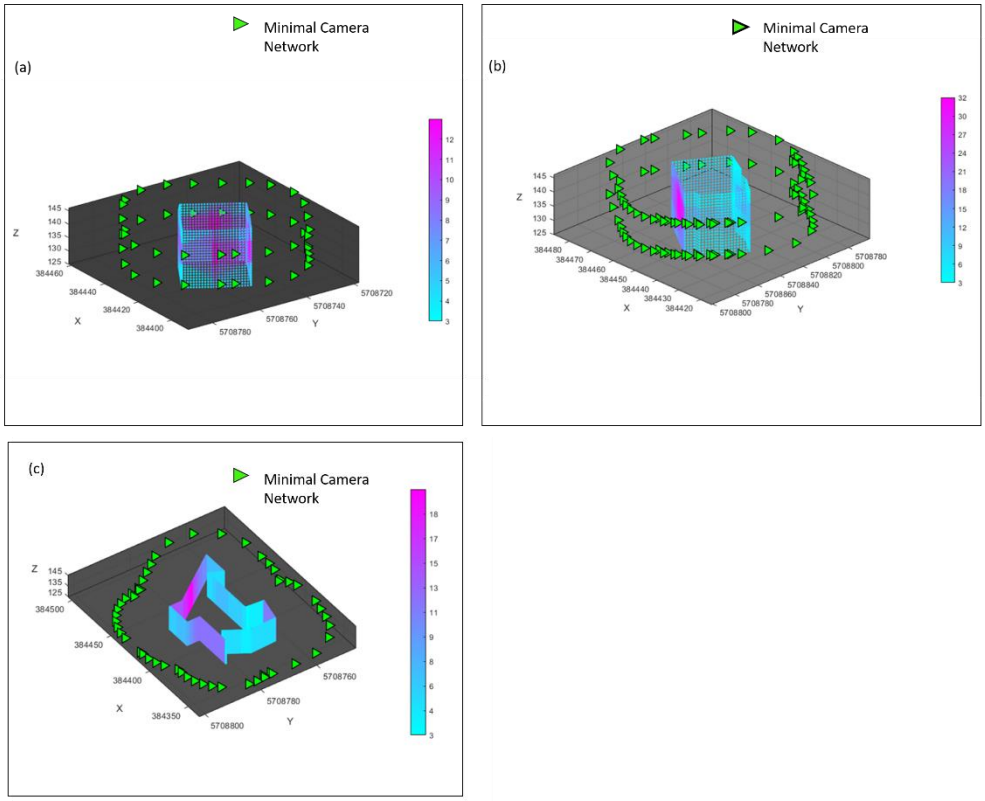


Figure 5.15 Minimal camera network for (a) building1 (b) building2 (c) building3

The visibility of each point in a total number of images was plotted in the graph as seen in figure 5.16 for initial dense camera network and minimal camera network after filtering for coverage. The spikes in the graph can be attributed to the flat facades of the building that received more coverage from the corner cameras in addition to the nadir cameras observing them already. The recurring pattern of a spike in the graph shows a section of the building is receiving higher coverage. Since the façade points are given in cyclic order of strips from bottom to the top of the building, this happened.

The second and third spikes in each pattern can be attributed to the points near the corners of the building. Since converging cameras were placed at shorter baselines, visibility of all the points near the exterior corners increases. The dips in each pattern in figure 5.16(b) and (c) is due to the interior corners that get captured in exactly 3 or 4 images since these points are occluded because of the main façade.

The minimal camera network after coverage-based filtering for building1 with 45 cameras, building2 with 96 cameras, building3 with 61 cameras along with the visibility of each point in a number of cameras are shown in figure 5.16(a), (b) and (c) respectively. The graph in figure 5.16 (a) for building1 substantiates the fact that coverage is more in the overlapping region of strips as can be seen from the gradual increase in the curve followed by a gradual decrease between the points 900 and 1600.

