School of Built Environment
Australasian Joint Research Centre for Building Information Modelling

Automated As-built Information Modelling for Construction Monitoring with Vision Data in Liquefied Natural Gas (LNG) Plants

Jian Chai

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

JULY 2017

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature: 

Date: 24/07/2017

Supervisor: Prof. Xiangyu Wang

Curtin-Woodside Chair Professor for Oil, Gas and LNG Construction and Project Management

Director of Australasian Joint Research Centre for Building Information Modelling

Curtin University

Co-supervisor: Dr. Hung-Lin Chi

Research Fellow, Curtin University

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Prof. Xiangyu Wang for offering me the opportunity to conduct my research, and a platform on which to do so. It has been a unique experience in pursuing a PhD degree at the BIM centre with its rich resources for study and research.

My thanks also go to my co-supervisor Dr. Hung-Lin Chi for his continuous and valuable support during my PhD. Without his support and contributions, I would not have reached this point. His patience, tolerance and immense knowledge has guided me over the past three years. I would also like to give a special mention to Dr. Changzhi Wu for his guidance with the research and to Mr Jiong Zhu for his collaboration with the scaffolding monitoring research.

Special thanks are given to Woodside for funding my study and research. In addition, I would like to thank the two anonymous reviewers, whose comments and suggestions have been valuable to this thesis.

Last but not least, I am grateful to my parents and sisters for their support throughout my life. They have provided financial and spiritual support along the way. Without their sacrifice and love, it would have been impossible for me to reach my present position in life. I give my special thanks to my girlfriend, who steered me to an unexpected but amazing life in Australia and stimulated me to survive the process of gaining my PhD. My parents, sisters and girlfriend, you are always there for me.

CONTENTS

ABSTRACT

Construction monitoring is of vital importance in ensuring a construction project is finished in compliance with time limitations and budget, especially for large projects such as Liquefied Natural Gas (LNG) plants. Effective project monitoring requires efficient and reliable as-built information collection and interpretation techniques. Currently, vision techniques with laser scanners and cameras are commonly used in construction sites, producing a large volume of 2D photos or 3D point cloud data. Even though this vision data contains a great deal of information regarding the as-built status of a construction project, the data is underexploited, and usually used for visualisation only or interpreted manually; it is not fully utilised in supporting construction management tasks. This research aims to develop automated methods of retrieving as-built information from vision data in the construction of LNG plants, including temporary structures like scaffolding, and to use these developed methods to improve construction productivity in the LNG industry.

This research presents a framework that integrates vision-based techniques including laser scanning and image capturing to monitor construction activities in LNG plants. Furthermore, the research investigates 3D point cloud collection with laser scanners and consumable cameras in LNG plants. Experiments were conducted using both laser scanners and cameras, and the performance of these is presented and analysed.

The research investigates class-level and object-level interpretation of the point clouds of LNG plants. It achieves semantic segmentation of industrial point clouds by using Markov Random Field (MRF) models. The segmentation process clusters

points into local patches and considers contextual information in a scenario with MRF. Tests with different data sets show that the implemented semantic segmentation approach achieves a high degree of accuracy after considering contextual constraints. Besides this, the research realises object-level industrial component recognition in cluttered environments by using local feature matching and clustering. Tests on different types of components show the feasibility and accuracy that the proposed method can achieve. The semantic interpretation approach can facilitate tasks such as rapid 3D environmental mapping of construction sites and progress tracking.

By developing scaffolding detection methods, this research contributes novel scaffolding monitoring methods, using cameras. The scaffolding detection method functions via a hierarchical strategy combining colour, topological patterns and layout of scaffolding in the photo. Furthermore, based on those recognised scaffolding components, more as-built information such as quantity and location of scaffolding can be inferred from the vision data. Moreover, this research implements a prototype system of the developed monitoring approach and tests the prototype in actual LNG plants. Extensive testing has shown that the developed scaffolding monitoring approach is feasible and effective in monitoring the various types of scaffolding used in LNG plants. The developed hierarchical scaffolding monitoring methods can be used in several scaffolding management tasks including material management, progress monitoring and dismantling planning. They can be further extended for use in other construction industries involving the utilisation of large amounts of scaffolding.

Keywords: as-built information, laser scanning, vision, scaffolding, construction monitoring, LNG plant

1 INTRODUCTION

This thesis presents research on the development of vision data interpretation approaches for promoting automated construction monitoring. Section 1.1 introduces the motivation and background of the research. Section 1.2 describes the objectives of the research and defines the research scope. Section 1.3 presents the organisation of the thesis.

1.1 Motivation and Background

As defined in the Project Management Body of Knowledge (Rose, 2013), "the monitoring and controlling process group consists of those processes required to track, review, and orchestrate the progress and performance of a project; identify any areas in which changes to the plan are required; and initiate the corresponding changes." Efficient and effective monitoring on progress and performance (product quality) are two important tasks of project monitoring and controlling. These ensure that a construction project complies with time and budget. Progress tracking involves tracking, reviewing and reporting construction progress; and quality control ensures the deliverables are as per the quality requirements defined. Both tasks depend on the effective capture and modelling of the as-built status of a project.

In the past decade, there has been an urgent need for fast and accurate project monitoring methods in the Architectural Engineering Construction/Facility

Management (AEC/FM) industry (Bosché et al., 2015). One common and promising technique for construction monitoring is the vision-based method. Research into improving the current construction monitoring process is worthwhile and necessary.

Given the importance of construction monitoring to the overall success of a project, recent studies have aimed to improve construction monitoring for more effective construction management in different aspects: (1) efficient and automated as-built data collection, (2) semantic interpretation of various data and extraction of construction-related knowledge, and (3) visualisation of information for effective decision-making.

As-built data collection: Researchers have investigated emerging technologies for automating and enriching raw data collection representing project status. Some handheld computing devices have been developed to improve data acquisition and sharing in current manual project tracking and controlling processes (Omar & Nehdi 2016). Sensing and tracking technologies, such as Radio Frequency Identification (RFID) and Global Positioning Systems (GPS), have also been studied to track and identify materials, equipment and personnel during the transportation, laydown, and construction phases (El-Omari & Moselhi 2011; Chai et al. 2017). Many works have been proposed which use laser scanning to speedily acquire geometric information on objects, and they are used for as-built modelling, progress tracking and quality control purposes (Bosché et al., 2009; Tang et al. 2010; Chai et al. 2016). Video and imaging techniques have also been used for tracking personnel, modelling as-built status and monitoring operation productivity (Park et al. 2011; Golparvar-Fard et al. 2012; Klein et al. 2012; Yang et al. 2016).

Semantic interpretation: The processing and interpretation of various raw data to produce knowledge needed for construction management plays a key role in construction monitoring and this has been widely studied in recent years. Localisation and identification data from sensors can be directly applied to analyse construction activities, and many researchers have focused on improving the localisation reliability and accuracy of sensing and tracking approaches (Ding et al. 2016). Laser scanning directly produces 3D point clouds capturing explicit

geometric information. This 3D information can be compared with the as-designed 3D model to analyse spatial or progress differences (Anil et al. 2011; Turkan et al. 2012). A few researchers are also investigating the semantic segmentation and object recognition of point clouds to create semantically enriched as-built building information models. Barazzetti (2016) developed methods for reconstructing parametric as-built models of complex shapes from photogrammetric and laser scanned point clouds using Non-Uniform Rational B-spline (NURBS) curves and surfaces. Image processing and computer vision techniques have increasingly drawn attention from the field of construction management and they have been investigated to efficiently monitor construction projects. 3D reconstruction of as-built models with photogrammetry and 3D computer vision has been widely used in past decades (Dai et al. 2013). Information retrieval from images has also been studied. For instance, Zhu and Brilakis (2010) used colour, texture and shape information to recognize concrete columns from images. Son et al. (2014) investigated the use of texture, colour and shape information in the classification of construction materials. Yang et al. (2016) investigated machine learning algorithms to recognise actions of construction workers using video, and this reached an average accuracy rate of 59% for 11 common types of actions.

Visualisation and decision-making: Another aspect of construction monitoring is found in making decisions according to interpretation of data and in visualising the results. Building Information Modelling (BIM) is an effective management concept and tool for synthesising as-built information and as-designed information to support decision-making. Navisworks, as a commercial BIM platform has been commonly used to colour-code and visualise construction progress (Marzouk & Zaher 2015). BIM is applied to identify and visualise hazards in construction sites by tracking on-site machines and personnel with wireless sensors (Park et al. 2016). Tang et al. (2010) used as-built BIM models to analyse and visualize surface flatness. Mixed Reality (MR) technologies, including Augmented Reality (AR) and Virtual Reality (VR), are effective in visualising information from construction sites, and this can enhance visualisation and increase user comprehension. Golparvar-Fard et al. (2009)

developed a system called D^{4}AR which used AR to visualise 4D models for monitoring construction progress. Zollmann et al. (2014) implemented an AR system for the on-site presentation of progress information by superimposing as-planned information onto the user's view of the physical environment.

Although extensive research has been conducted on various aspects of construction monitoring, the full potential of vision-based data for construction monitoring has not yet been realised. The huge potential of the vision-based monitoring technique and its significance to project management requires further investigation. The main concerns of this research on state-of-the-art developments are presented as follows:

As-built information modelling for existing facilities without as-designed models

There are many existing plants, buildings or infrastructures, which lack pre-existing BIM or CAD models. For example, more than 60% of residential buildings in Europe were built before 1990 (Economidou et al. 2011) and most do not have BIM models. Many automated or semi-automated approaches for as-built modelling for facilities are based on comparisons with as-designed BIM models, and these are known as "Scan vs BIM" schemes (Bosché et al. 2013). These methods compare point clouds aligned with as-designed models and thus cannot be applied to those existing facilities without as-designed 3D models.

Recent research on converting captured data into as-built models, without the assistance of as-designed BIM, known as "Scan to BIM", focused mainly on geometric rather than semantic information. In current practice, it is usually a manual process of converting point cloud data to BIM, making it costly and subjective. It is the core task of as-built modelling to detect and recognise predefined infrastructure elements. Thus, it will be of significance to investigate automated methods of segmenting and recognising point cloud without the assistance of as-designed models.

Limited attention to temporary structures on construction sites

Temporary structures are widely utilised in construction projects. Scaffolding is one kind of temporary structure commonly utilised in supporting materials, resources and permanent structures. Despite a large amount of research into automated construction monitoring, there has been little focus on temporary structures such as scaffolding. A Cyber Physical System (CPS) for scaffolding structure health monitoring was designed by deploying various wireless sensors on key nodes of structures and it was then evaluated with simulation (Yuan et al. 2016). Xu et al. (2015) tried to reconstruct 3D models of scaffolding with photogrammetry methods with 40% (of limited types) of the scaffolding components reconstructed being of good quality. Existing work on scaffolding monitoring is still in its initial stages and there are still many aspects that have not been appropriately addressed, such as the quantity estimation of erected scaffolding components. The development of a feasible, low cost and effective scaffolding monitoring approach would benefit scaffolding projects in terms of material management and progress tracking, further promoting the productivity of entire construction projects.

Recognition of irregular objects in industrial plants

Object recognition from point clouds involves recognising object components and their characteristics relevant to monitoring tasks from the acquired raw data. Current recognition methods can be categorised into model-based, data-based and other approaches (Golovinskiy et al. 2009; Mian et al. 2016), and these have been applied to recognise building objects in construction sites, such as floors, ceilings, windows, or walls (Vosselman & Dijkman 2001; Tang et al. 2010; Wang et al. 2015). While a lot of research has investigated methods of recognising building components, the problems of recognising industrial components have not been fully addressed. Many studies have proposed methods of recognising and modelling planes or pipes in industrial point clouds. However, other complex process components, such as valves, have not been considered. Recognition of those complex components could lead to the production of more complete as-built models. In addition, the effective object recognition of complex components also depends on the successful segmentation of objects. Local features-based approaches invariably suffer from over-segmentation

by splitting an object into several areas (Aijazi et al. 2013). Semantic segmentation of point clouds in industrial plants by considering high-level constraints can be helpful in the recognition of and the as-built modelling from point clouds. Here semantic segmentation means segmenting point clouds into semantic and functional parts as in a designed model.

Potential of automated vision data interpretation for construction management

Vision techniques can effectively capture on-site information at high speed and low cost. There is currently a great deal of vision data available on construction sites such as those from CCTV records and daily images. In the past, vision based monitoring has been applied to construction management tasks to improve efficiency and productivity. However, current usage of vision data depends heavily on manual interaction and limits its applications to high-level tasks. Advancement in image processing (Ruiz-Sarmiento et al. 2017) and machine learning (Qi et al. 2016) has greatly improved the performance of vision data recognition for high-level tasks and this should promote the interpretation of vision data for construction purposes. With a large volume of vision data available, along with novel methods on vision data understanding (Armeni et al. 2017), it is worthwhile investigating how these new methods can promote vision data interpretation in construction projects.

Given the current developments cited above, this research investigated methods of vision-based monitoring for Liquefied Natural Gas (LNG) plant construction and maintenance projects. The research will address the abovementioned concerns, and promote the productivity and automation level of monitoring processes in a LNG plant by using efficient data collection and automated data interpretation techniques. Without as-designed models, machine-learning techniques may be used to semantically classify and parse as-built scenes from point cloud data, and further interpret the quantity and other construction-related information from raw data. This research also uses a strategy based on local-feature matching to recognise irregular components in industrial plants. To consider temporary structures, this research developed an approach for automatically monitoring scaffolding with computer

vision techniques. These developed approaches are expected to tackle key problems and promote the development of a vision-based construction tracking system.

1.2 Objective and Scope

The objective of this research is to develop automated vision-based construction monitoring methods with object recognition techniques to improve the automation level and effectiveness of construction monitoring and control. Specifically, this research is focusing on two major challenges: 1) interwoven and self-occluded geometries of LNG objects, causing difficulties of effective segmentation and robust recognition; 2) linear features of temporary scaffoldings that are easy to be missed and mixed with permanent parts of LNG facilities due to similar materials and looking, so that causing difficulties of reliable monitoring of temporary structures.

In this research, vision-based monitoring takes laser scanners or cameras as input devices, uses computer vision methods to process input data and produces information related to monitoring tasks. Those monitoring tasks mainly include semantic segmentation of point clouds, industrial components recognition, scaffolding structure recognition and quantity estimation. These tasks are mainly to promote the interpretation level of 2D/3D vision data for construction management purposes from merely geometric to semantic or construction-related knowledge. For example, rather than manual counting, this research developed an automated method for scaffolding recognition and quantity estimation based on on-site photos, which can be used in the material management and progress tracking of a scaffolding project. In order to validate the developed monitoring approaches, this research collected data from real industrial sites to test its performance. The outcomes of the research could be useful in more effectively applying vision data to construction progress and quality control, along with the better management of scaffolding structures in a construction project.

To define the scope of this research, the vision-based monitoring processes were organised into three phases: raw data collection, as-built modelling and outputs related to specific applications, as shown in Figure 1.1. The data collection mainly

refers to capturing vision measurements including point clouds and photos. To make it clearer, this research uses vision data to represent both 3D point clouds and 2D photos. Point clouds come from laser scanning or 2D photos with Structure From Motion (SFM) techniques. Other information related to a construction project, such as project schedules, is also collected. The as-built modelling process represents interpretation of raw vision data to knowledge reflecting current construction status so as to support decision-making on different tasks. Specifically, it involves semantic segmentation of point clouds and scaffolding quantity estimation from photos. The produced as-built information can be used to support visualisation and decision-making tasks, which are related to specific tasks.

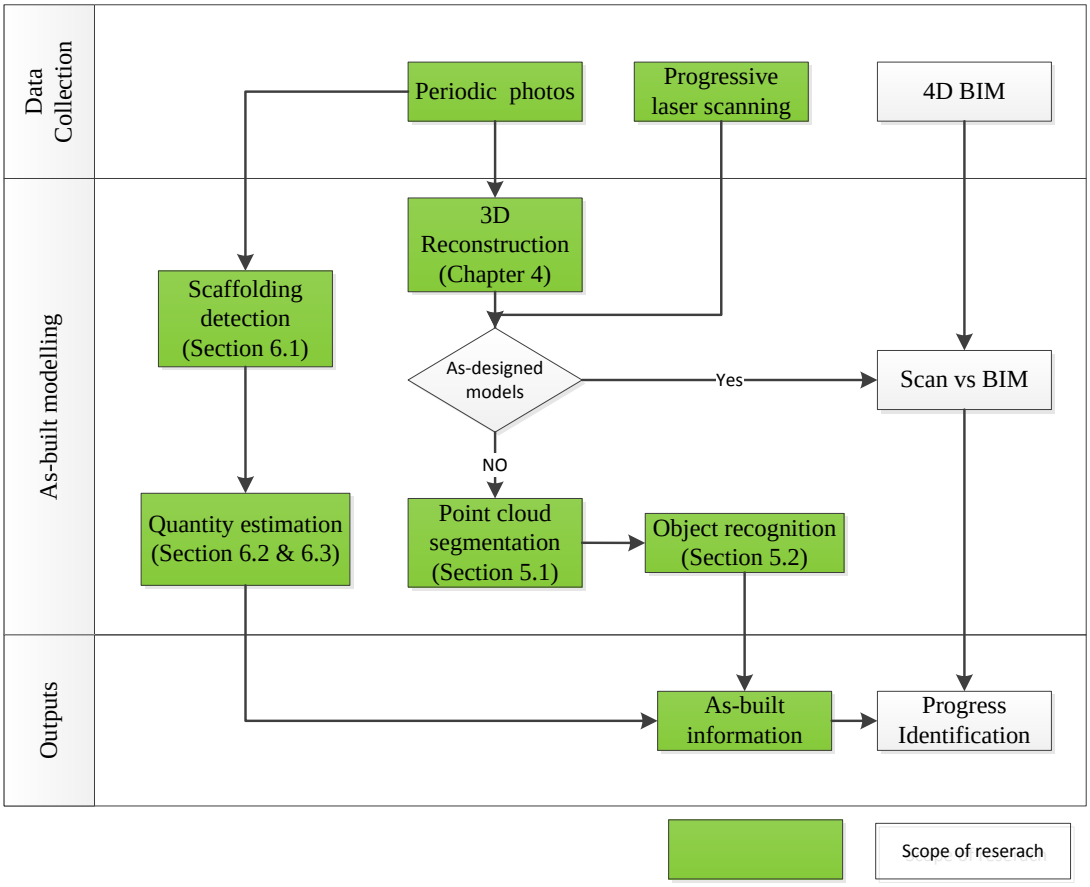

Figure 1. 1 Scope of the proposed research

The scope of this research project is shown in Figure 1.1, where green areas mark our main focus. It shows three major differences to previous research undertaken. Firstly, as-built information modelling and progress tracking with the "Scan vs BIM" strategy is outside the scope of this research. Secondly, previous research was mainly aimed at buildings or infrastructure. The subject of this research focused on the LNG plant, a typical kind of industrial plant. A LNG plant is a hazardous environment requiring intrinsic safety, which means the energy, electrical and thermal of all electrical equipment operated in the plant are limited to a level below that required to ignite a specific hazardous atmospheric mixture. Thirdly, this research includes scaffolding monitoring to tackle the challenge of temporary structure monitoring, which has seldom been addressed in previous research. Scaffolding components are large in quantity, small in size and scattered in distribution and those characteristics pose challenges in using vision-based monitoring techniques. Therefore, it is important to propose a reliable linear feature detection and validation approach for reliably detecting linear features and reliably quantify amounts of scaffoldings from images.

The tasks defined in the scope were to fill various gaps and challenges in the current monitoring process in a LNG plant construction or maintenance project. Firstly, the research compared characteristics of different methods of capturing point clouds when applied to a LNG plant. Compared to buildings, the severe occlusion and similar appearance of industrial components would incur challenges for point cloud production from photos which require line-of-sight. The completeness and accuracy of image-based 3D reconstruction when applied to LNG plants is evaluated. Due to the irregular and complex characteristics of LNG components, scaffolding is widely used during construction and maintenance phases to support materials and working personnel in LNG plants. Current monitoring of scaffolding construction and dismantling largely depends on manual reports and incurs large costs and results in low productivity during LNG projects. The research addresses the monitoring of the scaffolding widely utilised in a plant with a single photo. Large volumes of point clouds accurately reflect the as-built dimensions of LNG components. However,

current data processing involves intensive manual input. Automation of object recognition is of vital importance to improve efficiency and promote wider applications. Thus, the research investigates automated semantic segmentation of industrial point clouds produced from laser scanning or photos, and recognition of irregular components, without the assistance of as-designed models.

The main technical contributions of this thesis include: 1) extending Support Vector Machine method with Markov Random Field method and improved the recall and precision of object recognition from point clouds; 2) processing voxels rather than points for enhancing the computational efficiency; 3) improving the line detection, clustering, and validation methods to ensure detection of all scaffolding elements that are shown as linear features in images while eliminating incorrect detections of scaffolding elements.

1.3 Organisation of the Thesis

The chapters of the thesis are organised as follows:

Chapter 2 introduces several existing techniques related to construction monitoring from data collection, data interpretation to visualisation and decision-making. Previous work related to this research is also reviewed. Previous work on monitoring temporary structures in construction is also described, including its importance and its achievements. This chapter also describes the background and current research on point cloud segmentation and recognition.

Chapter 3 presents a brief framework of this research. This chapter illustrates the roles and possible applications of the developed monitoring approaches in existing construction management processes.

Chapter 4 describes two methods of collecting 3D data in industrial plants, including laser scanning and image-based 3D reconstruction. This chapter presents evaluations of the two methods by conducting experiments in a LNG facility. In addition, this chapter presents tests of 3D reconstruction results of scaffolding with different image

data. Performance of reconstruction with different data is compared in terms of accuracy and completeness.