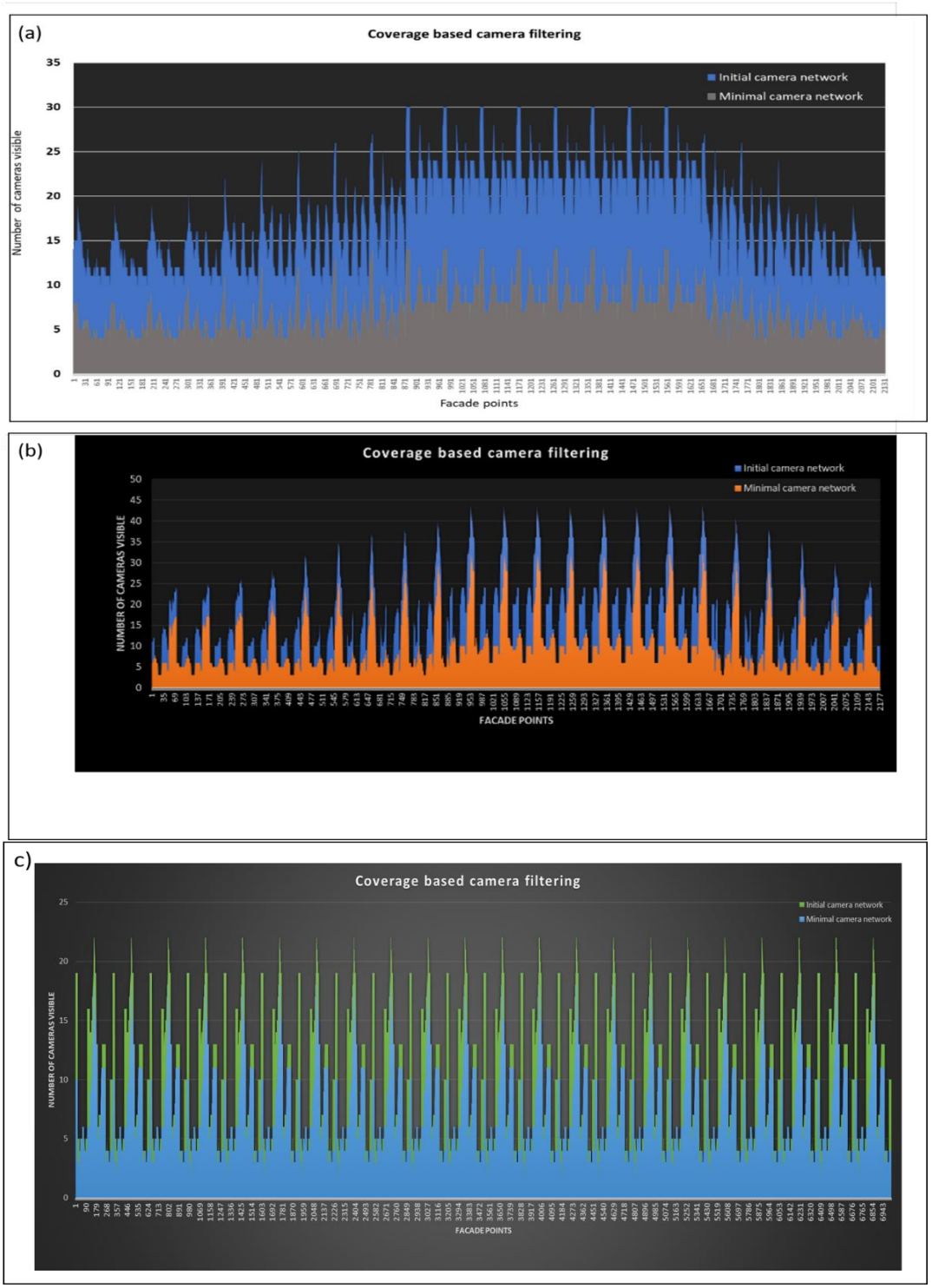


Figure 5.16 Comparison of observability of each point before and after coverage filtering for (a) building 1 (top) and (b) for building2 (middle) (c) for building3(bottom)

6. DISCUSSION

In this section, the advantages and limitations of the methods used and critical comments on the results are discussed. This section is grouped into five major subsections similar to chapter 4 and chapter 5, highlighting the overall key steps involved in this study:

6.1 Building extraction
6.2 Façade definition
6.3 Camera network design
6.4 Visibility analysis
6.5 Coverage filtering

6.1. Building extraction

The proposed flight planning method relies on delineating the building boundary from the rough DSM obtained from nadir flight. Therefore, accurate building segmentation was necessary to obtain the building outline. Mean based region growing method has its limitations in missing the building pixels in the segment when the value is beyond the set threshold. In this study, for the buildings observed in the DSM, the threshold reg_maxdist =5, maximum variations that can be included in the region gave an acceptable estimation of the building shape. However, for other data, this threshold may no longer hold and the user needs to estimate this value for segmentation. For example, if the variation in the building altitude at different regions is beyond this threshold, it was difficult to detect the entire building as a single segment. In building1 this was evident where the segmentation did not include tall towers on the outline which varied larger from the other parts of the building. In such cases, different segments of the building can be merged to get the overall footprint of the building.

Using GRVI for removing the tall trees, that hampered the precise boundary extraction worked well for building2 in this case. However, some trees in the other region of DSM were not identified because of the presence of flowers that had a different spectral signature. This misinterpretation can be easily incorporated

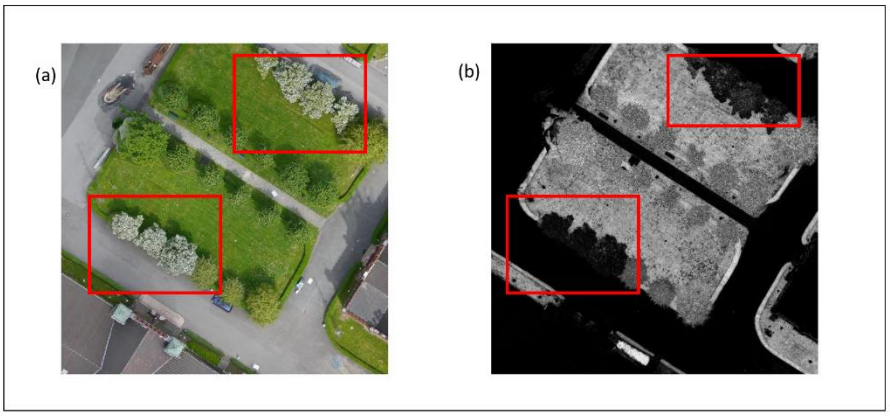


Figure 6.1 Misinterpretation of trees in GRVI method

into the building outline leading to irregular outline of the building. Figure 6.1 (a) shows the trees with different spectral value (other than green) in the orthoimage and (b) the corresponding GRVI with green vegetation represented in white and other features in black.

While tracing the boundary from building segment, that was cleared off from tree pixels one or two erroneous pixels were traced as a part boundary which was identified when plotting these coordinates as a polygon. These points were to be removed to get a good approximation of boundary. However, when there

were no trees overlapping the boundary, this problem did not pose a threat. While regularising the boundary, some intermediate points in the façade still existed and were considered an individual façade. If not removed, these points will lead to difficulties in visibility testing, as the test was implemented to check only the points in direct neighbouring facades. In most of the cases, boundary outlines were able to be retrieved with reasonable shapes. Overall, better segmentation of building can result in a good approximation of the building outline that can be used to plan the cameras.

In the case of building2, during regularisation of the outline, a small section of the building was compromised due to the presence of trees. This can be solved by adopting a better generalising strategy when the vertices of the boundary are too close to each other such as fitting a line with low residuals in the place of a sequence of nearby points due to trees. In this way, some structures close to the trees can be retrieved.

6.2. Façade definition

First of all, the building façade was assumed to be flat with no protrusion such as a balcony or hanging roof. In this way, points for the façade was generated based on the outline extracted from the DSM. This can lead to self-occlusion of some parts of the building based on the flying height of the UAV. Sometimes, minor inaccuracies due to variation in depth at these locations is possible. However, flying at a certain distance, these difference in depth will be negligible. The detection of corners of the building depends directly on the angles made between the meeting facades. In this study, for inner corners 30º to 160º and for outer corners, 190º - 350º were used which is true for most of the buildings. However, in an exceptional case, if the facades of buildings do not meet at corners in these angles, corner detection have to be altered for better detection. Since corner detection is necessary for planning a converging camera network that will minimise the wide baseline problems allowing a uniform depth. This step is critical to ensure the smoothness and connectivity of facades.