Chapter 5 presents a semantic segmentation approach with the Markov Random Field. The proposed segmentation approach performs more effectively and efficiently by considering contextual constraints in LNG plants and operating supervoxels rather than points. The chapter describes features of point clouds and their estimation methods. Supervoxelisation and classification with Support Vector Machine (SVM) are then presented in detail. At the end of this chapter, experiments are presented along with the analysis of the performance of the proposed method. This chapter also presents an approach to point-cloud-object recognition. Object recognition based on local features is described. Further, this chapter presents the performance of 3D Scale-Invariant Feature Transform (SIFT) based recognition when applied to industrial plants.

Chapter 6 presents scaffolding monitoring with photos. A hierarchical approach to scaffolding detection from photos is firstly presented. It then presents in detail the different evaluation methods for detection, including colour-based segmentation, depth thresholding and linear feature detection in various scenarios of scaffolding. Following this, methods estimating scaffolding quantities, including pattern matching and linear feature clustering, are also presented. Lastly, this chapter presents experiments in a LNG plant, evaluating the performance of the developed methods.

Chapter 7 summarises the proposed research and outlines future directions. The work, results and contributions are summarised in detail. Future work that could promote enhanced construction monitoring is also detailed.

2 LITERATURE REVIEW

In this chapter, previous work on the development of related construction monitoring technologies is presented and reviewed. Related works include developed techniques of data collection, information retrieval, visualisation and decision-making. These are described in Section 2.1, 2.2 and 2.3 individually. In particular, Section 2.4 presents previous research regarding the development of temporary structure monitoring approaches. In addition, Section 2.5 describes the available approaches for semantic segmentation and object recognition of point clouds.

2.1 Data collection techniques for construction monitoring

One basic task of monitoring a construction project is that of raw data collection which reflects information on a project. Conventionally, this is a manual process conducted through on-site inspections and reports. In past decades, there has been a great deal of work carried out on automating data collection by integrating innovative techniques. According to the equipment used, these techniques can be divided into three categories: Information Technology-based communication methods, wireless tracking and vision-based tracking methods.

Many Information Technologies have been used to reduce manual processes, and to promote information communication and sharing during a construction project. These include multimedia, email services, voice-based tools, Short Message Services

(SMS) and handheld computing tools (Omar & Nehdi 2016). Handheld computing tools, such as tablets, smart phones and personal digital assistants, enable on-site personnel to access a variety of information and to update and upload information at sites in real time. Interactive voice response can automatically collect voice messages from users and is used to improve the efficiency of collecting and recording on-site information. Multimedia information, including videos, photographs and voice recordings, enriches conventional reports by visualising and highlighting objects of interests. Many information management systems have been implemented to integrate multimedia with current management processes, which can improve the efficiency of data collection and communication, and minimise the time and cost of on-site information collection. For example, Chi et al. (2016) developed an inspection system combining mobile computing devices with a BIM platform for better visualisation and communication. iConstruct (2017), a Navisworks plugin, has extended the functionalities of the BIM platform, allowing 3D model components to be linked with related drawings, photos and messages from sites. Envision (2017), an Ennova cloud-based product, allows practitioners to communicate regarding construction progress through texts or images in a social-network-linked online platform. Other applications, such as material management (Tserng et al., 2005; Tsai et al., 2007), progress tracking (Abudayyeh, 1997; Abeid and Arditi, 2002), and quality control (Chen and Luo, 2014), have also been widely proposed using information technologies. Despite the benefits, these technologies still depend on manual interaction and require extensive user training. Thus, their perspectives on automated monitoring are limited.

One strategy of construction monitoring is based on localisation of on-site materials, resources and personnel by utilising wireless positioning techniques, such as WIFI or RFID. Localisation of targets offers significant potential in improving construction monitoring tasks by location-aware computing and context-aware information delivery. The Global Navigation Satellite System (GNSS) is the most commonly applied localisation approach in outdoor environments. Although it is high in accuracy and global coverage, it lacks support for indoor or congested environments.

There have been other technologies used for indoor areas, including Wireless Local Area Networks (WLAN), Ultra-Wide Band (UWB) and Indoor Global Positioning Systems (GPS). The WLAN positioning system is based on radio signals and its positioning accuracy can be as high as 2 metres. RFID is similar to WLAN and typically employs Received Signal Strength to infer the location of a tag or a reader. The RFID system is easily and cheaply deployed and it is widely used in construction. However, its accuracy varies depending on system configurations. It is difficult to attach RFID tags to all on-site objects and the deployment of a RFID system in large and complex scenes incurs heavy costs. An indoor GPS works like a satellite navigation system and comprises transmitters and a receiver. A receiver calculates a position by the triangulation based on signals from four transmitters. Although the indoor GPS is costly, its positioning accuracy is high, at around 2 centimetres. UWB determines the location of a tag using radio frequency and operates on bandwidth of over 1GHz. The accuracy of the UWB positioning system is about 50 centimetres. A detailed review on the performance of different indoor localisation methods was presented in Luo et al. (2011); and Khoury & Kamat (2009). These localisation or identification techniques have been applied in many construction monitoring tasks. For instance, Song et al. (2007) presented an approach using RFID for finding the precise location of a tagged component on construction sites. Naticchia et al. (2013) utilised a low-level Zigbee-based tracking technology to monitor possible interference on large construction sites. Vahdatikhaki et al. (2015) developed an optimisation approach to estimate the poses of excavators through the feedback of UWB sensors at site. Other applications have also investigated inventory management (Lu et al., 2011), enhanced safety (Lee et al., 2012; Teizer et al., 2013) and work permit control (Kelm et al., 2013).

Sensor networks are also deployed in acquiring data on the structural performance of constructed structures in a dynamic environment, such as deflection, load, strain and inclines. A sensor network consists of multiple small, lightweight and portable detection points called sensor nodes. Each is equipped with a transducer, a transceiver, computing unit and power source, allowing communication and

transmission among the nodes through the commands sent from a central computer. Traditional wired sensor systems use coaxial wires to transmit data collected by sensors which is expensive, inefficient, and susceptible to disturbance and inflexibility, especially in dynamic construction sites. According to development trends in microelectromechanical system (MEMS) technology in recent years, wireless sensing and integrated circuit technology have seen increasing interest and made promising achievements in applying wireless sensor networks to structural health monitoring in construction (Moon et al. 2011). Various wireless sensors, including load cells, switch sensors, accelerometers and displacement sensors, attached on key nodes of structures can acquire information on structural loading, movement and pressure, which is useful in preventing potential structure failures (Zhou & Yi 2013). The sensor can also effectively collect and send back data in real time by integrating a wireless communications network. By integrating virtual information on constructed objects, wireless sensor networks can also implement a cyber-physical system (CPS) (Hackmann et al. 2014), which can extract the data related to health monitoring and enables bidirectional coordination between physical systems and their virtual representations. However, there are still challenges remaining in the deployment and operation of a wireless sensor network, such as high cost, power supply and optimal localisation of sensors (Iqbal et al. 2015). This is especially true for a LNG plant as its large scale and intrinsically safe requirements can incur huge costs.

Other commonly used techniques are vision-based monitoring methods. Vision-based monitoring methods are approaches which collect site information through a vision capturing process, such as the utilisation of laser scanning and imaging. These techniques use vision sensors to acquire measurements of the surface of an object and they use point clouds or images as inputs representing the as-is status of a construction project. The characteristics of vision data include its intuitiveness regarding the human perception system, accuracy in describing object surfaces and completeness in reflecting situations in the environment. However, the requirements of data processing are significantly high in terms of the amount of data captured and

the computation power required for data analysis (Wu et al. 2014). Depending on different features of the collected data in different scenarios, further development of data processing approaches is usually required in order to achieve a reasonable level of efficiency and effectiveness.

There are different platforms for laser scanners, including terrestrial, mobile and aerial scanners. As laser scanning directly produces 3D point cloud measurements with high accuracy and speed, it is widely used in collecting on-site information for many tasks. These tasks include: construction progress measurement (S El-Omari & Moselhi 2008; Turkan et al. 2012), urban excavation monitoring (Su et al. 2006), earthwork volume estimation (Du & Teng 2007), dimension control (Bosché 2010), quality assurance and control (Gong et al. 2014), as-built modelling (Tang et al. 2010), health monitoring (Park et al. 2007), structure inspection (Son et al. 2002), damage assessment (Olsen et al. 2009), safety design and planning (Marks et al. 2013) and geometrical documentation (Gikas 2012).

Table 2. 1. Sensing technologies used in construction monitoring

Data Collection	**Laser scanning**	**Photos**	**Barcode/RFID/UWB**	**GPS**
Location	✓		✓	✓
Identification			✓	
3D model	✓	✓		
Material		✓		

Sensing technologies and their captured data are presented in Table 2.1, where a mark indicates whether a data collection method could produce corresponding information. It can be seen that there is no single method of capturing all kinds of information. The integration of different techniques can provide more data on a physical construction site and possibly promote the interpretation and understanding of construction activities.

2.2 Construction-related information retrieval methods

After data capture for construction monitoring, another essential task is the processing and interpretation of raw data to produce construction task-related information and knowledge. While an explicit 3D point cloud model is produced from a laser scanner, images can be used to produce a 3D point cloud model through structure from motion and multi-stereo techniques (Agarwal & Snavely 2009). With the advancement of computer vision and photogrammetry techniques, it is easy and convenient to produce 3D models with non-survey cameras. There are off-the-shelf systems that enable 3D reconstruction from unordered images, such as commercial products Photoscan (Agisoft 2017) and open source tools Visual SFM and MicMac (Galland et al. 2016). Although such developments are promising, there are still challenges for the 3D reconstruction of complex structures from photos. Furthermore, the geometric accuracy of 3D model from photos is not as high as that of a laser scanner. Previous research shows that only 40% of limited types of scaffolding components, including tubes and toeboards, are successfully reconstructed with good quality from images (Xu et al. 2015).

Despite achievements in 3D reconstruction, there remain many challenges in the recognition and interpretation of vision data, which limit the applications of vision data in high-level tasks. One strategy to extract information from 3D models produced by laser scanning or imaging is to employ a deviation analysis with as-designed models (Anil et al. 2013). Son and Kim (2015) proposed the reconstruction of a 3D model of industrial plants from point clouds based on connectivity relationships from Piping and Instrumentation documents (P&ID). When a 3D CAD or BIM model is presented, the "Scan-vs-BIM" approach is used to detect, recognise and identify point cloud models (Bosché et al. 2015). This approach first aligns a point cloud in the same coordinate system with a BIM model, and then recognises objects contained in the BIM model through matching and comparisons. Point clouds from images are pre-processed in a similar way to laser scanned data reconstructing 3D models. For scenes without as-designed models, knowledge-based recognition, which recognises model components by human expert knowledge about

the scene, the objects and their representations, is investigated to promote automated understanding and interpretation of point clouds (Hmida et al. 2013). Belsky et al. (2016) developed a prototype system to infer semantic information of building models based on a set of geometry and topology rules, which were compiled by domain experts. Other 3D objects like a Hough Transform and local features can be used to recognise point cloud objects without as-designed models (Chai et al. 2015).

Besides the reconstruction of 3D models, raw image data is also applied to detect semantic information and site activities with pattern recognition and machine learning approaches. Brilakis et al. (2006) proposed to recognize material cluster of an image by matching with sample material images and then to use recognized material to retrieve images in construction sites. Zhu et al. (2011) used artificial neural network to recognise bridge concrete columns based on visual features (Zhu & Brilakis, 2010) from images. Dimitrov and Golparvar-Fard (2014) developed a material classification approach to recognising construction materials and created a library containing 20 typical construction materials. The classification approach used statistical features of filter responses and a Support Vector Machine (SVM) classifier. The developed material classification technique was further integrated to recognise different elements of 3D models and to achieve operation-level construction progress monitoring (Han & Golparvar-Fard 2015). Yang et al. (2016) investigated SVM and various features to recognise actions of construction workers using videos. Although only an average accuracy of 59% for 11 common types of actions was reached, the work is promising for future real world applications.

Although current "Scan vs BIM" object recognition systems provide valuable information for construction sites, a large volume of vision data available in the construction industry is underexploited, and efficient processing and interpreting data from sensing technologies is still a challenge. Worthy of promise is the integration of laser scanning techniques with photo-based 3D modelling to promote the tracking and identification of on-site construction objects (Samir El-Omari & Moselhi 2008).

2.3 Data Visualisation and Decision Making

Initially, 2D visualisation was popularly used by displaying object trajectories on a 2D map, such as layout maps of a construction site. For instance, Pradhananga and Teizer (2013) overlaid GPS data on a construction site map, enabling users to interactively analyse equipment operations based on their locations. Many corporations (Geoforce, 2017; Track'em, 2017) provide tracking solutions for supply chain management and utilise 2D maps for large-scale construction vehicle positioning. 2D visualisation aids managers in making better decisions by tracking equipment and identifying site layout, particularly in the location information of subjects. However, those tracking solutions are difficult to extract due to the lack of third dimensional information in 2D maps, as is the case in the real world. Lacking of 3D information sometimes results in misinterpretations of on-site situations and further affects the quality of decision-making during construction. Even though this can be compensated for by using multiple 2D displays with different views of the space, the level of interpretation abilities required from a manager or an engineer is still considered high, especially for a complex construction case.

3D visualisation technologies including 3D CAD, BIM, Virtual Reality (VR) and Augmented Reality (AR) have been widely applied in construction management processes for site monitoring. These technologies provide more intuitive information for accurate decision-making during construction. A comparative study done by Hsieh and Lu (2012) showed a consistent improvement in efficiency of about 50% in the interpretation of monitoring data using 3D digital field models versus conventional 2D maps and bar charts. Teizer et al. (2013) proposed the use of a 3D virtual environment to track and visualise workers in a construction ironwork scenario for safety and productivity training. The resources (e.g. personnel, equipment and materials) of field operations can be tracked by the use of Ultra-Wide Band (UWB) and all the activities can be identified through the 3D environment. 4D visualisation (3D + time) for a workspace collision detection process has been proposed considering both time and space of activities in real time, which increases the accuracy of the detection. Besides this, 4D modelling considers changes inside

the constructed product, providing more information on changes of a further layout plan (EINimr et al., 2016). A 5D (4D + cost) virtual model has been proposed to visualise and simulate the emissions of a construction plant, and to identify potential hazards on construction sites (Wong et al. 2014).

All these 3D derived visualisations (e.g. 4D, 5D and nD) can be categorised as BIM applications because their models are all built based on a 3D virtual environment, and related 3D components are associated with additional information other than geometric properties only. BIM is treated as a visual database which integrates a building's dimensions and attributes. It is often used in static analysis and the comparison of construction processes (Underwood and Isikdag, 2009). By extending BIM with VR technologies, which puts more emphasis on the user's immersive experience and interactions, BIM models can be interpreted by managers more easily through enabling VR visualisation platforms, such as VR headsets (Heydarian et al., 2015), large-scale visualisation display (Gurevich and Sacks, 2014) and even mobile tablets used at sites (Kuenzel et al., 2016). In the research done by Chen and Teizer (2013), they proposed an interactive VR environment gathering real-time resource location data for safety and activity monitoring on construction sites. They also pointed out that one of the important challenges of VR applications is the integration of accurate and real-time field data. The data, such as spatial information of the as-built scene, can only become valuable in the VR visualisation once errors are filtered and data processed to convey true information.

As for AR applications of site monitoring, bringing the planned BIM models as virtual mockups to the site and overlapping such mockups with the real ones directly shows the derivations of projects. Golparvar-Fard et al. (2009) utilised AR to monitor construction progress and to display 4D construction models for management purposes. Schall et al. (2009) and Talmaki et al. (2013) used AR to display the positioning information and the layout of an underground infrastructure and showed how to mitigate any undesired damage to its structures. Zhou et al. (2017) implemented an AR approach for segment displacement inspections during a tunnelling construction. The result was positive in terms of feasibility, efficiency and

cost-effectiveness. AR-enabled applications could be one of the intuitive visualisation solutions, as long as the reliability of detecting and positioning algorithms is acceptable and the occlusion issues of virtual objects on real ones can be solved in advance (Behzadan et al., 2015).

In terms of data retrieval, vision-based data captured from sites is an appropriate resource which can be associated with BIM, and this can maximise the benefits given the homogeneity between the collected vision-based data and 3D models. Managers can rapidly and intuitively make comparisons between the design model and the actual situation of the construction. However, the data processing effort required is great, and therefore automated approaches are currently being studied. In the research work done by Bosché et al. (2010), the virtual environment has been used for recognition of BIM objects in laser scanned point clouds. Field tests gathering such scanned point clouds during the construction of a green field power plant project produced convincing results. Turkan et al. (2012) further improved such 3D visualisation systems which were enhanced by linking the 3D model and the construction schedule, effectively creating a 4D visualisation system. Similar research work was carried out by Golparvar-Fard et al. (2015), where field images were used as a data resource in building construction scenarios. The productivity analysis of the image analysis results for progress tracking can be directly displayed on the BIM platforms which better support project systems like billing. Turkan et al. (2013) developed an earned value tracking approach toward automated progress monitoring by using 3D imaging tools including laser scanning.

In summary, information-rich visualisation, such as 3D, 4D and 5D simulations for site monitoring can be beneficial to managers in identifying deviations and implicit risks through the current status of construction projects. A BIM-enabled visualisation along with the collected vision-based data can be one of the most intuitive solutions for managers in improving the quality of their decision making. The challenges faced currently in construction projects, including the tedious work effort in the pre-processing of the sensed data and the low level of automation, are still the main obstacles in the development of mature site monitoring visualisations.

2.4 Research on temporary structure monitoring

The term "temporary structures" refers to systems and assemblies used for temporary support or the bracing of permanent work during construction, and structures built for temporary use. The temporary structures are defined as the elements of civil engineering works, which support or enable permanent work. The structures included at sites are temporary support systems such as earthwork sheeting and shoring, temporary bracing, soil backfill for underground walls, formwork systems, scaffolding, and underpinning of foundations. Scaffolding is widely utilised in construction industry and is used especially when such construction is measurably high in elevation. A large amount of scaffolding is constructed and dismantled during the course of a facility construction for a large construction project, such as a Liquefied Natural Gas (LNG) plant. Effective management of scaffolding construction is important not only to the success of the scaffolding work but also to the entire construction project. In fact, well-designed and well-established scaffolding is indispensable as a basis for further operations to be conducted safely, smoothly, efficiently and successfully.

Some research has been conducted to optimise the performance of scaffolding design and planning with computer aided design. Wang et al. (2016) developed a scaffolding scoping and design platform by integrating mobile devices and BIM to promote the communication between stakeholders during the planning and design phases. Kim & Teizer (2014) proposed automated safety checking and planning algorithms to identify and visualise potential hazards related to scaffolding design. Although productivity and safety could largely be improved with BIM integrated into temporary structures, there is currently a severe lack of BIM models for temporary construction works. The lack of BIM for scaffolding means that there is insufficient information for analysing the performance of scaffolding projects.

Despite previous work on scaffolding planning and design, little of the research currently focuses on scaffolding monitoring during the erection and dismantling phases, where components are large in quantity, small in size and scattered in distribution. With the large quantity and high complexity of scaffolding systems,

Internet of Thing (IoT)-based methods can be used. However, these require the deployment of many sensors and infrastructures, which at present incur large extra costs. The Cyber Physical System (CPS) allows bidirectional coordination between virtual representations and physical systems and it shows promise in monitoring construction progress. A CPS for scaffolding structure health monitoring has been designed by Xiao et al. (2016), though its application in progress monitoring has not yet been proposed. As for laser scanning, a large volume of point cloud data makes data processing a time-consuming task and currently there is no fully automated method of semantic interpretation of scaffolding from point clouds. Methods from image and video processing are possible choices, but it would be still challenging to extract scaffolding components from the disarray of on-site objects, particularly when the scaffolding has multiple layers. 3D multi-view reconstruction is more suitable for distinguishing scaffolding with multiple layers, since depth information is included to assist recognition. However, it takes a very long computation time to reconstruct a 3D model, and previous research shows that only 40% of limited types of scaffolding components, including tubes and toeboards, are successfully reconstructed to a level of good quality (Xu et al. 2015). This limits the accuracy and feasibility of this approach for progress monitoring. Xu et al. (2016) developed automatic methods of classifying scaffolding point cloud into tubes and toeboards, and achieved an overall accuracy of 63% in a real building site test. Jung (2014) developed an automated approach to detecting failures in concrete shoring, which is similar to scaffolding. The approach establishes, learns and predicts possible structure failures with a Hidden Markov Model (HMM) based on edge images from video surveillance. The results demonstrated that vision technology is promising in performing comprehensive site monitoring tasks.

Previous research on scaffolding mainly focused on planning and designing for safety and productivity purposes. There is little in existing literature which focuses on automated monitoring scaffolding or other temporary structures during construction or maintenance. In addition, in current industry practice, most clients rely on reports from scaffolding contractors to know the actual quantity and progress

of scaffolding construction projects, while scaffolding contractors rely on manual on-site inspections by experienced scaffolding supervisors. There is a lack of efficient and objective methods of monitoring the actual status of scaffolding construction which lets clients know as-used quantities, rather than merely as-reported or as-ordered ones. The lack of effective methods for monitoring scaffolding construction in the physical world limits the control of a scaffolding project, leading to high cost and low efficiency. In this research, a vision-based approach is developed to estimate the quantity of erected scaffolding, which can monitor the process of scaffolding erection or dismantling.

2.5 Point cloud segmentation, classification and recognition

Point cloud segmentation, classification and recognition is the basis for further processing tasks and applications, such as geometric modelling, 3D object recognition, and scene understanding. Much work has been carried out to investigate segmentation and classification methods for point clouds in various scenarios including land cover, road environments (Lehtomäki et al. 2016), indoor environments of buildings, and Mechanical/Electrical/Plumbing (MEP) structures (Dimitrov & Golparvar-Fard 2015). Existing segmentation methods are mainly categorised into model-based methods, region-growing methods and clustering methods.