6.3. Camera network design

The camera network design was proposed to deliver a 3D model with an accuracy based on equation (7) and (8) for stereo pairs. This does not mean that the final model of the building will always have predicted theoretical accuracy. The accuracy of the 3D model depends on multiple factors in a real situation during the data collection such as calibration of the instruments, precise positioning, image matching on featureless surfaces, etc.

The overlap percentage used in the design can be increased if the building façade has minimal features which can act as tie points. This will ensure more homologous points in every image pair that will reduce error in matching. Varying the depth at which an object is observed can highly alter the quality of 3D model. This was seen from the computed theoretical accuracy of building3 flying at 40m away from the building compared to flying at a 20m distance in case of building1 & 2. Flying far from distance means, the scale at which the buildings are captured are less. Since the scale is the ratio of focal length to distance, this will adversely affect the GSD of the images.

The view planning is suited for obtaining a homogenous resolution for the facades assuming the nadir images can be obtained for the building that will provide similar resolution for the roof. Therefore, this method is useful for flat roofs. However, the camera network design with 80% front overlap and 40% side overlap ensured the coverage of building from foot to the top. For gable roofs, efficient planning is needed to get homogenous data on the roof similar to that of façade. By adopting a similar converging camera network in connecting the roofs to facades, problems in the connecting edges can be addressed. From the

DSM, the slope of the gable roofs can be calculated and cameras can be placed to face the normal of these planes

The convergent camera design in the exterior corners made sure to reduce the photo base distance between the façade cameras. This also acts as the oblique cameras covering a large number of the points of façade which can complement the nadir-looking cameras for 3D reconstruction later. Similarly, for the interior corners, adding additional images looking at different angles ensured better coverage (>=3 images). The effect of adding converging cameras in the view planning is shown in figure 6.2 where a wide baseline is reduced to short baselines by adding additional cameras at equal depth from the object.

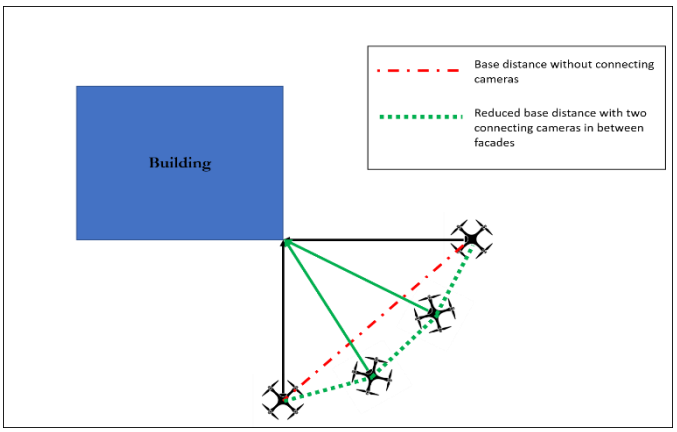


Figure 6.2 Convergent cameras reducing the distance between photo base for neighbouring facades

6.4. Visibility analysis

The mapping of cameras to facades helped to perform visibility analysis of points faster. For 10,000 points, the algorithm was able to deliver results in less than a minute. However, when tested for the curved facade, the algorithm did not perform well since the curved façade was defined with points closer to each other, each of which was treated as individual façade by the program which was designed to check only the direct neighbouring facade. Even though sufficient camera coverage was there, visibility check could not be performed from all the cameras as each camera is mapped only to a small segment of the curved façade.

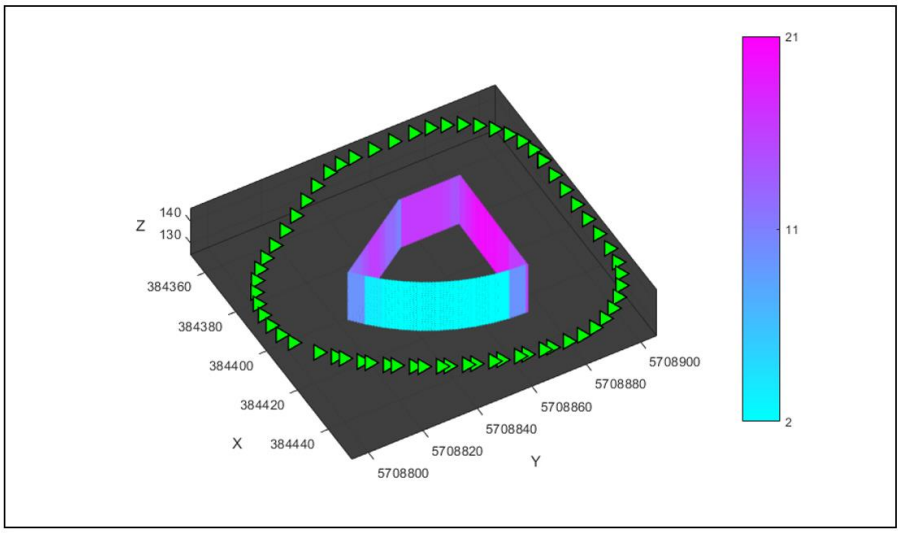


Figure 6.3 Visibility check not performing well for the curved facade

The limitation on the curved surface is shown in figure 6.3 where the visibility check shows that curved façade is not covered by 3 images whereas it is actually covered in more than 3 images.

In this visibility testing, the presence of obstacles such as trees in the vicinity was not considered. This could be one of the reasons for building2 to have higher visibility in the façade points which is partially occluded by nearby trees in reality. This problem can be addressed either by planning cameras around the trees trying to capture the occluded facades in oblique images. In DSM, only the crown of the tree is seen, However, if the tree is taller than the building, flying below the tree height may ensure coverage of façade apart from the occlusion caused by the stem of the tree. But this can also result in the shadows in images due to the tree.

Since the position of trees are known from the GRVI output, this can be used to check if a tree is affecting the line of sight from the camera to the façade point, by similar box ray intersection method. This can help to improve the reliability of visibility testing. However, the actual occlusion because of the trees cannot be avoided sometimes.

6.5. Coverage filtering

The coverage-based camera filtering provided better results for regular polygons with only exterior corners as seen in building1. When there were many interior corners in the building2, it was difficult to reduce the number of cameras as the number of cameras viewing at inner corners are rather limited and have a significant impact if removed. The cameras placed at inner corners have very less coverage in terms of the number of points covered by them, as it is aimed only at ensuring a better coverage in those interior corners while the scope of visibility is limited. Similarly, for a complex shape such as building3, the reduction percentage of the camera was low. However, the reduction was acceptable as the primary objective is the completeness of data acquired. Filtering of redundant images proved to reduce the processing time up to 60% as demonstrated by (Alsadik, Gerke, Vosselman, Daham, & Jasim, 2014).

The developed algorithm was able to reduce the number of cameras by up to 60% in less time. This will help to collect quality data with less effort in terms of data collection and processing.

7. CONCLUSION AND RECOMMENDATIONS

7.1. Conclusion

In this thesis, the algorithm was developed to compute the positions and orientation of cameras for UAV that are required to capture a building for the 3D model. For this, the DSM obtained from nadir images of the region was used as input. From this input, building footprint was extracted using region growing segmentation method. To remove the trees that were overlapping the boundary of the building segment, GRVI of the region was computed. The morphological operations were performed to remove the obstacles and close the openings in the building the segment. This segment was traced using Moore's neighbour algorithm to extract the outline of the building. From this boundary outline and the height of the building from DSM, assuming the walls to be vertical, façade point cloud was generated for the building. The orientation and corners of the building facades were determined from the boundary. The scheme of maximum to minimum in camera view planning was adapted. The initial dense camera network was produced with elementary photogrammetric block equations (Wolf & Dewitt, 2013) and the theoretical accuracy given by (Forstner, 1998). From this approach, the 3D model to be constructed was able to give a theoretical accuracy within the acceptable photogrammetric standards in order of few centimetres. Special attention to capture the corners efficiently was achieved by converging camera network. Visibility check of the façade points on the dense camera network was performed by an approach similar to the box-ray intersection. From the visibility of points in 'n' number of images, redundant cameras were identified and removed based on coverage based filtering for minimal camera network proposed by (Alsadik, 2015). The camera view planning, visibility analysis and the coverage-based filtering method was time efficient and could produce results in less than a minute depending on point cloud density. The reduction of the required number of images up to 60% in some buildings was an important factor to note. Thus, a minimal camera network for 3D modelling of a building with an average visibility of more than 3 images and estimated accuracy was produced using the algorithm. The tests were performed on two real buildings from the DSM and other simulated buildings to check the efficiency of the algorithm. The performance of the algorithm was good when building footprint was well defined in terms of regular geometric shapes. However, in the conditions of small sections of facades running inwards from the main façade, curved façade, obstacles in boundary it needs to be improved. The camera planning and visibility analysis can be still improved by considering the obstacles in the scene which causes occlusion of the building as discussed in section 6.4. Camera network design can be extended further in modelling the roofs with homogenous resolution as mentioned in section 6.2. The final output of the algorithm, the 6 exterior orientation parameters of the UAV to capture the data for 3D building model was produced. Thus, trying to automate the process of data capture for UAV, with less intervention from the pilot. However, this method needs to be tested on the field to check the efficiency of the algorithm in producing building models.

7.2. Recommendations

The study considered camera view planning only for the facades assuming the roofs can be captured by regular grid-based nadir images using UAV. However, for complex roof structures, to have a uniform resolution similar to that of façade surface, a detailed study needs to be undertaken by an integrated approach to include different view plans for each section of the building such as façade and roof. This can be planned by considering the orientation of the roof plane in 3D space and flying the drone at the same depth as that of capturing the façade.

Collision avoidance and better positioning of cameras in case of obstacles in the scene can be planned from the DSM data. The computed camera view points can be used to verify if there is any obstacle in a search space of few meters from its position and if there is any, it can alternatively select a viewpoint in the neighbourhood while avoiding a collision. But this step can be computationally intensive to perform for all the camera locations. Therefore, selecting the facades with obstacles, as demonstrated by building2 in this study and performing this check can yield better camera viewpoints.

The number of images in the process has considerably reduced after coverage filtering. The exterior orientation parameters obtained at the end of this process can be further optimised with a mathematical objective function minimizing the computed average error of object points (on the facades) and considering an optimal covariance matrix for the object point coordinates (Alsadik, Gerke, & Vosselman, 2013). In this process, some constraints are introduced for each parameter so that the points previously visible in cameras continue to be visible in the same cameras after the change in orientation at the end of the optimisation. At the end of this process, the theoretical accuracy of the façade reference points (object coordinates) improves considerably. This process is computationally intense; hence it was suggested by the author to use only when error after filtering of cameras exceeds the expectation of the project. The optimisation problem is further being investigated as follow up to this research.

The algorithm needs to be tested in the field to test the effectiveness and difference in accuracy of the planned 3D model and actual 3D model reconstructed from real images captured.

Once this algorithm proves to give considerable results in real cases, it can be developed as a GUI application for UAV, letting non-professionals use for data collection.