Model-based segmentation methods segment point clouds by the fitting of basic geometric shapes, such as planes, spheres, cones, and cylinders. 3D Hough Transform and Random Sample Consensus (RANSAC) algorithms (Fischler & Bolles 1981) are the most commonly used model-fitting methods. Schnabel et al. (2007) developed a RANSAC-based algorithm to detect shape primitives from an unordered point cloud. Much subsequent work has been developed to improve the performance of RANSAC by extending the restriction of primitive shapes and enhancing robustness to noises or outliers (Poudel 2013). These methods are sensitive to point cloud characteristics such as density and noise. They are of high

computation complexity when applied in large-scale datasets and can only detect limited types of shapes.

Region-growing or surface-growing methods are widely implemented in point cloud segmentation. Research was conducted to improve the performance of region-growing methods by investigating seed selection, stopping criteria, features for different scenes and the desired level of scale. For instance, Rabbani et al. (2006) proposed a region-growing algorithm with local connectivity and smoothness constraints to segment a point cloud into a set of smooth surface patches. Pu and Vosselman (2009), used a planar surface-growing method to segment a building point cloud into different planar patches, which were then classified into semantic elements such as wall, door and window. Dimitrov and Goparvar-Fard (2015) proposed multi-scale features accounting for surface roughness and local curvature to address variability in point clouds of MEP structures. Region-growing methods are not considered robust due to dependency and sensitivity on seed selection. Moreover, the method suffers from over segmentation when applied to scenes with large variations.

Another group of commonly used segmentation algorithms is that of clustering analysis, such as individual and contextual clustering methods. These individual clustering methods operate on local features of individual points. Local features include simple features like normal, colour, curvature and shape. Statistical features are also applied including Fast Point Feature Histogram (FPFH), Spin Image, and 3D Speeded Up Robust Features (SURF). There have been many clustering methods applied in point cloud segmentation and classification, including Maximum Likelihood classifiers, Support Vector Machines, Random Forests, Bayesian Discriminant Classifiers and AdaBoost. The performance of those clustering methods depends heavily on input features and thus there has been much work focusing on the estimation of representative and invariant features. A set of 2D/3D features have been assessed on their relevance to semantic classification of point clouds (Weinmann et al. 2013). In addition, these classification methods are independent classifications of each 3D point according to its own feature vector.

Despite high efficiency, these methods suffer the drawback of noisy appearance of point clouds (Weinmann et al. 2015).

Contextual classification approaches not only consider the features of each individual point but also account for the contextual relationship between neighbouring points by using statistical models of context, which can directly enable semantic interpretation of point clouds. These approaches include the Markov Random Field, Conditional Random Fields (CRF), and spatial inference machines modelling. Rusu et al. (2009) applied CRF to segment a point cloud into different types of geometric surfaces such as planes and cylinders by considering the contextual information of geometric surfaces. Koppula et al. (2011) developed a semantic labelling method of 3D point clouds for indoor scenes by considering the geometrical context with Markov Random Field and a scene labelled into walls, grounds, ceilings, printers, keyboards and so on. Lu and Rasmussen (2012) improved the efficiency of semantic labelling of points with simplified Markov Random Models by predicting point-wise potential with random forest classifiers or support vector machines. Their method divides point clouds in urban environments into six classes including vegetation, wire, pole, ground, façade and vehicle. Rather than individual points, the Markov Random Field can also operate on segments of points, which can improve robustness to noise and create efficiency of the segmentation. Shapovalov et al. (2010) proposed a Markov Random Field Model for the classification of urban point segments from a spatial index into ground, building, car, tree and low vegetation, and results showed greater accuracy and higher efficiency. These contextual classification methods can consider contextual constraints implied in a scene, making it promising for semantic interpretation of point clouds from construction sites.

Although these handcrafted features have been shown to be effective in representing geometric and spatial attributes, it is impossible for them to achieve an optimal balance between discrimination and robustness. Deep learning algorithms that learn abstract representations of objects at high level are promising for semantic and high level segmentation and classification. Recent years have seen great achievements in

image recognition and classification with deep learning approaches (Zhang et al. 2016). For instance, a texture recognition with Fisher Vector pooling of a Convolutional Neural Network (CNN) filter bank achieves a high accuracy of 81% in cluttered scenes (Cimpoi et al. 2015). Although it would be promising to investigate applications of deep learning in the semantic interpretation of construction scenes, an extremely large volume of training data is necessary for the fit of deep learning models. There is currently no available dataset at construction scenes, which limits the application of deep learning approaches.

In the field of computer vision, 3D object recognition is the task of correctly identifying objects presented in point clouds and determining their positions including location and orientation (Bariya & Nishino 2010). The field has a large number of applications, such as remote machine control, automatic assembly, remote sensing and robotics. According to the features used for recognition, current 3D object recognition methods are usually categorised into global feature based methods and local feature based methods (Bayramoglu & Alatan 2015). Global feature based methods use computational representations of an entire 3D object (Funkhouser et al. 2003) and are commonly used for shape retrieval (Castellani et al. 2008). However, global features require a successful segmentation of objects before recognition. They are therefore not suitable for partial objects in cluttered environments, or complex scenes where it is difficult to produce good quality segmentation. Severe occlusions in LNG plants make the global feature based method an ineffective option.

In contrast to global features, local features based methods use a group of surface descriptors in local areas, which are matched between sub-parts or regions of an entire object. Local feature based methods are typically composed of three steps: 3D keypoint detection, local feature description estimation and feature matching. At the keypoint detection step, 3D points with high information contents are extracted. At the stage of the feature description estimation, a numerical representation is produced to describe information in neighbourhood areas of keypoints. At the feature-matching step, feature descriptors are matched between a model library and the query object to infer identity of the object. Some typical keypoint descriptors are

Binary Robust Appearance and Normal Descriptor (BRAND) (Do Nascimento et al. 2013), Scale Invariant Feature Transform (SIFT) (Scovanner et al. 2007), Persistent Feature Histogram (PFH) (Rusu et al. 2008) and 3D SURF (Knopp et al. 2010). Local feature based methods are more suitable to occlusions and noises than global features. For instance, Huang and You (2014) proposed an industrial object detection system by matching PFH features and this achieved better results in cluttered industrial scenes. The advantages of local features in 3D object recognition make them promising in the detection of industrial components in LNG plants.

3 PROPOSED FRAMEWORK OF MONITORING

In this chapter, we introduce the overall roles of our proposed monitoring approach applied to an existing construction workflow and present the detailed processes of as-built information modelling in LNG plants.

As defined in our research scope, this research mainly focuses on the data interpretation of the vision data so as to enrich the semantic information of as-built models. The produced as-built information can be used for construction progress tracking, which is investigated in this research. But the semantic-enriched as-built model can also be applied to other construction management tasks, such as quality control, material management, productivity analysis and structure monitoring. This makes the developed approach different from as-built information modelling for progress tracking purposes. For progress tracking cases with existing as-designed 3D models available, the segmentation and recognition tasks can be reduced to a much simpler problem. The main advantage of the proposed as-built information modelling process is that the developed approach does not require as-designed 3D models, which makes it applicable to existing facilities without as-designed models, as well as to on-site temporary structures such as scaffolding.

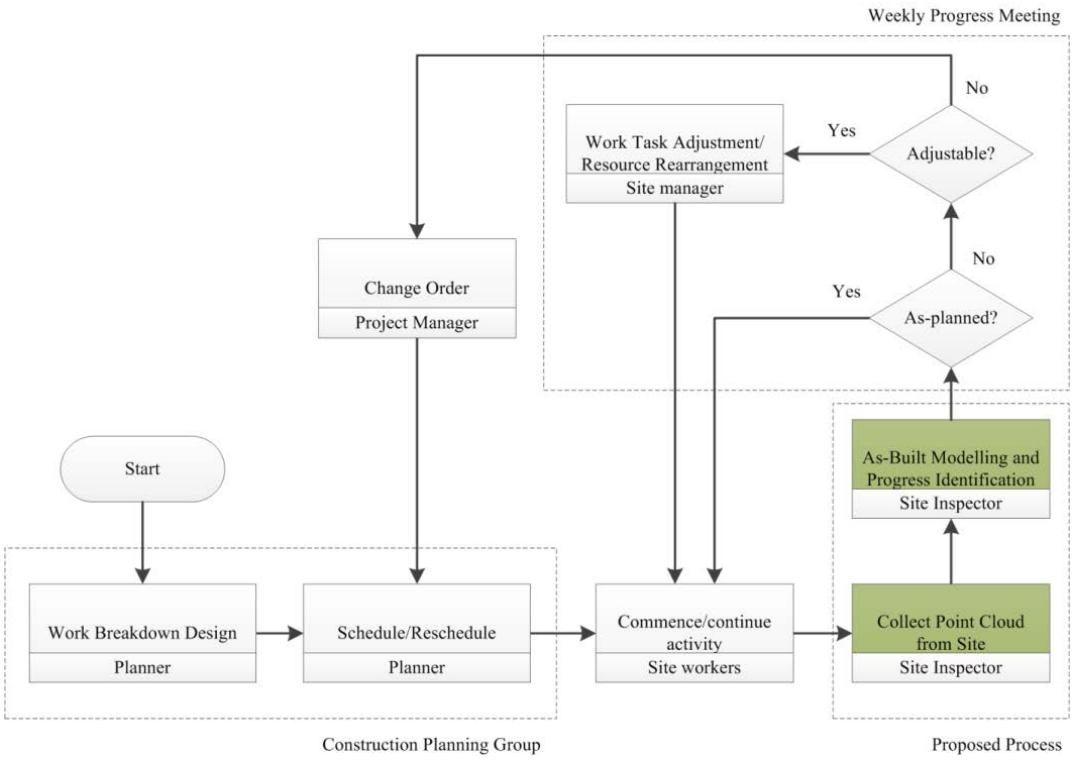

Figure 3.1. Roles of the proposed monitoring for progress tracking with vision-based data in the existing construction workflow

When used for progress tracking in the existing construction workflow, roles for construction monitoring are proposed, and a concurrent construction environment can be established by following the procedure shown in Figure 3.1. The conventional approach of the construction planning teams generates the schedule in terms of available resources (e.g. availability of contractors, heavy equipment and materials), the nature of construction sequences, owners' requirements and construction budget. Field construction teams then execute the plan. Once the field inspectors have conducted the proposed processes through collecting point clouds or images and identifying the current progress of the construction, detailed information regarding deviations with respect to the original plan can be clearly identified. The information and reports can be presented through visualisations during weekly progress meetings or monthly milestone meetings on-site. This guarantees that accurate results are generated and analysed in time. Site managers and project managers will then be able to determine whether the progress of the construction is on schedule. If there is

a work task that is behind schedule, they can identify whether it is adjustable at sites by rearranging current resources to speed up its progress, or a change in order could be raised to align with and rapidly satisfy real situations at sites. In the latter cases, planners can be notified in a relatively short time and they can start to reschedule the construction plan at an early stage, avoiding further time and costs.

This research focuses on both permanent and temporary structures in LNG plants to enable more complete status monitoring. For permanent structures in LNG plants, the as-built data is captured through laser scanning and is used to monitor and update their as-is conditions. For scaffolding as temporary structures, their as-built information is captured with cameras and their status is updated with daily or weekly photos. The as-built information modelling processes for both permanent structures and scaffolding are the key components for this research, which will be presented in detail in the following chapters. Figure 3.2. gives the overall strategy for information collection and processing.

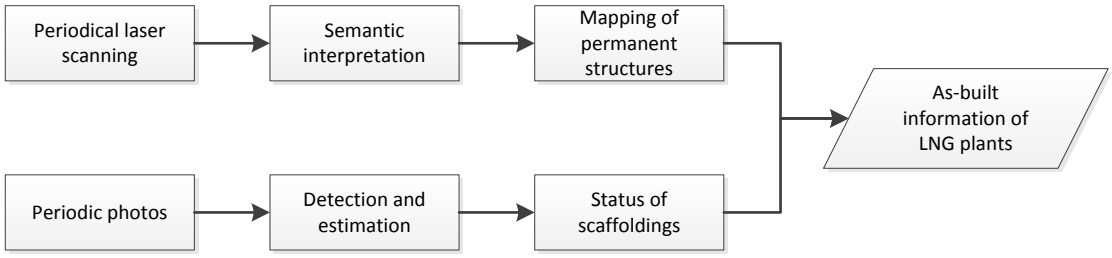

Figure 3.2. Overall strategy of data collection and processing

In the following chapters, detailed processes for each step in the overall strategy illustrated in Figure 3.2 will be presented respectively.

4 POINT CLOUD GENERATION IN INDUSTRIAL PLANTS

In this chapter, different methods of producing point clouds, including laser scanning and image-based reconstruction methods, are investigated. In Section 4.1, a general comparison between two methods is given. In Section 4.2 and 4.3, experiments are presented to test their performance in producing point clouds for LNG modules and scaffolding respectively.

4.1 Comparison between laser scanner and image-based reconstruction

There are two categories of approaches to the production of a 3D point cloud on a construction site, including laser scanning and image-based 3D reconstruction. Both laser scanners and image sensors can be deployed on various platforms, including cars, unmanned vehicles or aircraft, and they are of high applicability. Laser scanning can directly produce the point cloud measuring the distance between a laser scanner and an object surface. For large or complex sites, point clouds from multiple stations are aligned together to produce a complete point cloud of the whole scene. This can also help address occlusion problems, as a scanner requires line-of-sight conditions. A laser scanner is of high accuracy and high speed. Previous research

shows its measuring accuracy down to millimetres and it can collect millions of 3D points in minutes (Kršák et al. 2016). It is also able to directly produce 3D point clouds without much post-processing effort and thus requires little computation cost.

3D reconstruction from photos takes multi-view overlapping images as input. Point clouds are then generated through a series of processes: feature detection, matching, and 3D reconstruction. This can be achieved with conventional photogrammetry techniques with surveying cameras. In addition, this can also be achieved by 3D computer vision techniques, composed of self-calibration, SFM, and multi-view stereo (MVS) with consumable cameras. Due to rapid developments in recent years, 3D computer vision techniques make it possible to use photos captured by consumable (non-surveying-level) cameras or mobile phones to create 3D point clouds with off-the-shelf software. Depending on the cameras, software and photo capturing, the accuracy of the point cloud from photos with computer vision methods varies from centimetres to decimetres (Bhatla et al. 2012). Given the cameras used, the cost of computer vision methods is quite low compared to other methods. Although it is possible to produce a 3D model of high quality from photos with photogrammetry, surveying-level cameras and related equipment are required, and data collection and processing are subject to surveying regulations by highly trained personnel. The high cost of surveying-level equipment and the training required for the operator limits applications in construction projects. On-site photo collection is very fast and feasible, while the reconstruction process is of high computation cost and takes a long time, varying from hours to days, depending on the size of the scene.

The characteristics of different methods for producing point clouds are presented and compared in Table 4.1.

In this research, both laser scanning and non-surveying-level cameras are used to collect raw data on industrial facilities. Computer vision-based methods, SFM and related processing procedures in particular, are used to produce a point cloud model from photos captured with the cameras. The point clouds from laser scanning and photos are evaluated in the following sections.

Table 4.1. A comparison between 3D modelling with laser scanning and image-based reconstruction

Methods / Features	Laser scanning	Image-based reconstruction	
		Photogrammetry	SFM-MVS
Accuracy	High	Medium	Low
Speed	Fast	Slow	Slow
Cost	High	Medium	Low
Computation cost	Low	High	High

4.2 Performance evaluation in LNG components

Presented in this subsection are experiments on producing point clouds for permanent industrial structures with both laser scanners and cameras. As mentioned earlier, there are different methods of producing point cloud data, such as laser scanning, structure from motion techniques and stereo imaging. In a LNG plant, the main structures are permanent components, such as pipes and valves, which are essential parts of process plants. With regard to information from LNG components, there are special regulations subject to safety and production requirements, which limit available devices for data collection. Besides this, there are also many scaffolding components on-site, serving as temporary structures for construction or maintenance. Given the characteristics of different data collection methods and the complexity of industrial plants, it is necessary to investigate the performance of different data collection methods when they are applied in a LNG plant scenario. This research uses two kinds of components; pipes and valves, mentioned above as subjects, and it uses laser scanners and structures from motion techniques to produce point clouds of these components. The performance of the two methods is investigated by comparing their accuracy and completeness in producing a point cloud for an industrial module. Here, the completeness of the point cloud defines how complete the point cloud is with respect to the real object or scene it represents (Rukkanchanunt, 2014).

The experiments were carried out at a LNG process training site, where modules simulate the processes of a LNG plant and are used for training LNG workers. The site used is a typical operation site in the oil and gas industry and it was provided by the Australian Centre for Energy and Process Training (ACEPT), which runs just as a real process plant does, following all the operations and intrinsically safe regulations for training purposes. The site contains various LNG process modules, including metal components (steel frames, vessels, pumps, valves, pipes and compression tanks).

Figure. 4.1. Dehydration module for the laser scanning evaluation

On this site, a dehydration module representing a typical permanent structure of the site (see Figure 4.1) was scanned with a terrestrial laser scanner at multiple positions around it to reduce occlusion problems. The raw point clouds of multiple scanning were manually aligned together to produce data for the complete module. Figure 4.2 displays rough locations of different scanning spots, with five around the module and one on the top of the module.

The point cloud consists of more than 7 million points and is displayed in Figure 4.3. The point cloud with laser scanning is of high accuracy, down to millimetres, which

makes accurate geometric modelling possible. Line-of-sight conditions are necessary for successful scanning. For the outermost components in the module and components with few others around, laser scanning can produce a point cloud at a high completeness level, as high as 100% with multiple scanning.

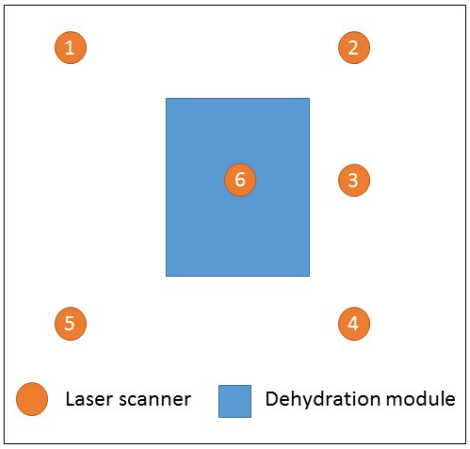

Figure 4.2. Locations of different scanning spots

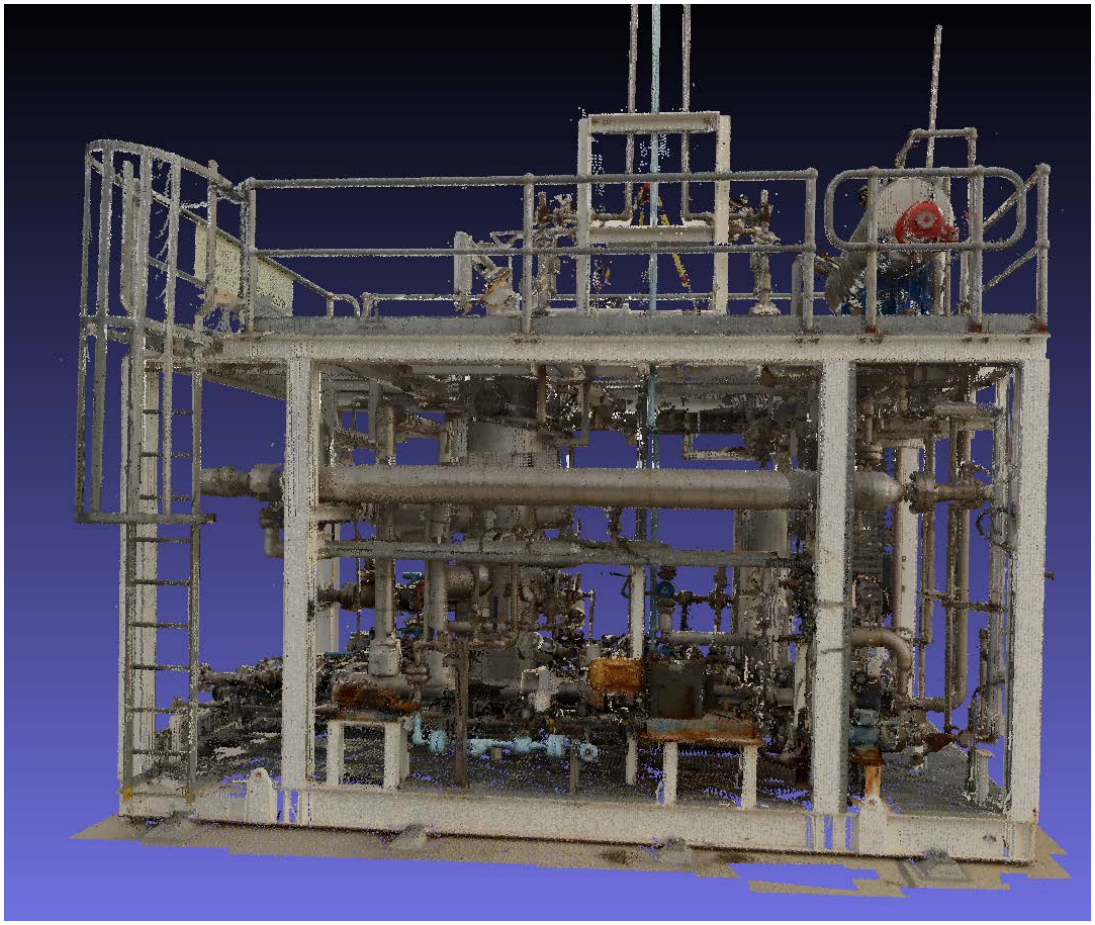

Figure 4.3. The point cloud of the dehydration module collected by a terrestrial laser scanner

The point cloud still suffers from incompleteness incurred by the occlusion among different components. Even though multiple scanning at different positions could reduce the impact of the occlusion, there are still many incomplete components in the point cloud. Figure 4.4 displays the point clouds of some components that are incomplete. As shown in the left figure, approximately half of the pipe facing the scanner is successfully scanned while other parts are missing. For the global control valve shown on the right, less than fifty percent of the whole valve remains. This is because this valve is inside the module surrounded by other components, and only limited positions for scanning can satisfy the line-of-sight requirement.