LIST OF REFERENCES

Acute3D. (2019). Retrieved March 5, 2019, from http://www.acute3d.com/software/

Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S. M., & Szeliski, R. (2011). Building Rome in a Day. *Commun. ACM*, *54*(10), 105–112. https://doi.org/10.1145/2001269.2001293

Agisoft. (2019). Retrieved March 5, 2019, from https://www.agisoft.com/

Ahmadabadian, A. H., Robson, S., Boehm, J., Shortis, M., Wenzel, K., & Fritsch, D. (2013). A comparison of dense matching algorithms for scaled surface reconstruction using stereo camera rigs. *ISPRS Journal of Photogrammetry and Remote Sensing*, *78*, 157–167. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2013.01.015

Alsadik, B. (2015). *Guided close range photogrammetry for 3D modeling of cultural heritage sites*. https://doi.org/10.3990/1.9789036537933

Alsadik, B., Gerke, M., & Vosselman, G. (2012). Optimal camera network design for 3D modeling of cultural heritage, (September). https://doi.org/10.5194/isprsannals-I-3-7-2012

Alsadik, B., Gerke, M., & Vosselman, G. (2013). Automated camera network design for 3D modeling of cultural heritage objects Automated camera network design for 3D modeling of cultural heritage objects. *Journal of Cultural Heritage*, *14*(6), 515–526. https://doi.org/10.1016/j.culher.2012.11.007

Alsadik, B., Gerke, M., Vosselman, G., Daham, A., & Jasim, L. (2014). Minimal camera networks for 3D image based modeling of cultural heritage objects. *Sensors (Switzerland)*, *14*(4), 5785–5804. https://doi.org/10.3390/s140405785

Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). *Speeded-up robust features (SURF). Computer Vision and Image Understanding* (Vol. 110). https://doi.org/10.1016/j.cviu.2007.09.014

Besada, J. A., Bergesio, L., Campaña, I., Vaquero-Melchor, D., López-Araquistain, J., Bernardos, A. M., & Casar, J. R. (2018a). Drone mission definition and implementation for automated infrastructure inspection using airborne sensors. *Sensors (Switzerland)*, *18*(4), 1–29. https://doi.org/10.3390/s18041170

Besada, J. A., Bergesio, L., Campaña, I., Vaquero-Melchor, D., López-Araquistain, J., Bernardos, A. M., & Casar, J. R. (2018b). Drone mission definition and implementation for automated infrastructure inspection using airborne sensors. *Sensors (Switzerland)*, *18*(4), 1–18. https://doi.org/10.3390/s18041170

Biljecki, F. (2017). *Level of details in 3D city models*. https://doi.org/10.4233/uuid:f12931b7-5113-47ef-bfd4-688aae3be248

Bircher, A., Kamel, M., Alexis, K., Oleynikova, H., & Siegwart, R. (2018). Receding horizon path planning for 3D exploration and surface inspection. *Autonomous Robots*, *42*(2), 291–306. https://doi.org/10.1007/s10514-016-9610-0

Buczkowski, A. (2018). Drone LiDAR or Photogrammetry? Everything you need to know. - Geoawesomeness. Retrieved March 1, 2019, from http://geoawesomeness.com/drone-lidar-or-photogrammetry-everything-your-need-to-know/

C Gonzalez, R., E Woods, R., & R Masters, B. (2009). *Digital Image Processing, Third Edition*. *Journal of biomedical optics* (Vol. 14). https://doi.org/10.1117/1.3115362

Cavegn, S., Haala, N., Nebiker, S., Rothermel, M., & Tutzauer, P. (2014). Benchmarking High Density Image Matching for oblique airborne imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *40*(3), 45–52. https://doi.org/10.5194/isprsarchives-XL-3-45-2014

Cefalu, A., Haala, N., Schmohl, S., Neumann, I., & Genz, T. (2017). a Mobile Multi-Sensor Platform for Building Reconstruction Integrating Terrestrial and Autonomous Uav-Based Close Range Data Acquisition. *International Conference On Unmanned Aerial Vehicles In Geomatics*, *42–2*(W6), 63–70. https://doi.org/10.5194/isprs-archives-XLII-2-W6-63-2017

Chen, S., Li, Y., & Kwok, N. M. (2011). *Active vision in robotic systems: A survey of recent developments*. *International Journal of Robotics Research* (Vol. 30). https://doi.org/10.1177/0278364911410755

Costante, G., Delmerico, J., Werlberger, M., & Valigi, P. (2017). Exploiting Photometric Information for Planning under Uncertainty, 1–16. https://doi.org/10.1007/978-3-319-51532-8_7

Costante, G., Forster, C., Delmerico, J., Valigi, P., & Scaramuzza, D. (2016). Perception-aware Path Planning, 1–16. https://doi.org/arXiv:1605.04151v1 1

Daftry, S., Hoppe, C., & Bischof, H. (2015). Building with drones: Accurate 3D facade reconstruction using MAVs. In *2015 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3487–3494). IEEE. https://doi.org/10.1109/ICRA.2015.7139681

Darmawan, H., Walter, T. R., Brotopuspito, K. S., Subandriyo, & I Gusti Made Agung Nandaka. (2018). Morphological and structural changes at the Merapi lava dome monitored in 2012–15 using unmanned aerial vehicles (UAVs). *Journal of Volcanology and Geothermal Research*, *349*, 256–267. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2017.11.006

Díaz-Vilariño, L., Khoshelham, K., Martínez-Sánchez, J., & Arias, P. (2015). 3D Modeling of Building Indoor Spaces and Closed Doors from Imagery and Point Clouds. *Sensors* . https://doi.org/10.3390/s150203491

Döllner, J., Buchholz, H., Nienhaus, M., & Kirsch, F. (2005). *Illustrative visualization of 3D city models*. *Proceedings of SPIE - The International Society for Optical Engineering*. https://doi.org/10.1117/12.587118

Dong, J., Gary Burnham, J., Boots, B., Rains, G., & Dellaert, F. (2017). *4D crop monitoring: Spatio-temporal reconstruction for agriculture*. https://doi.org/10.1109/ICRA.2017.7989447

DroneZon. (2019). Retrieved March 6, 2019, from https://www.dronezon.com/learn-about-drones-quadcopters/drone-3d-mapping-photogrammetry-software-for-survey-gis-models/

Emelianov, S., Bulgakow, A., & Sayfeddine, D. (2014a). Aerial laser inspection of buildings facades using quadrotor. *Procedia Engineering*, *85*(June 2015), 140–146. https://doi.org/10.1016/j.proeng.2014.10.538

Emelianov, S., Bulgakow, A., & Sayfeddine, D. (2014b). Aerial Laser Inspection of Buildings Facades Using Quadrotor. *Procedia Engineering*, *85*, 140–146. https://doi.org/https://doi.org/10.1016/j.proeng.2014.10.538

Engel, J., & Döllner, J. (2014). *Immersive Visualization of Virtual 3D City Models and its Applications in E-Planning*. *International Journal of E-Planning Research* (Vol. 1). https://doi.org/10.4018/ijepr.2012100102

Entrop, A. G., & Vasenev, A. (2017). Infrared drones in the construction industry: designing a protocol for building thermography procedures. *Energy Procedia*, *132*, 63–68. https://doi.org/https://doi.org/10.1016/j.egypro.2017.09.636

Erdelj, M., Natalizio, E., R. Chowdhury, K., & Akyildiz, I. (2017). *Help from the Sky: Leveraging UAVs for Disaster Management*. *IEEE Pervasive Computing* (Vol. 16). https://doi.org/10.1109/MPRV.2017.11

Farinella, G. M., & Mattiolo, G. (2006). Solutions to 3D Building Reconstruction from Photographs, (February).

Feifei, X., Zongjian, L., Dezhu, G., & Hua, L. (2012). Study on construction of 3D building based on uav images, *XXXIX*(September), 469–473.

Fischler, R., & Bolles, M. (1981). *Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography*. *Commun ACM* (Vol. 24).

Forlani, G., Roncella, R., & Nardinocchi, C. (2015). Where is photogrammetry heading to? State of the art and trends. *Rendiconti Lincei*, *26*(S1), 85–96. https://doi.org/10.1007/s12210-015-0381-x

Forstner, W. (1998). On the Theoretical Accuracy of Multi Image Matching , Restoration and Triangulation.

Furukawa, Y. (2010). Bundler - Structure from Motion (SfM) for Unordered Image Collections. Retrieved March 5, 2019, from http://www.cs.cornell.edu/~snavely/bundler/

Furukawa, Y., & Ponce, J. (2009). Pami08a.Dvi, *1*(1), 1–14. https://doi.org/10.1109/TPAMI.2009.161

Gao, X.-Z., Hou, Z.-X., Zhu, X.-F., Zhang, J.-T., & Chen, X.-Q. (2013). The Shortest Path Planning for Manoeuvres of UAV. *Acta Polytechnica Hungarica*, *10*(1). Retrieved from https://pdfs.semanticscholar.org/1fe4/b155ea9d03098fd4073c4f2ddc780ef6890e.pdf

Ghuneim, A. G. (2000). Contour Tracing. Retrieved March 16, 2019, from http://www.imageprocessingplace.com/downloads_V3/root_downloads/tutorials/contour_tracing_Abeer_George_Ghuneim/index.html

González, V., Monje, C. A., Moreno, L., & Balaguer, C. (2017). UAVs mission planning with flight level constraint using Fast Marching Square Method. *Robotics and Autonomous Systems*, *94*, 162–171. https://doi.org/10.1016/J.ROBOT.2017.04.021

Haala, N. (2013). The Landscape of Dense Image Matching Algorithms. *Photogrammetric Week '13*, 271–284.

Harwin, S., & Lucieer, A. (2012). Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from Unmanned Aerial Vehicle (UAV) imagery. *Remote Sensing*, *4*(6), 1573–1599. https://doi.org/10.3390/rs4061573

Hasegawa, H., Matsuo, K., Koarai, M., Watanabe, N., & Masaharu, H. (2000). *DEM Accuracy and the Base to Height (B/H) Ratio of Stereo Images*. *ISPRS Archives. XIXth ISPRS Congress Technical Commission IV. Vol. XXXIII-B4* (Vol. XXXIII).

Hirschm, H. (2007). Stereo Processing by Semi-Global Matching.pdf, 1–14. https://doi.org/10.1109/CVPR.2005.56

Hirschmüller, H. (2005). *Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information*. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Vol. 2). https://doi.org/10.1109/CVPR.2005.56

Hoppe, C., Wendel, A., Zollmann, S., Pirker, K., Irschara, A., Bischof, H., & Kluckner, S. (2012). Photogrammetric Camera Network Design for Micro Aerial Vehicles, (January).

Hu, F., Gao, X. M., Li, G. Y., & Li, M. (2016). DEM extraction from worldview-3 stereo-images and accuracy evaluation. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *2016–Janua*(July), 327–332. https://doi.org/10.5194/isprsarchives-XLI-B1-327-2016

Huang, R., Zou, D., Vaughan, R., & Tan, P. (2017). Active Image-based Modeling with a Toy Drone. Retrieved from http://arxiv.org/abs/1705.01010

Incekara, A. H., & Seker, D. Z. (2018). Comparative Analyses of the Point Cloud Produced by Using Close-Range Photogrammetry and Terrestrial Laser Scanning for Rock Surface. *Journal of the Indian Society of Remote Sensing*, 1–11. https://doi.org/10.1007/s12524-018-0805-z

insight3d. (2019). Retrieved March 5, 2019, from http://insight3d.sourceforge.net/

Isler, S., Sabzevari, R., Delmerico, J., & Scaramuzza, D. (2016). An information gain formulation for active volumetric 3D reconstruction. In *Proceedings - IEEE International Conference on Robotics and Automation*. https://doi.org/10.1109/ICRA.2016.7487527

James, M. R., & Quinton, J. N. (2014). Ultra-rapid topographic surveying for complex environments: The hand-held mobile laser scanner (HMLS). *Earth Surface Processes and Landforms*, *39*(1), 138–142. https://doi.org/10.1002/esp.3489

Jing, L., Yong-she, S., Jian-ronga, W., & Jun-fenga, Y. (2008). The research and design of the base-height ratio for the three linear array camera of satellite photogrammetry. … *Archives of the Photogrammetry,* …, (1), 757–760. Retrieved from http://www.isprs.org/proceedings/XXXVII/congress/1_pdf/130.pdf

Karaman, S., & Frazzoli, E. (2011). Sampling-based algorithms for optimal motion planning. *International Journal of Robotics Research*, *30*(7), 846–894. https://doi.org/10.1177/0278364911406761