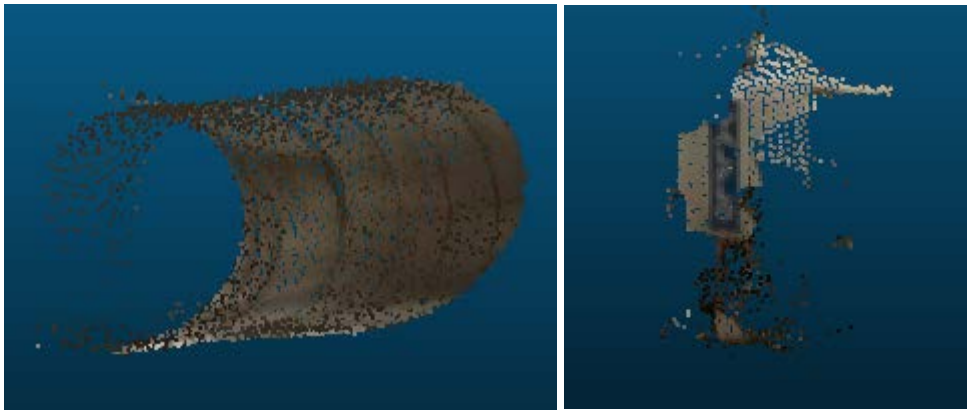

Figure 4.4. Incomplete components in the point cloud with laser scanning

In addition to the laser scanning, an attempt was made to produce a point cloud of the module with image-based reconstruction methods. A Digital Single Lens Refection (DSLR) camera was used to capture photos of the module at multiple viewpoints. The total number of photos was 389. Some of the photos are displayed in Figure 4.5. Agisoft PhotoScan software was then used to produce a point cloud, which is shown in Figure 4.6. The geometric configuration of all the photos is shown in Figure 4.7, where the locations and orientations of all photos relative to the module are displayed. The point cloud contains more than 35 million points. As can be seen from the figure, the SFM technique can successfully reconstruct overall structures of this module.

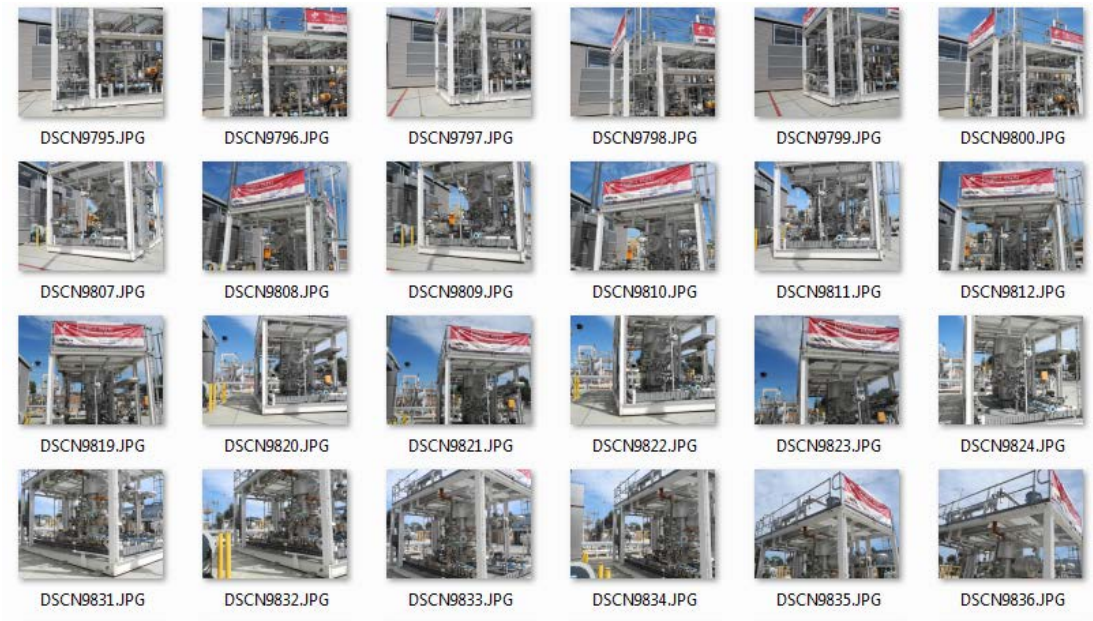

Figure 4.5. Photos used in the 3D reconstruction of the module

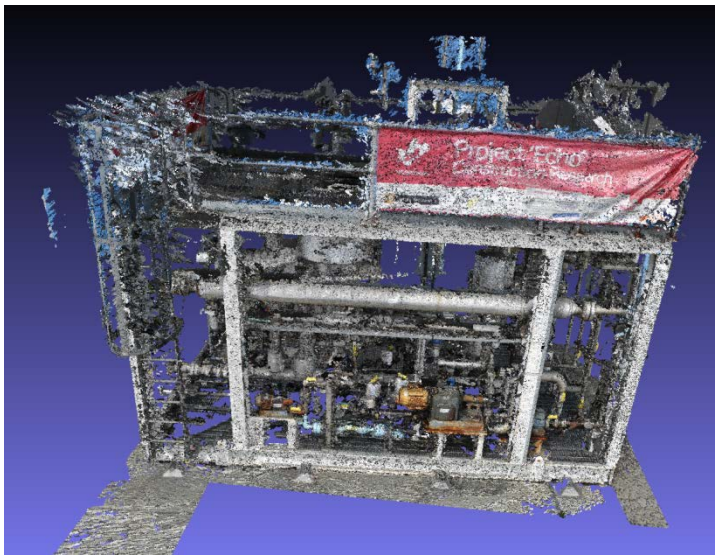

Figure 4.6. The point cloud produced from multi-view photos with Photoscan software

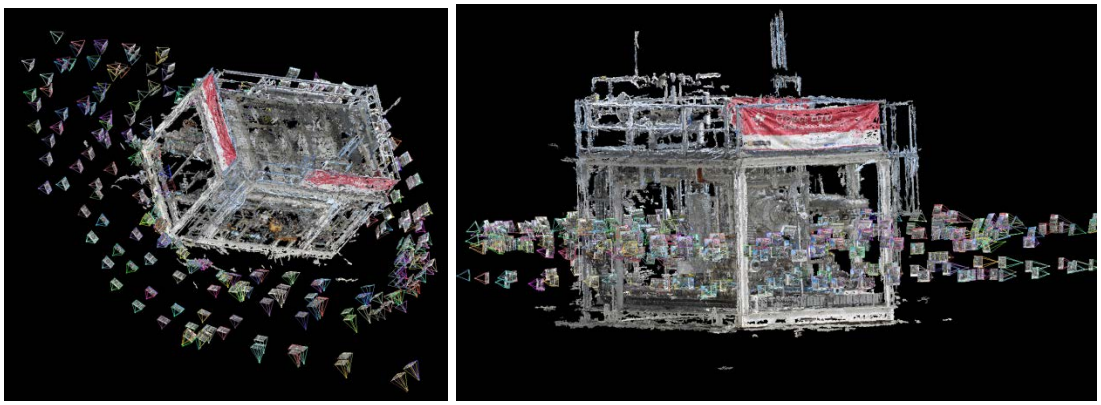

Figure 4.7. The configurations of all photos with positions and poses dispayed in the 3D space

There are more points in the point cloud than those produced from the laser scanner. However, the accuracy and completeness is not up to the standard of laser scanning, as can be seen in the Figure. Severe outliers exist in the point cloud and the accuracy of the point cloud is down to centimetres, making accurate modelling unfeasible. In addition, some parts of the module are missing or incomplete. This is also because of the severe occlusions caused by dense components in the module. Another possible reason is that most components are smooth steel with a similar appearance. The similarity and lack of texture in appearance could lead to mistakes in the matching, and failures in the reconstruction.

Displayed in Figure 4.8 are some components, whose point clouds produced by SFM are inaccurate or incomplete. For the manual valves in the first row, the overall surfaces are roughly reconstructed from photos, while details like flanges or bolts are missing. For the global control valve in the second row, the SFM method only produces a rough outline of the valve with details missing. For thin pipes and electrical components as shown in the third and fourth rows respectively, the SFM method fails to produce a correct point cloud model given most components are missing.

Figure 4.8. Incompleteness and inaccuracy in the point cloud produced by SFM: Each row represents a photo (left column), a point cloud produced by SFM (middle column) and a point cloud collected from laser scanning (right column)

The experiments show that it is challenging to create point clouds for LNG modules from photos. The problems are mainly severe occlusions in industrial plants and the similar appearance of man-made structures, which make the reconstructed point cloud incomplete and of low accuracy. It can still be useful for visualisation purposes.

However, there are many problems remaining for tasks requiring data with a high level of completeness and accuracy, such as as-built modelling.

Given the poor quality of the point cloud from photos in this case, the processing of point cloud segmentation and recognition presented in the next chapter (Chapter 5) is conducted on laser scanned data only.

4.3 Performance evaluations in scaffolding structures

In this subsection, we conduct a performance evaluation of producing a point cloud for scaffolding in a LNG plant with the SFM method and a photo collection process. Due to the temporal, flexible and dynamic natures of scaffolding tasks, it is time-consuming for laser scanning and unlikely to produce a massive amount of scanned data for the entire scaffolding structures particularly for temporary uses. In addition, scaffolding is mainly composed of thin tubes with similar appearance and it is considered to be difficult to build a point cloud from scanning. Besides, as a result of intrinsic safety requirement of a LNG plant, all electrical devices including laser scanners and cameras should be intrinsically safe, if those devices are used while the plant is in operation. For devices that are not intrinsically safe, there needs to be shutdown of the plant, which is impractical due to the impact on the productivity and cost. This requirement significantly limits the choice of affordable and effective data collection tools used on-site. Specifically, laser scanning is not a choice for monitoring scaffolding during plant operation.

Previous study and research shows that it is possible to reconstruct a scaffolding structure with high resolution cameras (Chaiet al. 2016). However, current available cameras meeting intrinsic safety requirements (named intrinsically safe cameras) suffer from high distortion, low resolution and limited focal length, which decreases the quality of the photos and creates a challenge for successful reconstruction. Therefore, it is meaningful to test the feasibility and performance of 3D reconstruction with intrinsically safe cameras when applied to scaffolding. To do this, this research conducted pilot experiments on reconstructing scaffolding with

intrinsically safe cameras, and a commercial DSLR camera, and it compares their reconstruction results.

As a commercial DSLR camera does not comply with safety regulations and thus cannot be tested in an oil and gas plant in operation, the scaffolding around a building was chosen as a subject different to that of a plant. It is meaningful since the aim is to test the performance of scaffolding reconstruction with photos. There is little difference between scaffolding structures used in a building and those in an industrial plant.

Presented in the following is the experimental performance of a commercial DSLR camera, and an intrinsically safe camera, on reconstructing scaffolding point clouds. The reconstruction took two groups of photos as inputs. One group included 83 photos and was captured with a commercial DSLR camera, while the other included 88 photos with an intrinsically safe camera. The DSLR camera was a Canon EOS 700D with a focal length of 39mm. The resolution of captured photos was 5184*3456. The intrinsically safe camera was a Pixavi Gravity X with an auto focus lens. The photos had a pixel resolution of 2448*3264 and were taken at an infinity focal length of 4.6 mm.

This study used Agisoft PhotoScan software to reconstruct the point clouds. Displayed in Figures 4.9 and 4.10 are the reconstructed models of the scaffolding structure with photos from two cameras respectively. Figure 4.9 shows that the whole scaffolding along a building façade is successfully reconstructed with individual tubes of the scaffolding clearly modelled. In contrast, the scaffolding modelled from the intrinsically safe camera is of very low accuracy, with overall reconstruction failing and details missing. As can be seen in Figure 4.10, large parts of the scaffolding tubes are missing and geometrical relationships are mainly incorrect.

Figure 4.9. Reconstructed scaffolding with photos from a commercial DSLR camera

Figure 4.10. Scaffolding reconstructed from photos with an intrinsically safe camera

The main findings of the experiment include: 1) it is feasible to reconstruct point cloud models of scaffolding from photos, should high quality cameras be available; the completeness of the reconstructed model makes some monitoring tasks possible; 2) the intrinsically safe camera fails to reconstruct point cloud model of scaffolding, making it infeasible to use point cloud to monitor scaffolding. For intrinsically safe cameras, completeness and accuracy are too low to perform monitoring tasks.

Based on the observations and experimental results conducted here, it is difficult for intrinsically safe cameras to produce an accurate and complete point cloud from

images of a LNG plant. This research developed a hybrid approach which included directly analysing those raw images and estimating the quantity of scaffolding structures. This is described in detail in Chapter 6.

4.4 Summary

In this chapter, we introduced two categories of methods of capturing 3D as-built models of industrial facilities, including laser scanning and image-based reconstruction. The two methods can produce a 3D point cloud model of a construction site. We tested and compared the performance of producing 3D as-built models of LNG components in terms of their accuracy, speed and cost. The results show that compared to laser scanning, image-based reconstruction suffers from more severe occlusion problems in LNG facilities and it does not effectively produce a complete point cloud model. Textureless appearances of process plants make them difficult to detect and match with feature points across images. Laser scanning is more applicable in dense environments such as LNG plants. To effectively capture the as-built information of scaffolding, we tested image-based methods to reconstruct a 3D model of scaffolding. Both high quality DSLR cameras and off-the-shelf intrinsically safe cameras were tested in different scenarios, including building sites and LNG facilities. The results show that it is possible to create a point cloud model of scaffolding with high quality DSLR cameras. However, intrinsically safe cameras required by LNG plants failed to produce complete and accurate 3D point clouds.

5 SEMANTIC INTERPRETATION OF INDUSTRIAL POINT CLOUDS

In this chapter, we propose a method for a semantic scene understanding of industrial point clouds by addressing critical problems including semantic segmentation and irregular object recognition. The proposed method hierarchically interprets a 3D scene by firstly segmenting the point cloud into different categories, and then recognising objects within an interested category. The proposed segmentation method accounts for semantic constraints and prior knowledge of the scene based on recent high-level image processing techniques developed in the computer vision field.

In Section 5.1, we describe a developed method for semantic scene parsing of industrial point clouds. The developed approach divides an entire industrial scene into different categories and is mainly based on Markov Random Field, considering the contextual information in industrial plants. This approach also uses supervoxelisation to improve robustness and computational efficiency. In Section 5.2, a developed method for recognising irregular objects from point clouds was described. The recognition method uses a local-feature based matching strategy and enables an object-level scene understanding. In Section 5.3, we present experiments

validating the proposed methods, where experimental data is detailed and quantitative results are presented.

5.1 Semantic segmentation via Supervoxel and Markov Random Field

Automatic segmentation and object recognition of point clouds is necessary to achieve automated monitoring with laser scanning. Without the recognition of an object, it is not possible to track or monitor the status of the scene at sites. For a conventional manual interpretation, it is easy to understand point clouds interactively, despite it being a time-consuming task with significant labour costs. However, there are still challenges for a computer system in fully interpreting point clouds at a high level. In this research, a graph-based approach was implemented to automate segment point clouds of an industrial plant into semantic objects. The graph-based approach is a semantic segmentation method that merges low-level context with object-level class knowledge. This makes it more suitable for high-level tasks, e.g., object recognition. Xiong and Huber (2010) used a similar strategy to create semantically rich 3D building models from laser scanner. However, their approach works on planar patches and are designed to extract only planar components such ceilings, walls, and floors (Xiong et al. 2013), making it unsuitable to industrial plants.

The process of the proposed semantic segmentation approach is displayed in Figure 5.1. The raw point cloud is over-segmented into patches named supervoxel, using a supervoxelisation method. A feature vector is estimated for each supervoxel. The probabilities of each supervoxel belonging to a class are then predicted with a Support Vector Machine (SVM) classification, which is trained by the manually labelled data. With each supervoxel as a node and their connectivity, a Markov Random Field is constructed, where the data cost of a node is its classification probability, and smooth energy is estimated according to its feature vector. The point cloud is then semantically segmented by optimising the energy of the Markov Random Field.

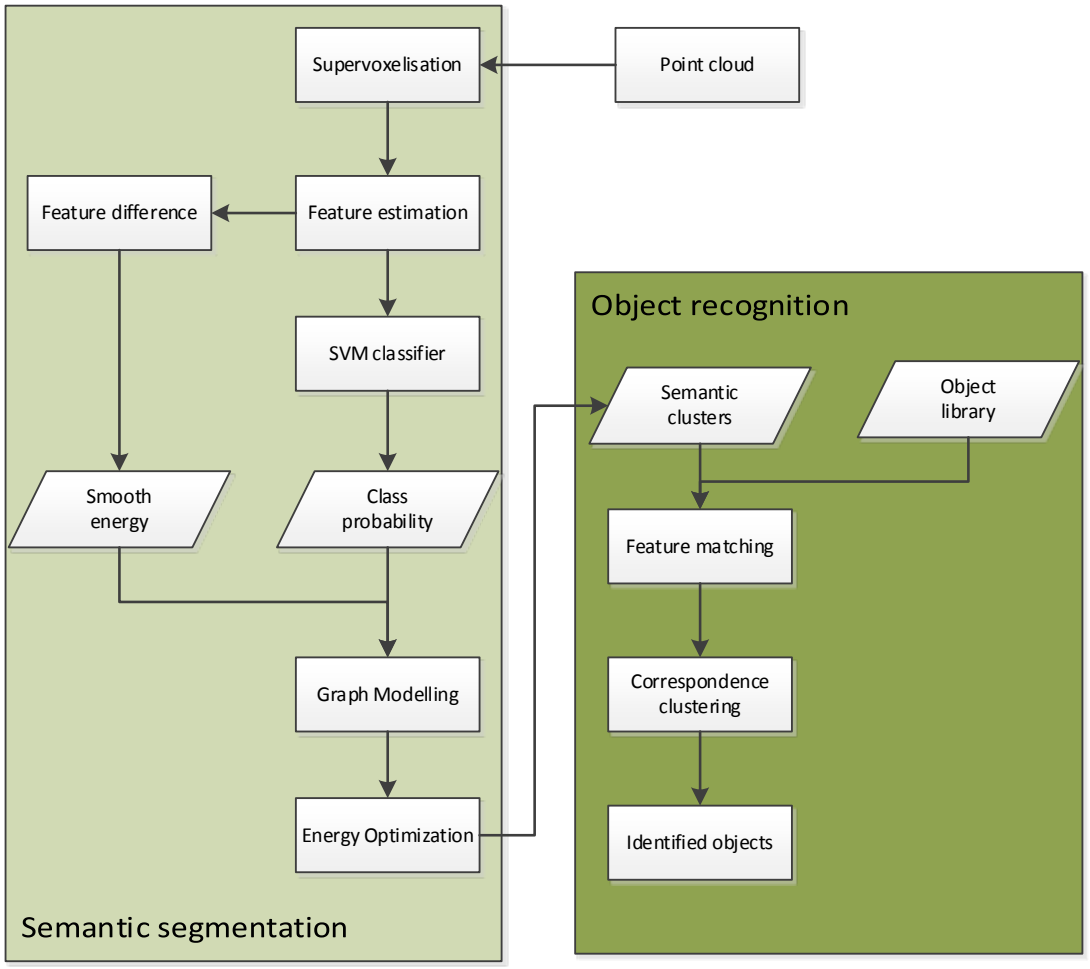

Figure 5.1. A process of semantic interpretation of an industrial point cloud

5.1.1 Supervoxelisation

Many previous works on graph-cut segmentation have been point-based, which means that the point cloud is processed according to each single point and its features. Although it is easy to estimate features of a point, point-based classification approaches suffer from high computational cost and low accuracy. A point cloud is usually greater than millions of points, depending on its density and size. For example, the collected point cloud of a LNG module is composed of more than 7 million points, while a LNG plant contains thousands of modules. The large volume of data makes direct approaches high in computational cost. Furthermore, features of a point are estimated according to its local neighbourhood, which makes it easily affected by noise and thus reduces the accuracy of the classification.

The cost of point-level processing leads to the development of mid-level inference schemes, which do not operate on points directly but rather use groups of points, namely supervoxel. A Voxel Cloud Connectivity Segmentation method (VCCG) was developed by Papon et al. (2013) to over-segment point clouds into supervoxels. In our research, we applied this approach due to its ability to conform to object boundaries during the segmentation.

Here, we briefly introduce the VCCG approach used in the developed process. The approach consists of four basic steps including adjacency graph, spatial seeding, features and distance measure, and flow constrained clustering. The approach firstly builds a 26-adjacent voxel graph by searching a kd-tree (Rusu & Cousins 2011) of the point cloud and then produces some seed points based on the neighbour size and gradient of a voxel. A 39-dimensional feature for each voxel is estimated which includes spatial coordinates, colour in CIELab space (Ganesan et al. 2010) and Fast Point Feature Histograms (FPFH). The FPFH are pose-invariant features describing local surface properties of points. All voxels are assigned to supervoxels iteratively using a local k-mean clustering (R. B. Rusu et al. 2009).

Table 5.1. The processed results of the supervoxelisations on two point clouds

Item	**Points**	**Supervoxels**
Point cloud 1	1,041,893	7,614
Point cloud 2	247,136	1,303

To illustrate the effect of the point cloud supervoxelisation, an example is given of point clouds of two scenes. The results of the supervoxelisation are displayed in Figure 5.2, where the left images show raw point clouds of industrial components and the right ones show those after the supervoxelisation. The first row in Figure 5.2 represents the point cloud for the first scene and the second row represents the point cloud of the second scene. Each supervoxel is coded with different colours. In Table 5.1, the numbers of points and supervoxels for two point clouds are presented. It shows that 7,614 supervoxels are produced from 1,041,893 points for the first scene and 1,303 supervoxels are produced from 247,136 points for the second scene. As

numbers of supervoxels are much less than numbers of points in corresponding point clouds, this pre-processing effort largely reduces the computational complexity of the point cloud segmentation.