Kobayashi, Y. (2006). *Photogrammetry and 3D city modelling*. https://doi.org/10.2495/DARC060211

Kolbe, T. H. (2009). *CityGML-OGC Standard for Photogrammetry? OGC Standard for Photogrammetry?* Retrieved

from http://www.ifp.uni-stuttgart.de/phowo/2009/presentations/270Kolbe.pdf
Kriegel, S., Rink, C., Bodenmüller, T., & Suppa, M. (2013). *Efficient Next-Best-Scan Planning for Autonomous 3D Surface Reconstruction of Unknown Objects. Journal of Real-Time Image Processing* (Vol. 10). https://doi.org/10.1007/s11554-013-0386-6
Kroon, D.-J. (2008). Region Growing. Retrieved from https://nl.mathworks.com/matlabcentral/fileexchange/19084-region-growing?focused=5098324&tab=function
Lowe, D. (1999). *Object Recognition from Local Scale-Invariant Features. Proceedings of the IEEE International Conference on Computer Vision* (Vol. 2).
Majdik, A. L., Verda, D., Albers-Schoenberg, Y., & Scaramuzza, D. (2015). Air-ground Matching: Appearance-based GPS-denied Urban Localization of Micro Aerial Vehicles. *Journal of Field Robotics*, *32*(7), 1015–1039. https://doi.org/10.1002/rob.21585
Martin, R. A., Rojas, I., Franke, K., & Hedengren, J. D. (2016). Evolutionary view planning for optimized UAV terrain modeling in a simulated environment. *Remote Sensing*, *8*(1), 26–50. https://doi.org/10.3390/rs8010026
Mathias Rothermel, Wenzel, K., Fritsch, D., & Haala, N. (2012). SURE - Photogrammetric Surface Reconstruction from Imagery. *Proceedings LC3D Workshop*, 1–9.
MeshLab. (2016). Retrieved March 5, 2019, from http://www.meshlab.net/
Mostegel, C. (2011). Supplementary Material : UAV-based Autonomous Image Acquisition with Multi-View Stereo Quality Assurance by Confidence Prediction, 1–7.
Motohka, T., Nasahara, K. N., Oguma, H., & Tsuchida, S. (2010). Applicability of Green-Red Vegetation Index for remote sensing of vegetation phenology. *Remote Sensing*, *2*(10), 2369–2387. https://doi.org/10.3390/rs2102369
Mouget, A., & Lucet, G. (2014). *Photogrammetric Archaeological Survey with UAV. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* (Vol. II-5). https://doi.org/10.5194/isprsannals-II-5-251-2014
Murtiyoso, A., & Grussenmeyer, P. (2017). Documentation of heritage buildings using close-range UAV images: dense matching issues, comparison and case studies. *The Photogrammetric Record*, *32*(159), 206–229. https://doi.org/10.1111/phor.12197
Nex, F., Gerke, M., Remondino, F., Przybilla, H. J., Bäumker, M., & Zurhorst, A. (2015). Isprs benchmark for multi-platform photogrammetry. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*(3W4), 135–142. https://doi.org/10.5194/isprsannals-II-3-W4-135-2015
Noor, N., Aiman, A., Bin, A., Abdullah, A., & Hashim, M. (2018). *3D City Modeling in reconstructing Traditional Malay Towns in Kelantan 3D city modelling in reconstructing Traditional Malay town in Kelantan*.
OGC. (2012). OGC CityGML Standard. *OGC Member Approved International Standard*, 1–344. https://doi.org/OGC 12-019
Palazzolo, E., & Stachniss, C. (2018). Effective Exploration for MAVs Based on the Expected Information Gain. *Drones*, *2*(1), 9. https://doi.org/10.3390/drones2010009
Pix4D. (2019). Retrieved March 5, 2019, from https://www.pix4d.com/
Rakha, T., & Gorodetsky, A. (2018). Review of Unmanned Aerial System (UAS) applications in the built environment: Towards automated building inspection procedures using drones. *Automation in Construction*, *93*, 252–264. https://doi.org/https://doi.org/10.1016/j.autcon.2018.05.002
Remondino, F., & El-hakim, S. (2006). Image - based 3D Modelling : A Review IMAGE-BASED 3D MODELLING : A REVIEW, (December 2017). https://doi.org/10.1111/j.1477-9730.2006.00383.x
Roca, D., Lagüela, S., Díaz-Vilariño, L., Armesto, J., & Arias, P. (2013). Low-cost aerial unit for outdoor inspection of building façades. *Automation in Construction*, *36*, 128–135. https://doi.org/https://doi.org/10.1016/j.autcon.2013.08.020
Roo, B. De, Bourgeois, J., & Maeyer, P. De. (2017). Advances in 3D Geoinformation Systems. In *Advances*

in 3D Geoinformation (pp. 323–335). https://doi.org/10.1007/978-3-540-72135-2

Scaramuzza, D., Achtelik, M. C., Doitsidis, L., Friedrich, F., Kosmatopoulos, E., Martinelli, A., … Meier, L. (2014). Vision-controlled micro flying robots: From system design to autonomous navigation and mapping in GPS-denied environments. *IEEE Robotics and Automation Magazine*, *21*(3), 26–40. https://doi.org/10.1109/MRA.2014.2322295

Themistocleous, K., Cuca, B., & Hadjimitsis, D. G. (2014). 3D documentation of Fabrica hills caverns using terrestrial and low cost UAV equipment, (November).