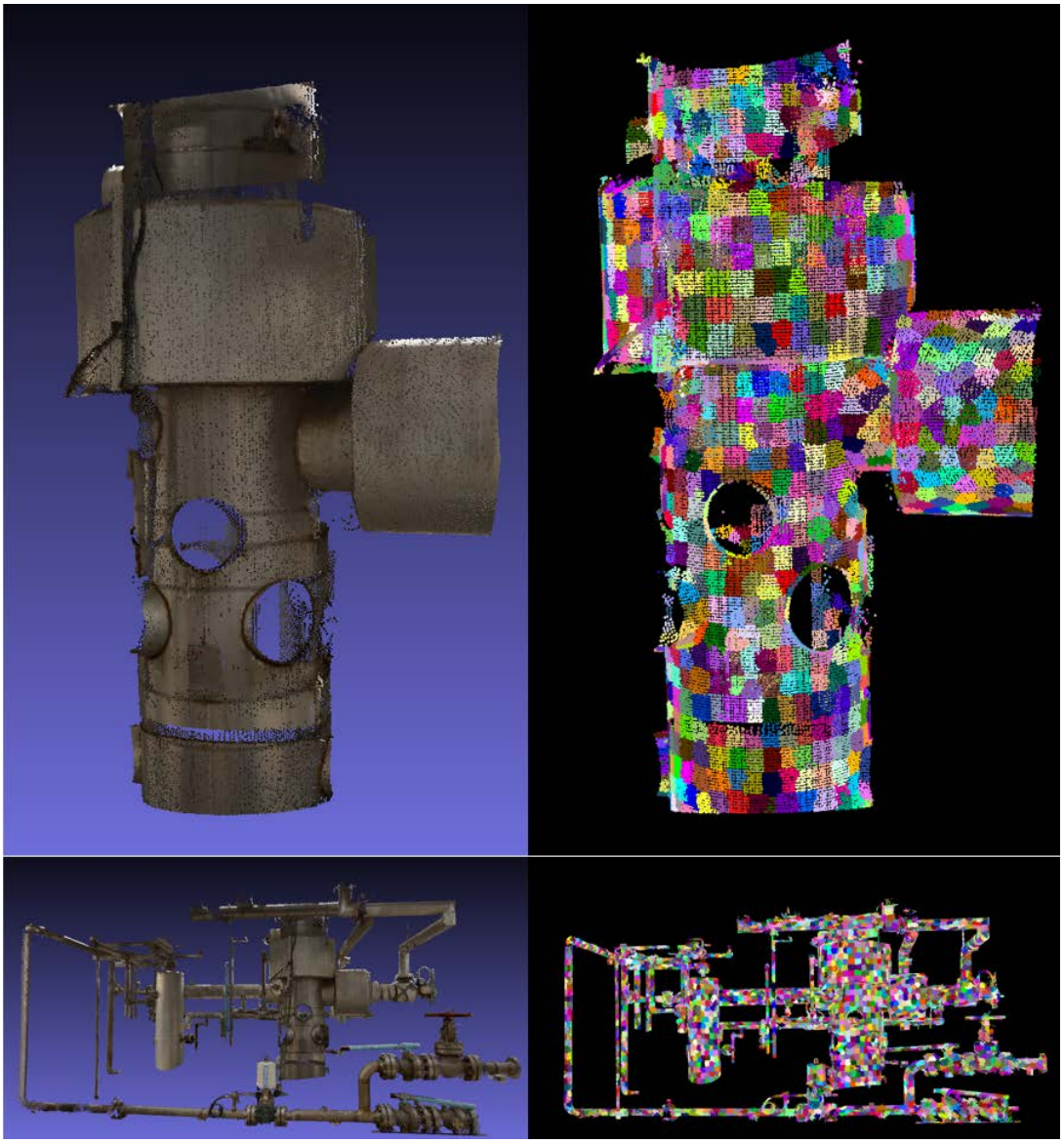

Figure 5.2. Supervoxelisations on point clouds of industrial components

5.1.2 Features of supervoxel

The feature vector of a supervoxel, which describes its characteristics, is the key to differentiating supervoxels belonging to different objects. Here, we produce a feature vector of a supervoxel by combining different kinds of basic features, including curvature, normal, colour, and shape features. All the features are estimated according to points belonging to their supervoxels.

The colour of a point $\boldsymbol{I_j} = [\boldsymbol{R_j} \quad \boldsymbol{G_j} \quad \boldsymbol{B_j}]^{\boldsymbol{T}}$, in RGB colour space, can be directly obtained with a laser scanner collecting illumination information as well. The average of colours of all points in a supervoxel is used as the colour for the supervoxel. This colour feature can contribute to an easy classification of objects with different colours.

The normal and principal curvatures of a supervoxel are estimated based on the eigenvectors of the covariance matrix of points in the supervoxel (Mitra et al. 2004). For a supervoxel $\boldsymbol{y_i}$, let $\{\boldsymbol{x_j \epsilon y_i} | \boldsymbol{j = 1, 2, \cdots, k}\}$ denote points in this supervoxel. The covariance matrix of the supervoxel can be calculated as follows:

$$\boldsymbol{C_i} = \frac{\boldsymbol{1}}{\boldsymbol{k}} \sum_{\boldsymbol{x_j \epsilon y_i}} (\boldsymbol{x_j} - \overline{\boldsymbol{x_c}})(\boldsymbol{x_j} - \overline{\boldsymbol{x_c}})^{\boldsymbol{T}} \tag{5.1}$$

where $\overline{\boldsymbol{x_c}}$ denotes the centroid of the supervoxel and $\overline{\boldsymbol{x_c}} = \frac{\boldsymbol{1}}{\boldsymbol{k}} \sum_{\boldsymbol{x_j \epsilon y_i}} \boldsymbol{x_j}$.

Let $\boldsymbol{\lambda_i}$ $(\boldsymbol{i = 1, 2, 3})$ denote the eigenvalues of the covariance matrix. Suppose that $\boldsymbol{\lambda_1 > \lambda_2 > \lambda_3}$ and $\boldsymbol{v_i}$ $(\boldsymbol{i = 1, 2, 3})$ denote the corresponding eigenvectors.

The normal vector $\boldsymbol{N_j}$ of the supervoxel is calculated as the eigenvector of the covariance matrix corresponding to the smallest eigenvalues, and the principal curvature $\boldsymbol{P_j}$ is calculated as the eigenvector of the covariance matrix corresponding to the largest eigenvalue. The curvature feature of points lying on different kinds of surfaces, such as a plane and a cylinder, are different.

In addition, shape features are also included, which describe the scatter, planarity and linearity characteristics of a supervoxel (Weinmann et al. 2015). Shape features

are also calculated by analysing eigenvalues of the covariance matrix of a supervoxel. Specifically, shape features can be estimated as follows:

$$\begin{cases} \boldsymbol{s_1} = \frac{\lambda_3}{\Lambda} \\ \boldsymbol{s_2} = \frac{\lambda_2 - \lambda_3}{\Lambda} \\ \boldsymbol{s_3} = \frac{\lambda_1 - \lambda_2}{\Lambda} \end{cases} \quad 5.2$$

where $\boldsymbol{\Lambda = \lambda_1 + \lambda_2 + \lambda_3}$, and $\boldsymbol{s_1}$, $\boldsymbol{s_2}$ and $\boldsymbol{s_3}$ are scatter, planarity and linearity features respectively.

Then, we use $\boldsymbol{S_j = [s_{j1} \quad s_{j2} \quad s_{j3}]^T}$ to denote the shape features of a supervoxel.

The feature difference between two supervoxels is estimated as the sum of weighted distances of different features as in Equation 5.3:

$$\boldsymbol{\delta_{Fij} = \gamma_1 \|I_i - I_j\| + \gamma_2 \|C_i - C_j\| + \gamma_3 \|P_i - P_j\| + \gamma_4 \|S_i - S_j\|} \quad 5.3$$

where $\boldsymbol{\gamma_1, \gamma_2, \gamma_3,}$ and $\boldsymbol{\gamma_4}$ denote the weights of different features. The detailed information of the feature vector used here can be seen in Table 5.2.

Table 5.2. Features of each supervoxel

Feature	**Label**
Normal	$\boldsymbol{N_j}$
Principal curvature	$\boldsymbol{P_j}$
Shape feature	$\boldsymbol{S_j = [s_{j1} \quad s_{j2} \quad s_{j3}]^T}$
Colour	$\boldsymbol{I_j}$
Total dimension of supervoxel features	12

5.1.3 Sematic segmentation with Markov Random Field

To consider contextual information for interpreting point cloud data, a graph cut approach based on supervoxels for semantic segmentation of the point cloud of an LNG plant is implemented in this subsection. The segmentation is defined as a supervoxel labelling problem. This is based on the observation that point cloud

regions are always homogenous and neighbouring pixels have similar properties. Markov Random Field (MRF) is a probabilistic model that captures such contextual constraints.

A point cloud of supervoxels is denoted as $\boldsymbol{Y} = \{\boldsymbol{y_i} \in \mathbb{R}^{\boldsymbol{B}}, \boldsymbol{i} = \boldsymbol{1}, \boldsymbol{2}, \cdots, \boldsymbol{m}\}$, where *B* is the dimension of the feature vector of a supervoxel, and *m* is the number of supervoxels. The point cloud consists of some predefined classes $\boldsymbol{L} = \{\boldsymbol{l_i} \in \boldsymbol{R}, \boldsymbol{i} = \boldsymbol{1}, \boldsymbol{2}, \cdots, \boldsymbol{n}\}$, where *n* is the number of classes and $\boldsymbol{l_i}$ is the label of each class. A supervoxel belonging to a region is specified by a label $\boldsymbol{w_i}$. The segmentation becomes a labelling process to produce an optimal label set $\boldsymbol{W} = \{\boldsymbol{w_i}, \boldsymbol{i} = \boldsymbol{1}, \boldsymbol{2}, \cdots, \boldsymbol{m}\}$. Optimal labelling can be obtained by maximising the posterior probability (Kato & Pong 2006), which is represented as Equation 5.4:

$$\widehat{\boldsymbol{w}} = \mathbf{arg}\ \max_{\boldsymbol{w} \in \boldsymbol{\Omega}} \boldsymbol{P}(\boldsymbol{w}|\boldsymbol{Y}) \qquad 5.4$$

where $\boldsymbol{\Omega}$ represents all possible labelling.

Maximising the posterior probability is equal to minimising the energy function of the model:

$$\widehat{\boldsymbol{w}} = \mathbf{arg}\ \min_{\boldsymbol{w} \in \boldsymbol{\Omega}} \boldsymbol{E}(\boldsymbol{Y}, \boldsymbol{L}, \boldsymbol{W}) \qquad 5.5$$

The energy function of the Markov model can be written as a sum of unary and pairwise terms as in Equation 5.6:

$$\boldsymbol{E}(\boldsymbol{W}) = \sum_{\boldsymbol{i}} \boldsymbol{\Phi_i}(\boldsymbol{w_i}) + \boldsymbol{\gamma} \sum_{\boldsymbol{ij} \in \boldsymbol{N}} \boldsymbol{\Psi_{ij}}(\boldsymbol{w_i}, \boldsymbol{w_j}) \qquad 5.6$$

where $\boldsymbol{\Phi_i}$ is the unary energy of a supervoxel, $\boldsymbol{\Psi_{ij}}$ is the pairwise energy of two adjacent supervoxels, $\boldsymbol{\gamma}$ is a regularisation parameter balancing the unary energy and pairwise energy, and $\boldsymbol{N}$ denotes the set of all pairs of adjacent supervoxels in a supervoxelised point cloud. The labelling process is now modelled as a MRF with a first-order neighbourhood system.

In the proposed approach, the unary term of the energy function is estimated according to the probability of a supervoxel belonging to a class. The probability is predicted with SVM. The unary term of the energy is proportional to its probability.

The unary energy for each supervoxel is defined as the negative log-probability for each class from the SVM classifier:

$$\boldsymbol{\Phi_i(w_i) = -\log \mathbf{P}(w_i|y_i)} \tag{5.7}$$

where $\mathbf{P}(\boldsymbol{w_i}|\boldsymbol{y_i})$ denotes the probability of a supervoxel $\boldsymbol{y_i}$ belonging to a class $\boldsymbol{w_i}$.

The pairwise term is estimated according to the feature vector difference between neighbouring supervoxels. The pairwise term indicates the energy incurred by neighbouring supervoxels belonging to different object classes. The pairwise term imposes a contrast-sensitive smoothness prior (Schindler 2012), which means the model prefers labelling where adjacent supervoxels have the same label. The pairwise term is defined as follows:

$$\boldsymbol{\Psi_{ij} = \delta_{ij}/d_{ij}} \tag{5.8}$$

where $\boldsymbol{d_{ij}}$ represents the Euclidean distance of two adjacent supervoxels and $\boldsymbol{\delta_{ij}}$ is a function describing label differences between supervoxels:

$$\boldsymbol{\delta_{ij}} = \begin{cases} \mathbf{exp}\left\{-\frac{\|F_i - F_j\|^2}{2\beta}\right\} \emptyset(i,j) & w_i \neq w_j \\ \mathbf{0} & w_i = w_j \end{cases} \tag{5.9}$$

where $\boldsymbol{F_j}$ denotes the feature vector of a supervoxel and $\boldsymbol{\beta}$ is the mean square difference of feature vectors of all supervoxels. $\emptyset(\boldsymbol{i,j})$ represents a category-cross energy weight for two neighbouring supervoxels belonging to different categories.

The pairwise contrast-sensitive energy is based on the assumption that a point cloud is usually smooth with label changing occurring at the boundary between different areas of features. In addition, the pairwise energy can constrain the neighbouring relationship between different classes by imposing different energy costs. In current

practice, the possibility of the neighbouring relationship between two individual classes of objects is not uniform and generally limited. For example, in a LNG plant with surrounding environment, all industrial components in the plant are enclosed with fences or walls. Therefore, there is no possibility that any industrial component is adjacent to trees or other vegetation. In this case, a high weight $\emptyset(\boldsymbol{i},\boldsymbol{j})$ for the energy cost can be imposed on the neighbouring relationship between the two classes.

Table 5.3. $\emptyset(\boldsymbol{i},\boldsymbol{j})$ values for different neighbouring categories in LNG plants

Category	**Ground**	**Trees**	**Fence**	**Industrial**	**Building**
Ground	0	1	1	1	1
Trees	1	0	2	2	2
Fence	1	2	0	2	2
Industrial	1	2	2	0	1
Building	1	2	2	1	0

In Table 5.3, we give an example of the category-cross energy weight for common objects in a LNG plant. As can be seen in the table, the weight between the same categories is all set to zero. This is in accordance with the previous smooth assumption. As all other objects usually stand on the ground, making it reasonable to be neighbouring to a ground supervoxel, so the weights between the ground and all other labels are set to 1. Conversely, trees and industrial components are not supposed to be adjacent to each other, so their weights are set to 2, which is higher than a reasonable relationship. Besides this, as the weight is only determined by labels of supervoxels, the weight is symmetrical and unidirectional.

To predict the unary term of a supervoxel, SVM was applied, as developed by Chang et al. (2011), to estimate the probability of a supervoxel belonging to a class. SVM is a popular machine learning method for classification problems. It takes a training dataset as input and produces a classifier, which is used to classify raw point clouds. In this research, the dataset is produced by establishing a point cloud library of different classes of industrial components. Depending on the scene, these classes

include ground, trees, fences, industrial modules and so on. Firstly, all components in the training library are supervoxelised with the VCGG algorithm. Then, the features of each supervoxel are estimated as described. A multi-class SVM classifier is trained based on those features, which is then used to predict the test point cloud data.

The regularisation term $\boldsymbol{\gamma}$ has a great influence on the performance of this approach. In this research, the value of $\boldsymbol{\gamma}$ is decided by experiments. The above energy minimisation problem can be solved by graph-cut algorithms. This research utilises the graph-cut algorithm proposed by Delong et al. (2012) to produce an optimal solution for the MRF.

5.2 Irregular object recognition with local feature matching

The segmentation method in the previous subsection aims to divide a whole scene into different categories. In this subsection, a 3D object recognition method attempts to recognise components in each category to understand a scene with a higher level of detail.

Current methods of recognising point cloud objects for as-built modelling mainly consist of two categories. One is the "Scan vs BIM" strategy, which identifies objects by comparing point clouds with their aligned as-designed models. This kind of method requires as-designed 3D models, which limits applications for existing facilities without BIMs. The other kind is based on feature matching and clustering, commonly used in the computer vision field.

LNG component recognition is used to find industrial components in each cluster of points and determine its pose in a coordinate reference system. To recognise the type of component, this research developed a library-based recognition approach, which recognises a point cloud component by matching it with a pre-defined component library.

In an industrial plant, most components are manufactured according to certain standards and their types can be identified by distinguishing features determined

from prior knowledge. Given this, a point cloud library of industrial components is created, where there are corresponding point-cloud models for every type of industrial component. Some components for industrial plants are shown in Figure 5.3, including manual valve, global control valve, meter and electromotor. Their related point clouds in the library of industrial components can be seen in Figure 5.4.

Figure 5.3. Components in industrial plants

Figure 5.4. Components in the library for recognition purposes

Because of the complexity of industrial plants, there is a great deal of occlusion and noise in point clouds, and it is possible that errors could occur during the segmentation process. Previous research has shown that local-feature-based recognition methods are more robust against occlusions and clutters in real world scenarios compared to global features (Guo et al. 2014). This research therefore implements a local-feature-based 3D object recognition approach to matching segmented laser-scanned data with industrial components in the pre-defined library.

The local-feature-based approach recognises an object by matching and clustering features in a scene against the model features. The performance of the feature-based recognition depends heavily on the local feature used in the algorithm, which should be distinguishable enough among different kinds of objects but repeatable within the

same one. This research uses 3D Scale-Invariant Feature Transform (SIFT) features (Scovanner et al. 2007) for the recognition. The 3D SIFT key points are detected as local extremes in the scale space using the difference-of-Gaussians pyramid technique. Histograms of gradient orientations in a neighbouring patch of key points are used as feature descriptions, as they are invariant to scale and rotation and are able to withstand an occlusion and noisy data. 3D SIFT key points and feature descriptors are estimated for both the industrial components library and the target point cloud of the scene. Matching according to 3D SIFT features produces point correspondence between the scene and models in the library. The matching process employs the ratio of nearest-neighbour distances.

The Hough voting method is used to cluster feature correspondence and generate hypothesis candidates in Hough space (Guo et al. 2014). Each point in Hough space represents a transformation from the library to the scene. All possible transformations estimated according to feature correspondences and peaks in the Hough parameter space (rotation and translation) are chosen as candidates. To remove false candidates and improve the accuracy of the transformation estimation, candidates for the transformation between a library component and a point cloud of the scene are verified and refined using an Iterative Closest Point (IPC) algorithm (Besl and McKay 1992). Examples of the matching process for the industrial component recognition can be seen in Figure 5.5, where red points represent point cloud objects, yellow points represent query objects and green lines show matched point correspondence.

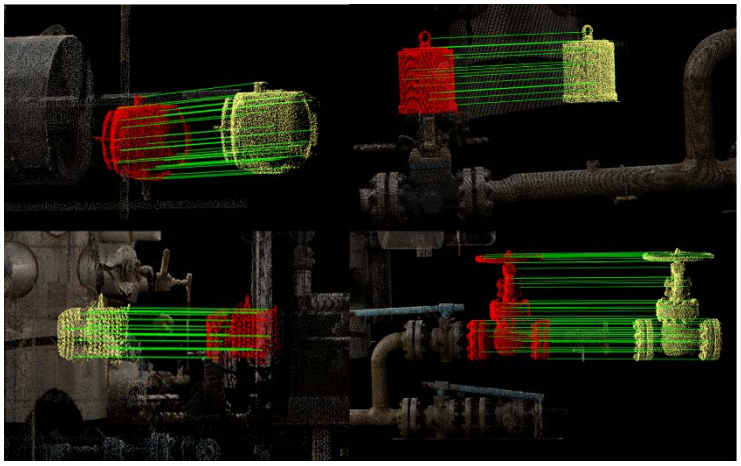

Figure 5.5. A matching process of the component recognition

5.3 Experiments of point cloud interpretation approaches

In this subsection, experiments on evaluating the proposed point cloud interpretation algorithms are conducted and described. The proposed approaches are implemented using VC^{++} platform and Point Cloud Library.

5.3.1 Point cloud data

The data used in these experiments includes the point cloud of a LNG process training site. It is shown in Figure 5.6. The point cloud was scanned using a Leica Scanstation at five scanning spots and it consisted of more than 1.2 million points. In addition, the performance of the proposed approach was tested with the point cloud of a plant pump ducting room. As can be seen in Figure 5.7, the point cloud of the pump ducting room is mainly composed of pumps and equipment. The data is available in public on the Internet (E57 Data Examples 2010). Both of the collected point clouds are firstly pre-processed to remove isolated points by using an outlier-filtering algorithm.