Torresan, C., Berton, A., Carotenuto, F., Di Gennaro, S., Gioli, B., Matese, A., … Wallace, L. (2016). *Forestry applications of UAVs in Europe: A review. Int. J. Remote Sens.* https://doi.org/10.1080/01431161.2016.1252477

Vasquez-Gomez, J. I., Sucar, L. E., Murrieta-Cid, R., & Lopez-Damian, E. (2014). Volumetric Next-best-view Planning for 3D Object Reconstruction with Positioning Error. *International Journal of Advanced Robotic Systems*, *11*(10), 159. https://doi.org/10.5772/58759

Vetrivel, A., Gerke, M., Kerle, N., Nex, F., & Vosselman, G. (2018). Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *140*, 45–59. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2017.03.001

Vetrivel, A., Gerke, M., Kerle, N., & Vosselman, G. (2015). Identification of damage in buildings based on gaps in 3D point clouds from very high resolution oblique airborne images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *105*, 61–78. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.03.016

Vladimirovich Kim, N., & Alekseevich Chervonenkis, M. (2015). *Situation Control of Unmanned Aerial Vehicles for Road Traffic Monitoring. Modern Applied Science* (Vol. 9). https://doi.org/10.5539/mas.v9n5p1

Wang, L., Xu, Y., Li, Y., & Zhao, Y. (2018). Voxel segmentation-based 3D building detection algorithm for airborne LIDAR data, 1–26.

Wolf, P. R., & Dewitt, B. A. (2013). *Elements of photogrammetry with applications in GIS* (4th ed.). McGraw-Hill Education.

Wu, C. (2011). VisualSFM. Retrieved March 5, 2019, from http://ccwu.me/vsfm/

Xie, H. (2018). Weighted Voxel : a novel voxel representation for 3D reconstruction, (August), 1–5. https://doi.org/10.1145/3240876.3240888

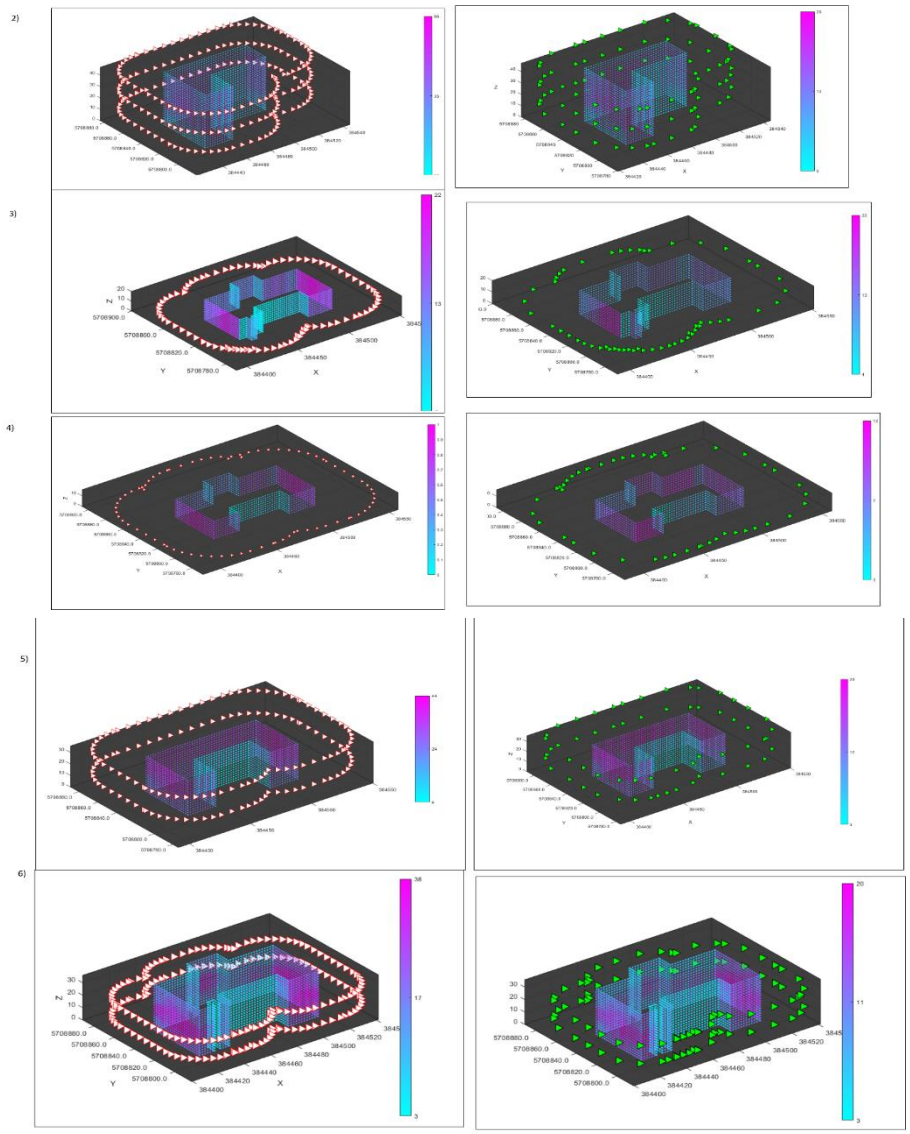
Zakaria, A., Abdul Shukor, S. A., Basah, S. N., Rushforth, E., & Wong, R. (2015). 3D Terrestrial Laser Scanner for Managing Existing Building. *Jurnal Teknologi*, *76*(12). https://doi.org/10.11113/jt.v76.5895

Zhang, Z., Gerke, M., Vosselman, G., & Yang, M. Y. (2018). A patch-based method for the evaluation of dense image matching quality. *International Journal of Applied Earth Observation and Geoinformation*, *70*, 25–34. https://doi.org/10.1016/J.JAG.2018.04.002

APPENDIX

Shape id	Cameras in DCN	Cameras in MCN	Highest coverage	Total façade points	Façade points visible in DCN	Façade points visible in MCN	Average visibility of points	Perimeter of building	Height of the building	Depth (D)
2	219	83	29	2784	89527	40227	14.4	224.6	45	30
3	89	57	14	1892	17776	9636	5	326	20	36
4	76	47	12	1467	16749	9405	6.4	312	15	42
5	148	60	20	2595	57992	26222	10	314.4	15	35
6	203	95	20	3268	53007	25911	7.9	327.9	35	26

APPENDIX 1 Table showing the results of visibility analysis and coverage filtering for synthetic building with the properties of it



APPENDIX 2 The dense camera network (left) and minimal camera network (right) with visibility of facade point cloud for synthetic buildings