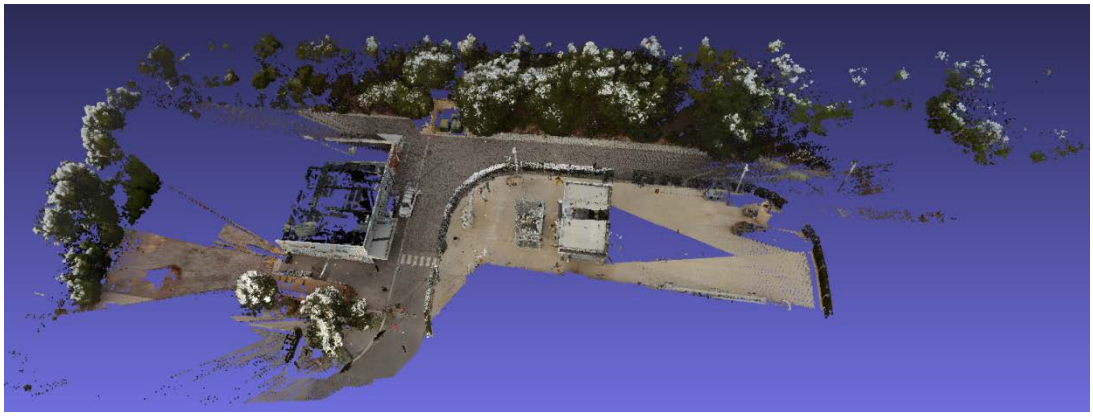

Figure 5.6. The point cloud of a LNG process training site

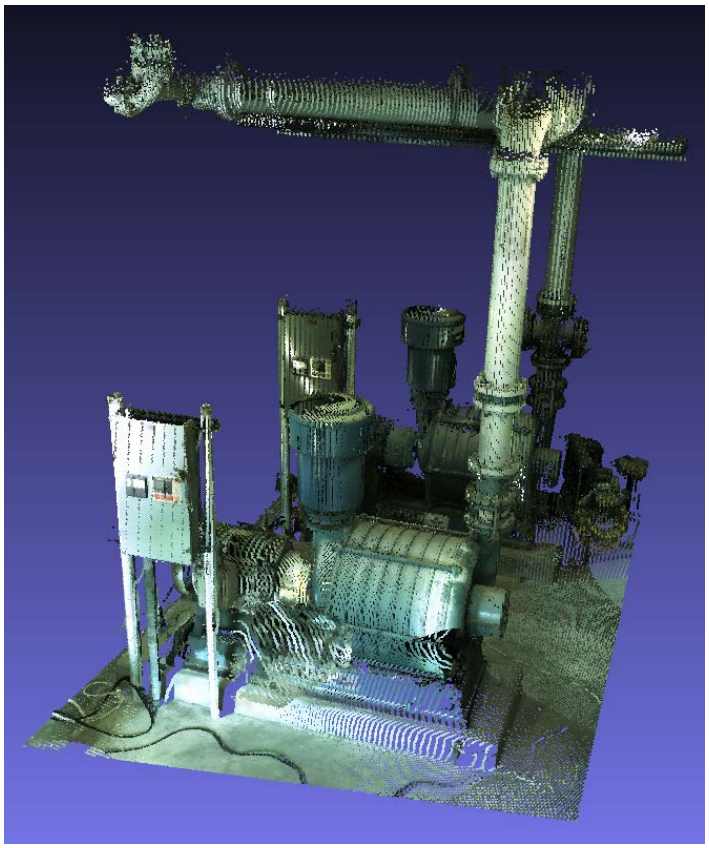

Figure 5.7. The point cloud of a pump ducting room

5.3.2 Results and discussions of semantic segmentation

The results of semantic segmentation of the pump room data are presented in Figure 5.8, where different clusters are labelled with different colours. The clusters include

pipes with different diameters (red or green), controlling boards (blue), pumps (yellow) and the ground (cyan). Figure 5.8(a) shows segmentation results with SVM only, while Figure 5.8(b) displays results of the proposed method integrating SVM and MRF. As can be seen in the Figure, most points are clustered into correct semantic clusters with the developed method, while there are small incorrect patches inside each cluster. There are also some errors around the adjacent areas between two objects and the adjacent line suffers from minor noise. Compared to the approach with SVM only, the developed approach produces an effective and more coherent segmentation, as small-scattered point clusters are significantly reduced by considering the neighbouring context. This can be seen from the fact that there are fewer patches with a different colour inside a semantic cluster. The effect of contextual constraints is clearly displayed in the Figure.

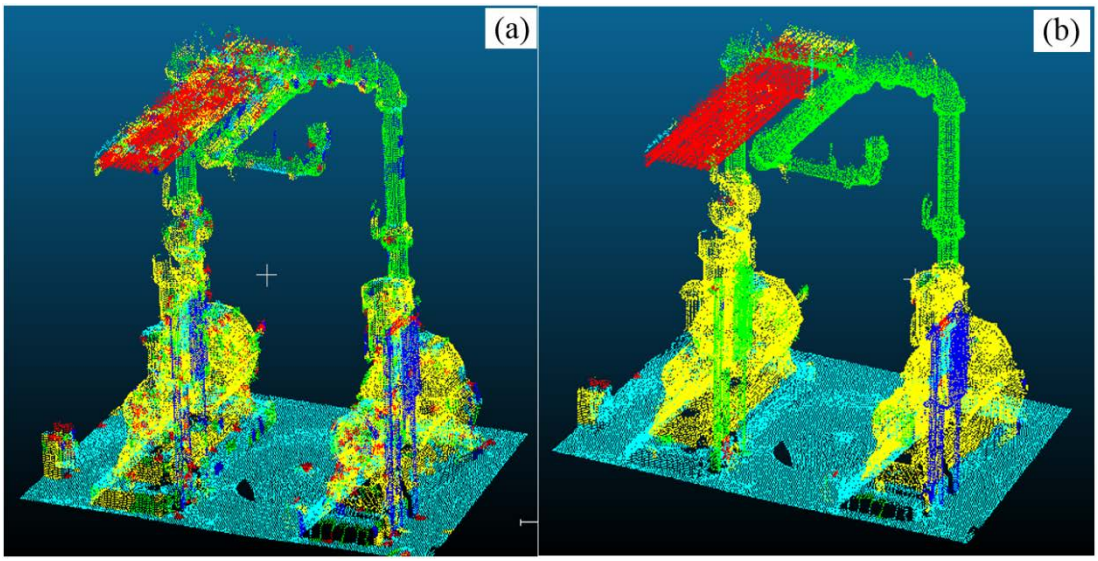

Figure 5.8. Segmentation results of the point cloud of the pump ducting room: (a) segmentation results with SVM only and (b) segmentation results with SVM and MRF

To give a quantitative evaluation on the performance of the developed approach, we present two indicators: recall and precision of the semantic segmentation results. The two indicators are estimated as follows:

$$\boldsymbol{precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FP}}$$
$$\boldsymbol{recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \qquad 5.10$$

where $\boldsymbol{TP}$ represents the number of supervoxels that are correctly segmented into a category, $\boldsymbol{FP}$ represents the number of supervoxels that are wrongly segmented into the category and $\boldsymbol{FN}$ represents the number of missed supervoxels.

Table 5.4. The performance of semantic segmentation on pump ducting room data

Objects	**SVM**		**SVM and MRF**	
	Recall (%)	Precision (%)	Recall (%)	Precision (%)
Pipe 1	59.73	66.17	91.14	93.53
Pipe 2	61.45	60.61	71.23	71.59
Control Panel	29.24	74.96	44.48	95.53
Electromotor	75.71	65.12	93.30	74.11
Ground	76.39	74.65	85.27	91.26

Detailed results of the pump room case are shown in Table 5.4. In the table, each row represents a category and each column represents corresponding indicators. The table shows that before considering contextual information without MRF, the precision of using SVM only is just over 60%, while it shows that the precision of the semantic segmentation with MRF is over 71% for all categories. The recall rates for each category with SVM only are 59.73%, 61.45%, 29.24%, 75.71% and 76.39% respectively, while those after using MRF are 91.14%, 71.23%, 44.48%, 93.30% and 85.27% respectively. The average recall rate increases from 60.50% to 77.08%. Although the recall rate of the control panel is only 44%, there is still an increase of 15% compared to that of using SVM only. The recall for the remaining categories is over 70% while that without using MRF is approximately 60%. As shown in Figure 5.8, one of the control panels is misclassified as pipes, which leads to a low recall rate. To sum up, after considering contextual information among classes with MRF, the overall quantitative indicators including both recall and precision show a big

increase and results show that the proposed approach is feasible and effective for semantic segmentation.

The results of the LNG process training site case are presented in Figure 5.9. The whole point cloud is divided into four semantic categories labelled with different colours, including: fence (blue), trees (green), ground (cyan) and industrial components (red). Figure 5.9(a) shows segmentation results with SVM only, while Figure 5.9(b) displays results of the proposed method integrating SVM and MRF. For the semantic segmentation with MRF, the pairwise cost weights among different categories are set given the prior information as specified in Table 5.3. As mentioned previously, the fence is the only neighbour to the ground in practice, and as a result, all other labels except the ground neighbouring the fence will incur large energy costs. Figure 5.9 shows that the overall segmentation with the developed method is correct while a small portion of the fence is mislabelled as industrial components. Similar to the results of the pump room, errors are most commonly identified along adjacent lines between two clusters, especially those between the ground and the industrial components. Despite the errors, the overall results achieved make the proposed approach useful to a rapid 3D mapping of LNG plants with laser scanning.

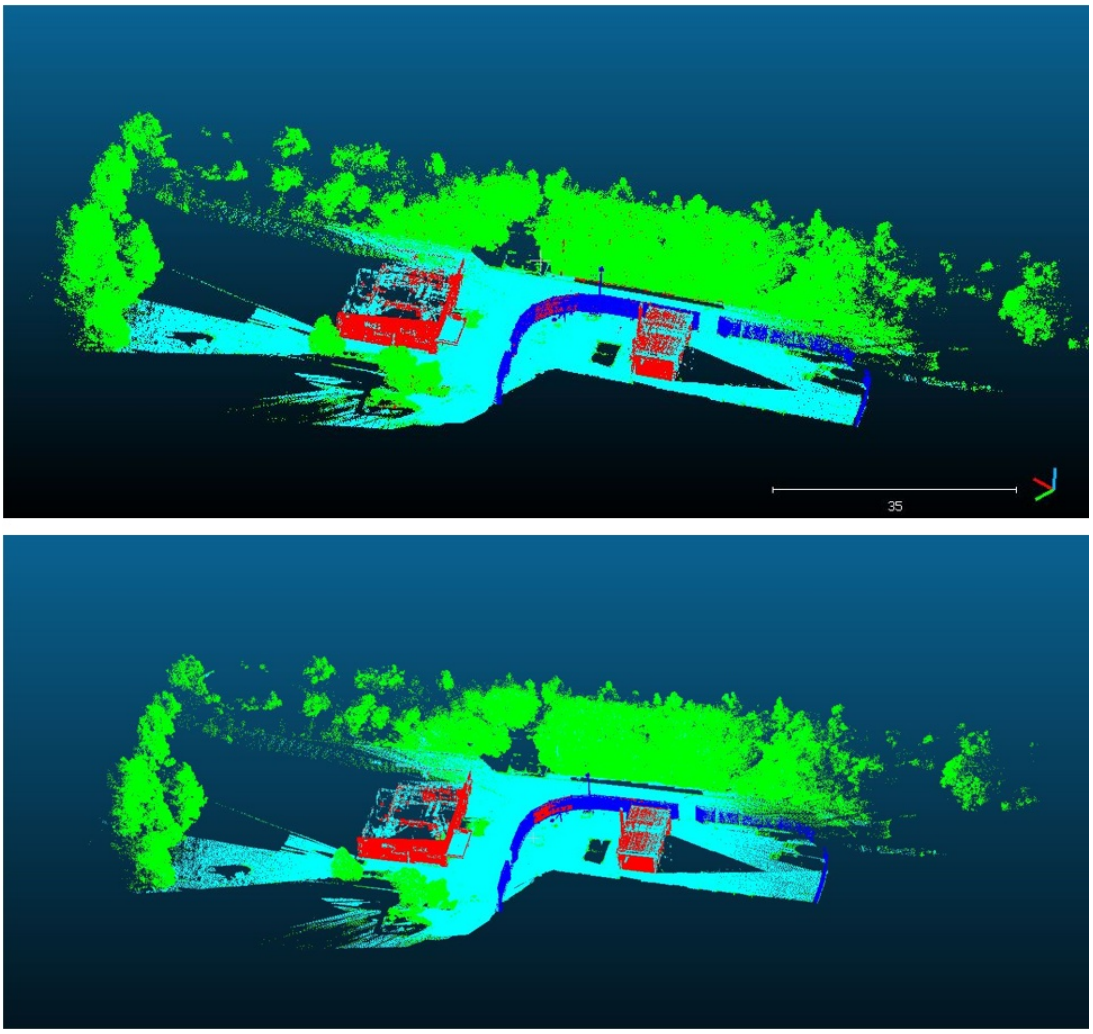

Figure 5.9. Results of semantic segmentation of the LNG process training site

Quantitative results are also shown in Table 5.5. The recall rates for four categories with SVM and MRF are 95.13%, 97.91%, 83.29% and 57.91% respectively while those with SVM only are 95.07%, 97.52%, 82.98% and 57.13% respectively. Both recall and precision rates see a large increase after using contextual information. The recall for industrial components is around 58%. The results show that components lying directly on ground are mislabelled as ground, which causes a low recall rate. In addition to industrial components, all other objects display a high recall rate, which is higher than 83% and a high precision rate, which is greater than 88%.

Table 5.5. The performance of the semantic segmentation on the LNG training site data

Objects	**SVM**		**SVM and MRF**	
	Recall (%)	Precision (%)	Recall (%)	Precision (%)
Fence	95.07	98.17	95.13	98.91
Trees	97.52	94.10	97.91	94.78
Ground	82.98	90.93	83.29	91.23
Industrial components	57.13	79.81	57.91	88.46

5.3.3 Results and discussions of the component recognition

To make a quantitative analysis for evaluating recognition efficiency and effectiveness, we tested the implemented algorithm on selected types of components. The recognition results for two types of valves are presented. As shown in Table 5.6, the overall success rates of recognition are 89% for valve type I and 100% for valve type II. One reason for failed recognition is the severe occlusions of components inside the LNG module, which means that only a small amount of components could be scanned. The average time spent on recognition mainly depends on the size of the point cloud data, as the matching algorithm tests all of the candidate points within corresponding categories. The results show that an average of 238 seconds is spent on recognising valve type I and an average of 400 seconds is spent on valve type II.

Table 5.6. Recognition results for valves in collected point clouds

Component	**Total Amount**	**Recognised Amount**	**Average Time Spent on Recognising Each Component(s)**
Valve Type I	8	7	238
Valve Type II	5	5	400

5.4 Summary

In this chapter, we developed an automated method of understanding as-built objects from point clouds in industrial facilities. The method integrates contextual information with segmentation methods based just on local features by using the Markov Random Field to enhance a class-level interpretation. The contextual information includes neighbouring relationships between different kinds of objects. In addition, supervoxelisation is used to group neighbouring points into local clusters, which makes the following segmentation more robust and more efficient than methods based on individual points. Moreover, an object-level interpretation was achieved with local feature matching and clustering. Experiments were conducted to test the proposed approach with different datasets. Quantitative results showed the feasibility and performance from the tested data.

6 SCAFFOLDING MONITORING WITH IMAGES IN LNG PLANTS

This chapter presents methods of monitoring scaffolding structures in a LNG plant. Scaffolding is commonly used as a supporting structure in the construction, maintenance and dismantling phases of oil and gas projects. To monitor the status of scaffolding construction, this research developed a vision-based strategy. Depending on the characteristics of different types of scaffolding, the developed strategy initially used different approaches for detecting scaffolding from raw data. Based on detected scaffolding, the strategy then inferred the current quantity and progress of the scaffolding construction.

Figure 6.1 displays the overall process of the developed strategy on monitoring scaffolding with vision-based methods. The developed process takes photos of scaffolding as inputs. If the scaffolding in photos displays a different appearance against its background, the process uses colour-based segmentation to detect the scaffolding. Failing this, the process seeks to reconstruct 3D models of the scene where there are sufficient multi-view input photos. The reconstructed 3D model can produce a depth map that enables the detection of the scaffolding. After the

scaffolding detection, the process uses template matching to find repetitive patterns and estimates the quantity of the scaffolding. In situations with failed scaffolding detection or no repetitive patterns, the process uses linear feature detection and analysis to estimate the quantity of the scaffolding.

The chapter is organised as follows: each section represents the individual steps of the proposed strategy as shown in Figure 6.1. Section 6.1 presents scaffolding detection approaches, with image segmentation based either on colours (described in subsection 6.1.1) or 3D reconstructed models (described in subsection 6.1.2). As for the images which cannot be reconstructed with 3D models, subsection 6.1.3 presents an alternative: a linear feature analysis method for scaffolding detection. Once the scaffolding is identified, Section 6.2 presents methods of estimating scaffolding progress with template matching. Section 6.3 presents the linear feature-analysis approach of still images for estimating scaffolding quantity and progress. Finally, experiments and discussions are presented in Section 6.4.

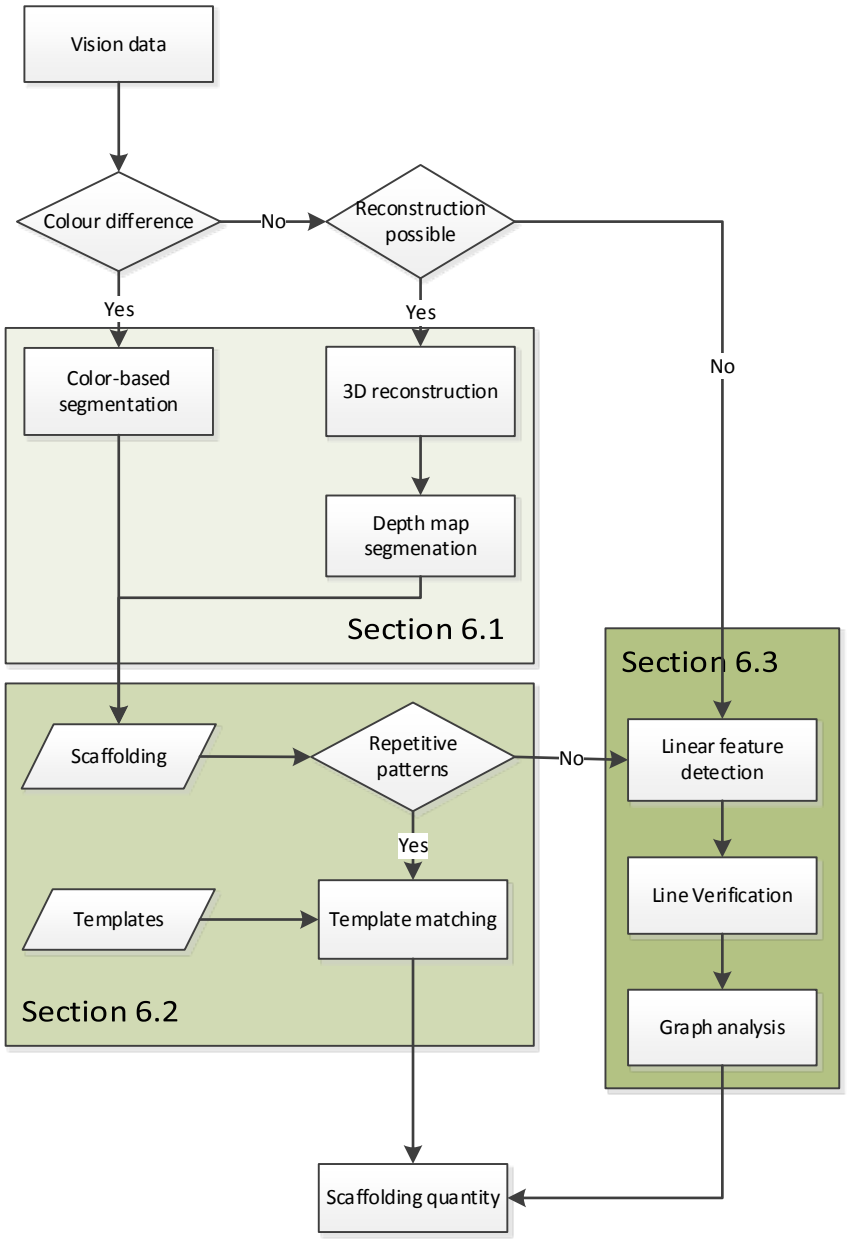

Figure 6.1: Process of the developed scaffolding monitoring strategy

6.1 Methods of scaffolding detection from photos

A key step of the vision-based monitoring approach is to detect scaffolding effectively from images. However, it is not really possible to use a general approach to detect all types of scaffolding structures from images, given the fact that there can be significant differences among different types of scaffolding. In addition, different backgrounds have influences on the choice of segmentation methods. Given this, a

hierarchical strategy of scaffolding detection is proposed, according to types of scaffolding and characteristics of backgrounds. For each possible scenario, different detection methods are proposed.

In this research, experiments were also conducted to evaluate the performance of various image segmentation methods for dividing the scaffolding and backgrounds. Specifically, colour-based and depth-based segmentation approaches were exploited.

6.1.1 Scaffolding detection with colour-based segmentation

Where there were differences between scaffolding and backgrounds in their appearance, straightforward image segmentation techniques were used based on features such as colour and texture, to segment scaffolding into raw images (Kaur et al., 2016). Here, we used the Gaussian Mixture Model (GMM) and graph cut segmentation algorithm (Szeliski et al. 2008). As mention in Section 5.1, graph cut is a segmentation method minimising the energy of a MRF. Graph cut algorithms consider contextual constraints among neighbouring pixels, which is a more robust method used to eliminate outliers and local noise.

Differing from point cloud segmentations, each node in the graph represents a pixel and each edge represents a connection between a pair of neighbouring pixels. The weight of an edge measures the dissimilarity between corresponding connected pixels. The unary energy of the graph is estimated with a GMM and the smooth energy is estimated according to feature differences of connected pixels. The segmentation method uses the colour values of each pixel in HSI (stands for Hue, Saturation, and Lightness) colour space as features.

The GMM is a statistical model that describes characteristics of data in the parameter space as a linear combination of Gaussian distribution (Permuter et al. 2003). Assuming that pixel features corresponding to both scaffolding and background in an image follow Gaussian distributions, the probability of a pixel belonging to each category can be estimated by the GMM, as in Equation 6.1 and Equation 6.2:

$$\boldsymbol{f(x) = \sum_{i=1}^{k} p_i N\left(x \middle| u_i, \sigma_i^2\right)} \qquad 6.1$$

$$N(x|u_i, \sigma_i^2) = \frac{1}{\sigma\sqrt{2p_i}} exp \frac{-(x-u_i)^2}{2\sigma_i^2} \quad 6.2$$

where $\boldsymbol{x}$ is the feature vector of a pixel in an image; *k* is the number of categories; $\boldsymbol{p_i}$ are weights of each categories; and $\boldsymbol{u_i}$ and $\boldsymbol{\sigma_i}$ are the mean and standard deviation of a category respectively.

The parameters of the GMM are estimated with an Expectation Maximization algorithm (Szeliski et al. 2008).

Figure 6.2 presents an example of the colour-based segmentation approach. The left image shows the raw photo of the scaffolding and the right one displays results of the segmentation, where different objects are coded with different colours.

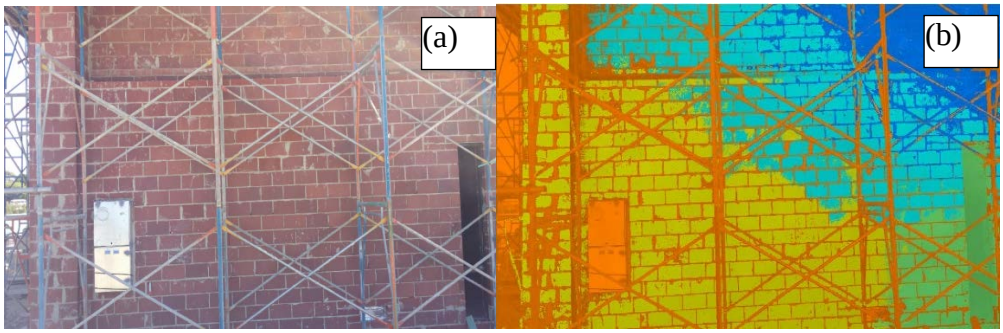

Figure 6.2. Scaffolding detection with colour-based segmentation: (a) raw image and (b) segmentation result with graph cut

6.1.2 Scaffolding detection with depth differences

In cases where the scaffolding and backgrounds have a similar appearance, it is difficult to detect the scaffolding with appearance-based segmentation (Dey et al., 2010). For instance, some scaffolding tubes are of similar colour to pipes in LNG facilities, as shown in Figure 6.3. To achieve an effective segmentation of these types of scaffolding, a depth-based approach is developed to distinguish the scaffolding in the images.

Figure 6.3. The scaffolding and backgrounds with similar appearance

This depth-based approach is based on the observation that scaffolding is usually erected on the surface of a facility. The spatial configuration of scaffolding against the surface produces possible distance differences from a viewpoint, and thus a local extrema in a depth map. The depth map is an image which relates to the distance of an object from a viewpoint.

A depth map is essential to understanding the 3D geometry of a scene and it can be estimated with stereo matching methods (Yang, 2015) from binocular photos or multiple photos. Stereo matching finds a correspondence between different photos of the same object and there are many existing tools available (Dall'Asta & Roncella 2014). To illustrate the effects of the depth information, Figure 6.4(b) shows a depth map corresponding to the raw image shown in Figure 6.4(a). It shows that there are obvious differences between scaffolding tubes and other areas in the depth map.

Given their differences in a depth map, the scaffolding and backgrounds thus can be segmented in a simple way. To be specific, a thresholding method is used. Thresholding changes pixels under a certain threshold into zero (treated as background) and pixels above the threshold into one (treated as scaffolding). A critical part of this method is to determine thresholds that divide pixels of images into background and foreground. The Otsu method (Sezgin, 2004) is applied to find the threshold that minimises the overlap between the background and foreground.

After thresholding, the scaffolding can be segmented into the foreground of a binary image while the remaining areas are the background.

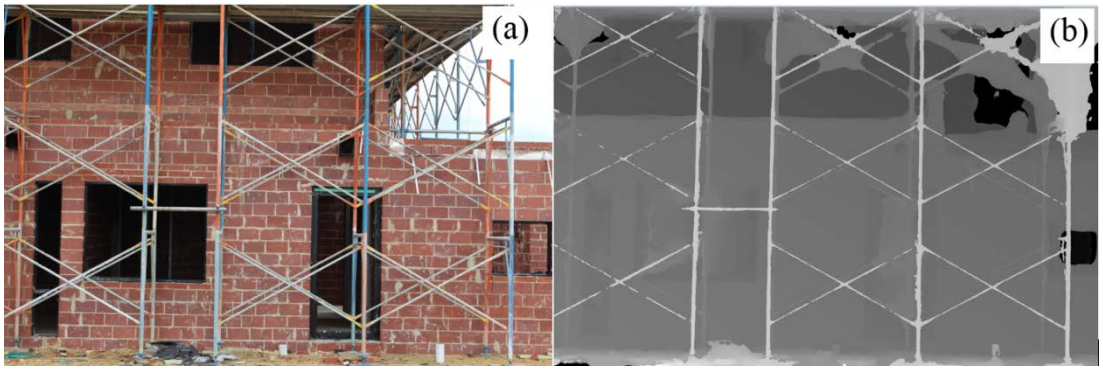

Figure 6.4: Scaffolding detection with depth differences: (a) raw image and (b) corresponding depth map

6.1.3 Scaffolding detection with linear feature analysis

In this section, an automated detection method of scaffolding is introduced by using photos with linear feature analysis.

This study proposed scaffolding detection approaches aimed at recognising scaffolding structures in images, abstracting the geometric structures of scaffolding as a graph and deriving final progress estimation results based on the graph analysis. The flowchart of the proposed method is presented in Figure 6.5. Scaffolding grid lines, which are used to determine the actual scaffolding structure, are detected first through extracting horizontal and vertical linear features in an image. In order to extract the horizontal and vertical components while eliminating other disturbing components, a Gaussian derivate filter (Shrivakshan, 2012) is used as a directional filter to obtain the edge map. A Hough transform is then used to detect horizontal and vertical lines in the image. With the initial detected lines being further grouped and judged, actual scaffolding segments are then drawn out.

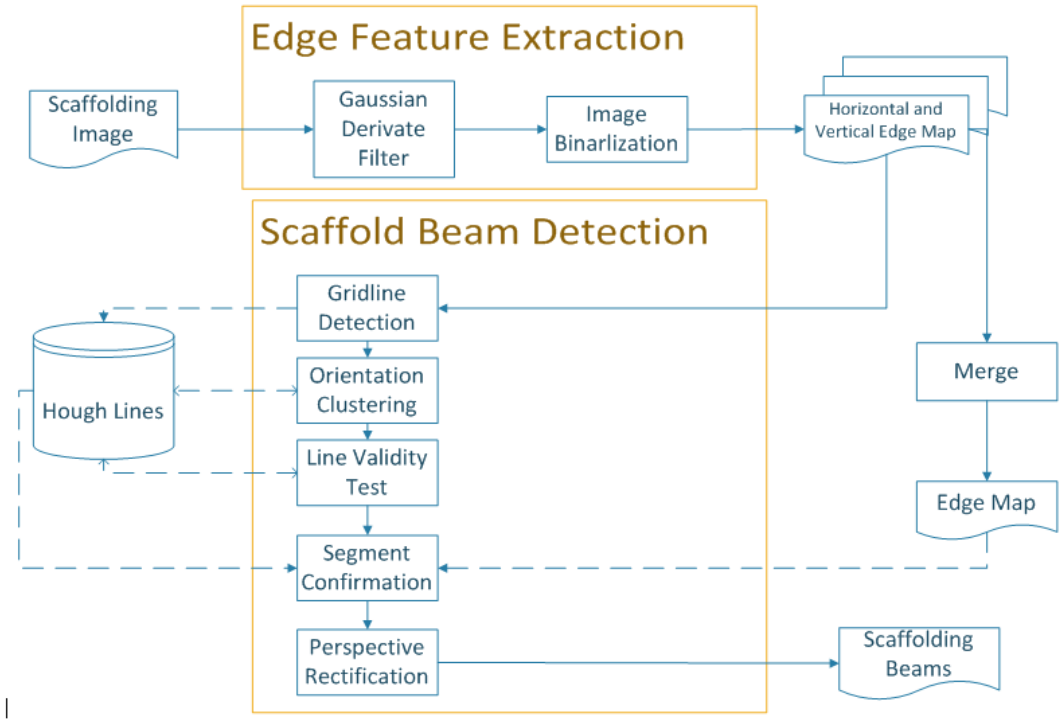

Figure 6.5. Flowchart of proposed scaffold beam detection method

The rest of this subsection is organised in a top-down sequence to introduce the scaffolding beam detection process. Starting from edge detection, it moves to scaffolding grid line recognition throughout the image and then progresses to breaking grid lines into line segments in order to represent a single scaffolding beam.

6.1.3.1 Edge feature extraction

Feature Extraction is used to emphasise and extract potential edges which represent scaffolding grid lines in a scaffolding image for further Hough transform (Duda & Hart 1972). A Gaussian derivate filter, a high-pass filter derived from the Gaussian Filter, is adopted to extract high-pass components, including edges, into the image. The image is processed by discrete convolution with a specific kernel matrix through introducing a differential equation to the kernel matrix of the Gaussian filter.

The kernel matrix of Gaussian filter with a radius of $\boldsymbol{r}$ is denoted in Equation 6.3:

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{-r,-r} & \cdots & \mathbf{G}_{-1,-r} & \mathbf{G}_{0,-r} & \mathbf{G}_{1,-r} & \cdots & \mathbf{G}_{r,-r} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \iddots & \vdots \\ \mathbf{G}_{-r,-1} & \cdots & \mathbf{G}_{-1,-1} & \mathbf{G}_{0,-1} & \mathbf{G}_{1,-1} & \cdots & \mathbf{G}_{r,-1} \\ \mathbf{G}_{-r,0} & \cdots & \mathbf{G}_{-1,0} & \mathbf{G}_{0,0} & \mathbf{G}_{1,0} & \cdots & \mathbf{G}_{r,0} \\ \mathbf{G}_{-r,1} & \cdots & \mathbf{G}_{-1,1} & \mathbf{G}_{0,1} & \mathbf{G}_{1,1} & \cdots & \mathbf{G}_{r,1} \\ \vdots & \iddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_{-r,r} & \cdots & \mathbf{G}_{-1,r} & \mathbf{G}_{0,r} & \mathbf{G}_{1,r} & \cdots & \mathbf{G}_{r,r} \end{bmatrix} \quad 6.3$$

To maximise the performance in the discrete convolution, the elements of the Gaussian filter are determined as shown in Equation 6.4:

$$\mathbf{G}_{x,y} \propto e^{-2\sigma^2} I_x(\sigma^2) I_y(\sigma^2) \quad 6.4$$

where $\boldsymbol{\sigma} = \frac{r}{2}$, and $\boldsymbol{I_\alpha(t)}$ is a Modified Bessel function of the first kind.

The derivation of the above kernel matrix is calculated and called the kernel matrix for the Gaussian derivate filter. As gradient direction is always perpendicular to the edge direction, $\boldsymbol{\partial_x \mathbf{G}}$ and $\boldsymbol{\partial_y \mathbf{G}}$ are used to filter out the horizontal components and vertical components respectively. As for the calculation of the derivate, the following Equation 6.5 is applied towards the *x* direction and *y* direction respectively:

$$\begin{aligned} \boldsymbol{\partial_x \mathbf{G}_{x,y}} &= \frac{1}{2}\left(\mathbf{G}_{x+2,y} + \mathbf{G}_{x-2,y} - 2\mathbf{G}_{x,y}\right) \\ \boldsymbol{\partial_y \mathbf{G}_{x,y}} &= \frac{1}{2}\left(\mathbf{G}_{x,y+2} + \mathbf{G}_{x,y-2} - 2\mathbf{G}_{x,y}\right) \end{aligned} \quad 6.5$$

The reason for filtering out the horizontal and vertical components separately is to separate the detection of horizontal and vertical lines in order to eliminate potential interference at a single edge map, and thus simplify the calculation. If the original image is a colour image, a Gaussian derivate filter can be applied to each channel respectively before the image is converted into grayscale. Filtered images can be binarised using a proper threshold to obtain the edge maps. An example of edge feature detection is displayed in Figure 6.6. Figure 6.6(b) and Figure 6.6(c) show the horizontal and vertical edge maps respectively, while Figure 6.6(d) combines them as a final edge map, which clearly illustrates the main structures of scaffolding beams.

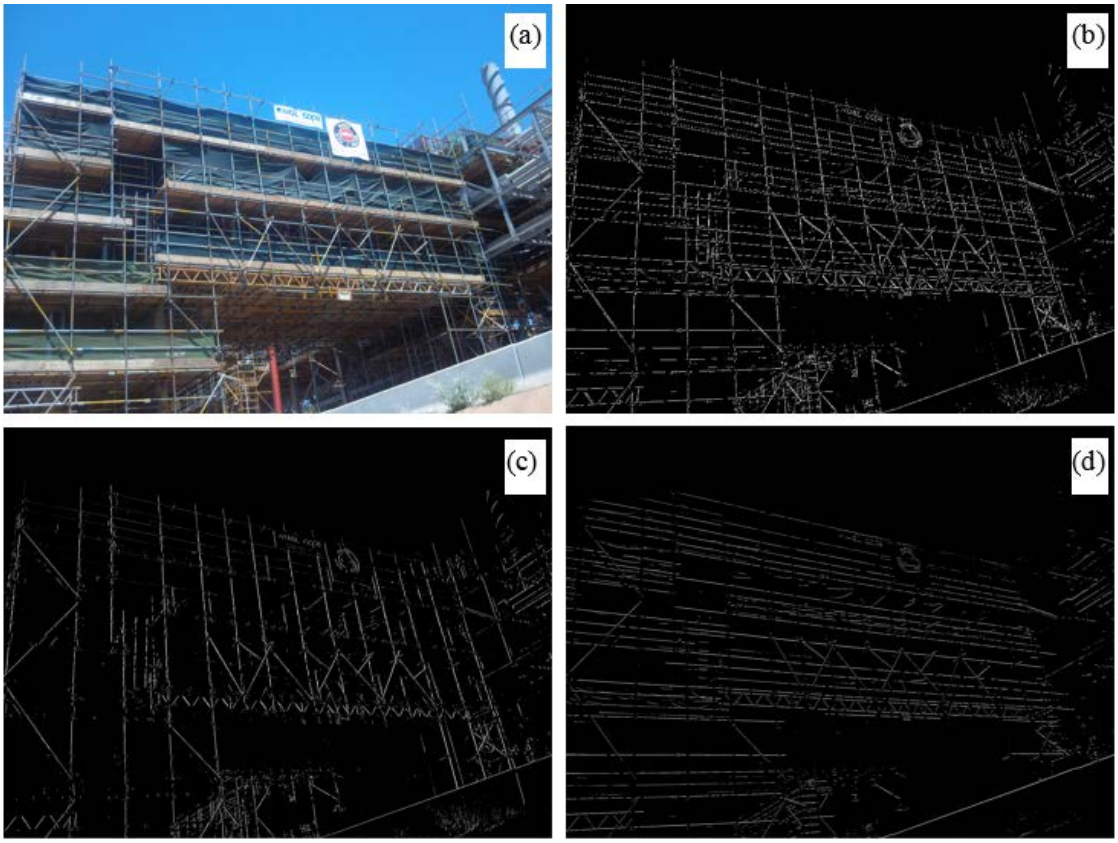

Figure 6.6. Results of scaffolding feature detection: (a) raw image; (b) horizontal edge map; (c) vertical edge map; and (d) combined edge map

6.1.3.2 Scaffold grid line detection

After extracting edges in the image, the proposed method converts visible edges on the edge map into lines with a Hough transform (Duda & Hart 1972) .

A Hough transform is a powerful tool for line detection that can be used to generate candidate scaffolding grid lines. It is based on mapping from the image space to a parameter space called Hough space. Given a point $(\boldsymbol{x}, \boldsymbol{y})$ on the image space, each straight line in the image going through that point can be represented in a polar equation as in: $\boldsymbol{x}\cos(\boldsymbol{\theta}) + \boldsymbol{y}\sin(\boldsymbol{\theta}) = \boldsymbol{\rho}$, and can therefore be mapped to the parameter space as a point $(\boldsymbol{\rho}, \boldsymbol{\theta})$. The curve: $\boldsymbol{x}\cos(\boldsymbol{\theta}) + \boldsymbol{y}\sin(\boldsymbol{\theta}) = \boldsymbol{\rho}$ in the parameter space represents all straight lines which go through the point $(\boldsymbol{x}, \boldsymbol{y})$ on the original image. If a straight line exists in the edge map, then all the points in this line

share the same point $(\boldsymbol{\rho}, \boldsymbol{\theta})$, thus resulting in a peak in the parameter space. By finding those peaks in the parameter space which meet a certain intensity threshold, length threshold and neighbourhood elimination rules, the lines existing in the edge map can be determined. Figure 6.7(a) is an illustration of the Hough space after the Hough transform has been used. Horizontal and vertical lines are detected respectively according to their corresponding edge maps, for the purpose of reducing mutual interference, and then combined as the results show in Figure 6.7(b):

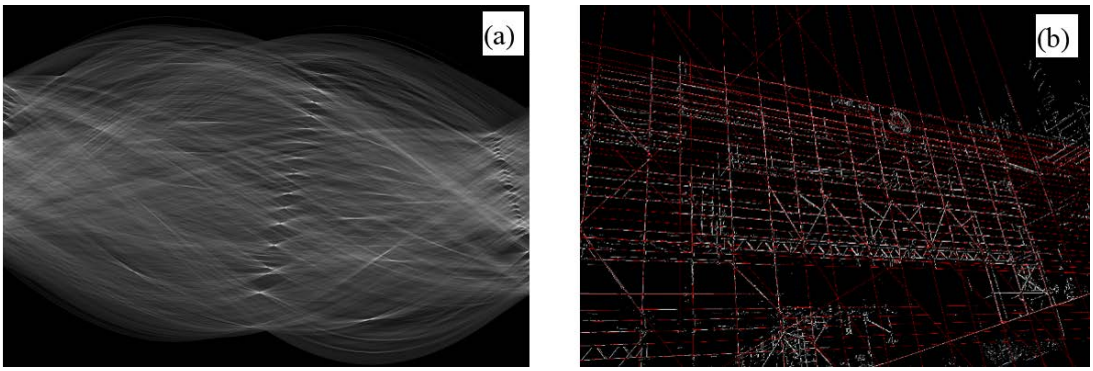

Figure 6.7. (a) Hough space of combined edge map; and (b) results of line detection

In the Hough line detection process, as mentioned above, three parameters, including an intensity threshold, a length threshold and a neighbourhood restraint rate should be set. Lowering the values of these three parameters increases the number of lines detected, which leads to an increase in the recall rate of scaffolding beams and thus, a decrease in precision, since more undesired lines are involved. There are two main aspects responsible for bringing in those undesired lines:

1. There are other irrelevant objects besides scaffolding such as trees, cars, nearby buildings and other construction materials. The edge map would contain noises for counting in these edges which is unwanted and may give out undesired lines.
2. Energy is concentrated around the Hough peaks in Hough space. Lowering the neighbourhood restraint rate is helpful when the distances

between parallel lines are small, but this may lead to redundant lines representing the same scaffolding beam.

Figure 6.7 (b) shows an example of line detection where detected lines are marked in red on the edge map. Almost all scaffolding beams are correctly detected. However, in the bottom section where trees and complex structures lie, the noise is high and therefore many undesired lines appear. Since only scaffolding grid lines are expected, further processing of the detected lines is necessary to remove outliers.

6.1.3.3 Orientation clustering

Clustering lines according to their orientations is helpful for both eliminating undesired lines caused by irrelevant objects and enhancing the robustness of the algorithm when photos are taken in apparent tilt. Single-linkage clustering is adopted to cluster all detected lines on the image plane. It is a hierarchical clustering method in a bottom-up fashion. Each line, as a sample, is regarded as a cluster at the beginning, then clusters are repeatedly merged according to the shortest distance between samples from two different clusters. The method is also called nearest-linkage distance, as written in Equation 6.6:

$$\mathbf{d}(\mathbf{X},\mathbf{Y}) = \mathbf{min}\{\mathbf{dist}(\mathbf{x},\mathbf{y}) \mid \mathbf{x} \in \mathbf{X}\ \mathbf{and}\ \mathbf{y} \in \mathbf{Y}\} \tag{6.6}$$

The merging process stops when the distance between clusters is significantly large, usually above a given threshold. Thus the eventual number of clusters is not pre-defined but calculated automatically.

By selecting two clusters representing horizontal and vertical lines, almost all inclined lines, especially paralleled diagonal beams in the scaffolding, can be automatically discarded. Figure 6.8 shows an example of the orientation clustering of a scaffolding photo. Because the orientation degree of a line only spans one $\boldsymbol{\pi}$ area, the clustering is carried out twice with orientation degrees ranging between $\left(-\frac{\boldsymbol{\pi}}{\mathbf{2}}, \frac{\boldsymbol{\pi}}{\mathbf{2}}\right]$ and $[\mathbf{0}, \boldsymbol{\pi})$ respectively. Judging by similar nearest-linkage distance, the cluster nearest to $\mathbf{0}°$ in the $\left(-\frac{\boldsymbol{\pi}}{\mathbf{2}}, \frac{\boldsymbol{\pi}}{\mathbf{2}}\right]$ range is regarded as a vertical line cluster, and the cluster nearest to $\mathbf{90}°$ in the $[\mathbf{0}, \boldsymbol{\pi})$ range is regarded as a horizontal line cluster.

Figure 6.8 shows an example of the line orientation clustering in two ranges, where line orientations are displayed in degrees and different clusters are marked with different colours. The left image shows clustering results in the range $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and the right image shows clustering results in the range $[\mathbf{0}, \boldsymbol{\pi})$.

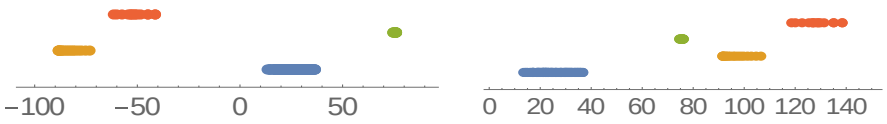

Figure 6.8. Clustering results of line orientations (in degrees)

6.1.3.4 Line validation

Though many unexpected lines could be removed through the clustering and selecting process, there still remain some questionable lines which have a similar orientation to common scaffolding grid lines. It is impossible to distinguish these lines from scaffolding grid lines with orientation only. To remove these lines, this research proposed a "Line Validity Test" method. The method is based on the perspective relationship of lines in an image.

The basic assumption of this line validity test, called Parallel Assumption, is that all horizontal beams or vertical beams should be parallel to each other. This assumption is reasonable, since almost all the scaffolding established nowadays would fit this in terms of structural stability, constructional convenience and costs. Based on the Parallel Assumption, another two significant restrictions, which scaffolding grid lines must adhere to, can be made:

Restriction I: Lines within a cluster should never intersect within the image. They may intersect outside an image, if prolonged, since there is prospective distortion of the image, which will be discussed in Restriction II. However, this should never occur inside the image. In view of this inference, it can be claimed that all the horizontal or vertical lines which intersect with those of the same kind, are suspect and require further judgement.

Restriction II: Ideally, for a series of parallel lines in the space, the slope or the reciprocal of the scope of their projection lines and their coordinates in the image plane should fit in a linear model. According to Perspective Geometry theory, all parallel lines in the space will remain parallel if they are projected onto a plane parallel to them (Hartley & Zisserman 2004). Otherwise, they will intersect at the same point, called the vanishing point, if they are projected onto a plane where they are not parallel to each other. These projected lines are prolonged long enough to show vanishing points, as illustrated in Figure 6.9. In this figure, the image plane is represented as a rectangle.

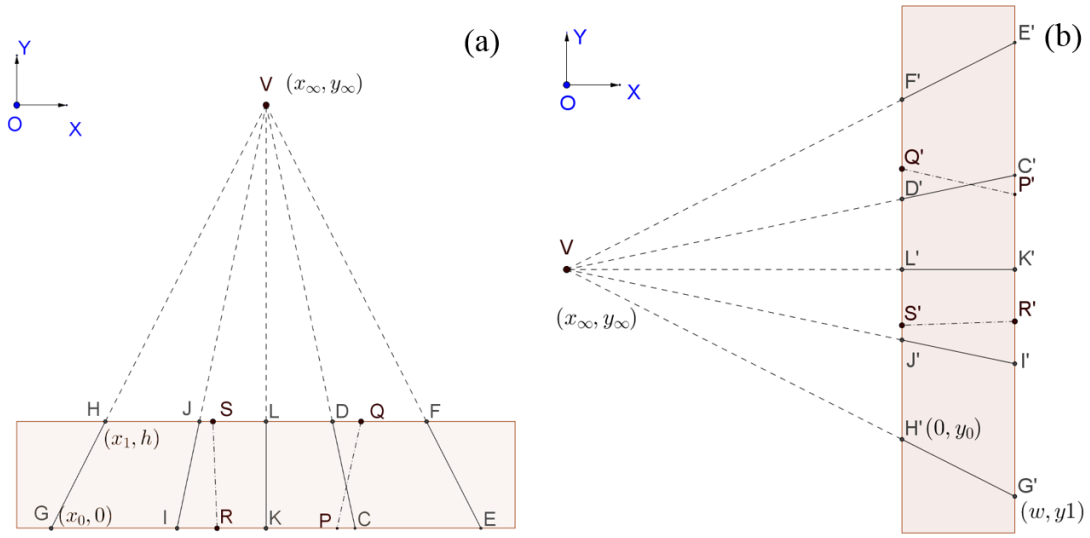

Figure 6.9. Relationship among projections of lines on a plane not in parallel with them: (a) vertical lines; and (b) horizontal lines

Ideally, all these lines (or their extensions) parallel to each other in physical space, such as vertical lines denoted as Line CD, EF, GH, IJ, KL, would share the same vanishing point O in an image (plane), denoted as $(\boldsymbol{x_\infty, y_\infty})$. In contrast, lines not parallel to each other, such as Line PQ and SR, will not go through the vanishing point. For each line, its intersection with the image upper boundary is denoted as $(\boldsymbol{x_1, h})$. With these notations, the slope of the vertical lines $\boldsymbol{k}$ can be represented as Equation 6.7:

$$\frac{1}{k} = \frac{x_\infty - x_0}{y_\infty} \quad 6.7$$

Since the vanishing point (x_∞, y_∞) and image height h are shared by all parallel lines and are therefore constants in the above equation, it is clear that a linear relationship exists between $\frac{1}{k}$ and x_0.

As for horizontal lines, their vanishing point would appear at the left or right side of the image, thus the equation will be slightly different**,** as shown in Equation 6.8:

$$k = \frac{y_\infty - y_1}{x_\infty} \quad 6.8$$

And a linear relationship exists between k and y_1.

Due to the inaccuracy involved in Hough line detection, in actual cases these lines may not intersect exactly at the same ideal vanishing point but they should intersect near it. Also, $\left(\frac{1}{k}, x_0\right)$ or (k, y_1) may not fit an exact linear model. However, by using least square fitting, the approximate linear relationships of $\left(\frac{1}{k}, x_0\right)$ or (k, y_1) could be estimated. Moreover, the relationship estimated should not be too far from the actual cases. If there is any sample point too far from the estimated linear model, its corresponding line on the image will be suspect with regard to its validity. This means that those lines in physical space which are not parallel to the majority of other lines are therefore not a part of the scaffolding grids. The following Figure 6.10 illustrates the process of the linear regression. In this figure, subfigure 6.10(a) and 6.10(b) show original sample points representing the horizontal lines and vertical lines detected in the scaffolding image shown in subfigure 6.10(c) as well as their estimated linear models respectively. In both subfigures, the horizontal axis denotes $\frac{1}{k}$ and the vertical axis denotes x in pixels. Red dots represent lines detected in the scaffolding image and blue lines represent the fitted linear model.

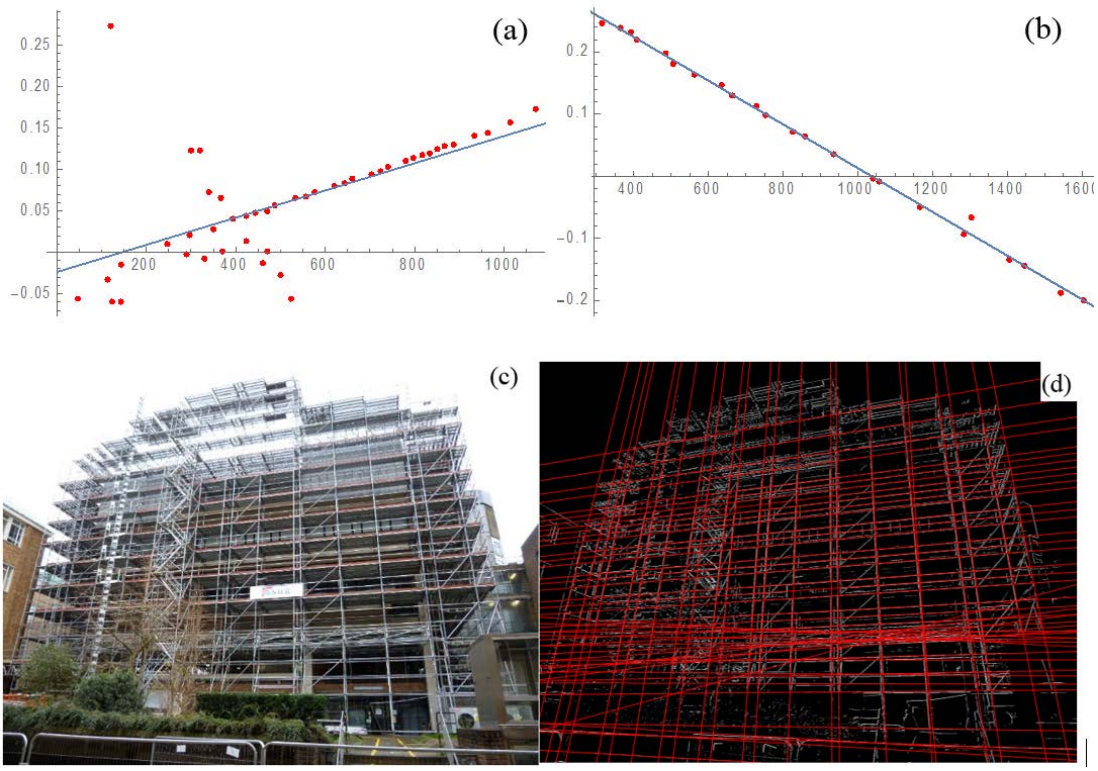

Figure 6.10. Linear regression results: (a) regression results of horizontal grid lines; (b) regression results of vertical grid lines; (c) original scaffolding image; and (d) detected grid lines

As shown in Figure 6.10, the vertical components obviously fitted better with the linear model. Comparing this with Figure 6.10(d), vertical lines also matched well with the real scaffolding original image. However, the horizontal part contains many outliers, also be found in Figure 6.10(d), which are obviously irrelevant to the scaffolding.

So far, two restrictions for scaffolding grid lines have been drawn from the basic Parallel Assumption. Restriction I points out that the target of further line validity judgement should be crossing lines, and Restriction II helps provide a quantity method of judging which lines are more likely to fit the prospective rules of scaffolding beams.

According to Restriction I, all the lines which intersect with other lines are suspect. However, it is obvious that removing these lines altogether is not a wise decision as

it may remove both unconvincing lines and convincing ones. Thus, some strategies must be followed to make an orderly examination and removal of those suspect lines. This research designed a strategy based on the intersection relationship and the above Restriction II.

If each line is regarded as a node on a graph, and a link exists between two nodes if their corresponding lines are intersecting, the intersecting relationship between lines can be observed intuitively as shown in Figure 6.11(a).

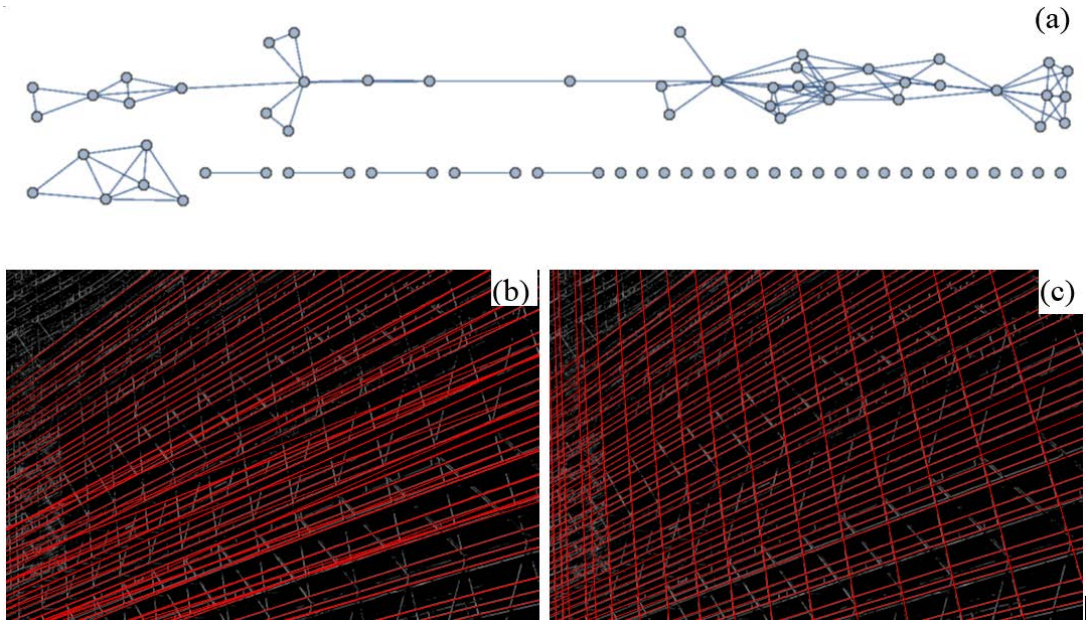

Figure 6.11. (a) A graph example of the intersection relationship between lines; (b) horizontal scaffolding grid lines before the detection; and (c) horizontal and vertical scaffolding grid lines after the detection

This intersecting relationship graph can be divided into several connecting components. Each connecting component with more than one node suggests the existence of intersection among those nodes. The largest connecting component involves more suspect lines, especially those significant outliers which usually create numerous intersections. It is therefore safer to deal with them first. The smallest connecting components, formed with only two lines, require more effort in clarification of the suspect lines. Each time when examining a connecting component, as in Restriction 2, a linear model will be fitted with least square fitting with current

remaining lines. The line with the largest deviation to the fitted model will be removed. The process iterates until there is no intersection among lines in this component. The detailed flowchart of the algorithm is shown in Figure 6.12 as follows:

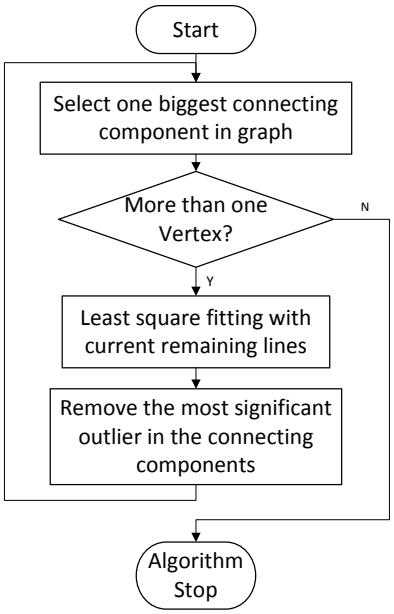

Figure 6.12. Flowchart of the line validation algorithm

Testing shows that this method effectively removes all questionable lines, and retains the main grid lines of the scaffolding. Figure 6.11(b) and Figure 6.11(c) show an example, where the questionable horizontal scaffolding grid lines are removed after the detection.

6.1.3.5 Scaffold beam confirmation

Via the above steps, the grid lines of scaffolding have been successfully obtained. However, actual scaffolding beams may just be a part of these grid lines, as there may be space for gaps and openings. In order to survey the precise number of scaffolding beams, these grid lines need to be separated into line segments, and it is also necessary to finally confirm whether a line segment is actually related to a real scaffolding beam. Horizontal and vertical grid lines are represented as a whole grid, and divided into line segments by intersecting points on the grid. To confirm valid

line segments, the overlap between each line segment and the edge map obtained from the feature extraction should be examined. The edge image is firstly dilated with a given radius to make up for the inaccuracy involved in the Hough line detection, since these detected lines may not overlap with the lines in the edge map exactly. Those line segments with the percentage of the overlap above a given threshold would be regarded as valid scaffold beams.

An example of the scaffold beam confirmation is displayed in Figure 6.13. Specifically, Figure 6.13(a) shows the edge map marked with grid lines after the dilation; Figure 6.13(b) shows the overlapped part of grid lines and dilated edge map; Figure 6.13(c) and Figure 6.13(d) show the segment confirmation results overlapped on the edge map and the original image respectively.

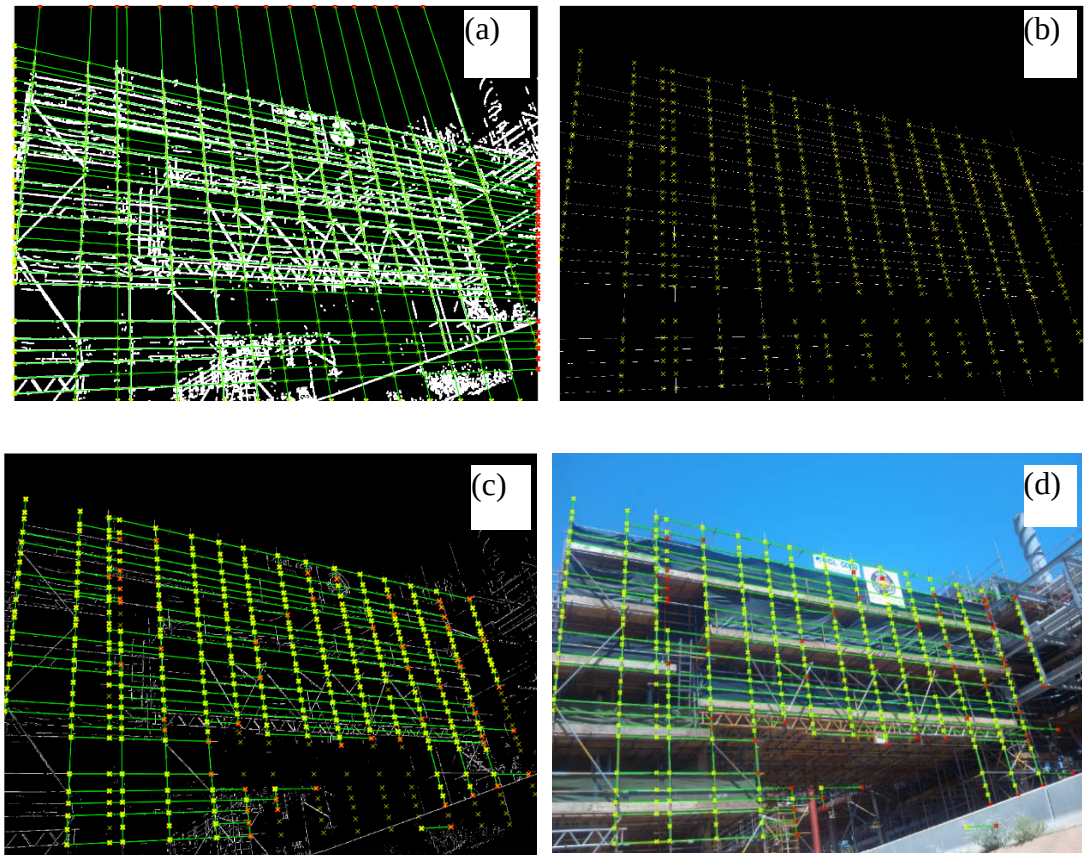

Figure 6.13. Illustration of line segments confirmation process: (a) detected grid lines overlapping on the edge map; (b) overlapping area between grid lines and edge

maps; (c) confirmed line segments overlapping on the edge map; (d) confirmed line segments overlapping on the raw image

6.1.3.6 Perspective rectification

Scaffolding can be mainly represented as a basic structure of rectangles. Thus, quantities of scaffolding can be estimated by the numbers of those rectangles. However, because of perspective transformation, a rectangle in physical space will be a quadrangle in the image. It is necessary to rectify the image of a rectangular structure into a rectangular shape before detection. Moreover, in the rectified image all the scaffolding lying on the same plane in physical space is of the same scale. Thus the length of a scaffolding beam can be obtained by the length of its corresponding line in the rectified image. Figure 6.14 shows the effect of homograph transformation on rectifying a quadrangle into a rectangle.

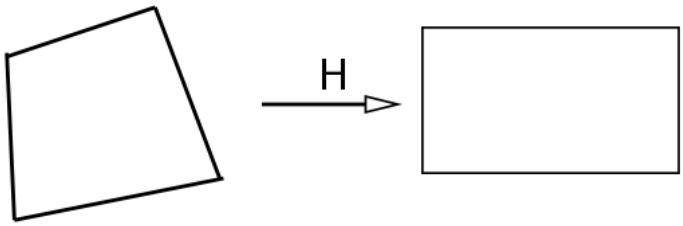

Figure 6.14. Rectification of a rectangle with homograph transformation: original shape (left) and rectified shape (right)

According to single-view geometry of imaging, there exists a homograph transformation between two images of a plane (Hartley & Zisserman 2004). The homograph transformation is represented as Equation 6.9:

$$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \\ \mathbf{1} \end{bmatrix} = \boldsymbol{H} \begin{bmatrix} \boldsymbol{x'} \\ \boldsymbol{y'} \\ \mathbf{1} \end{bmatrix} \qquad 6.9$$

where $\boldsymbol{H}$ is a $\mathbf{3} \times \mathbf{3}$ homograph matrix for the rectification, and $[\boldsymbol{x} \quad \boldsymbol{y} \quad \mathbf{1}]^{\boldsymbol{T}}$ and $[\boldsymbol{x}' \quad \boldsymbol{y}' \quad \mathbf{1}]^{\boldsymbol{T}}$ are homography coordinates of a corresponding point in raw and rectified images respectively.

Given the coordinates of at least four corresponding points between two images, the homograph matrix can be estimated with a least square optimisation (Agarwal et al. 2005). In the developed approach, the four vertexes of a rectangle are used to estimate the homograph matrix. The original image is then rectified according to the homograph transform.

6.2 Scaffolding quantity estimation via template matching

After the detection of scaffolding beams, two approaches can be used to determine the quantity of scaffolding that has been established and captured in the image. The first method is to carry out the pattern matching process in order to find existing assembling patterns of the detected beams in a pre-defined library. This method is discussed in this section (Section 6.2). Another method that follows the concept of line feature analysis is introduced in the next section (Section 6.3).

Scaffolding is typically a manually assembled structure, which mainly comprises some fixed patterns. Typical patterns in a scaffolding structure are listed but not limited as shown in Figure 6.15. Based on this observation, it is possible that quantities of the scaffolding can be estimated by counting the numbers of those patterns. This transforms the problem into detecting pre-defined patterns in a library from segmented images.

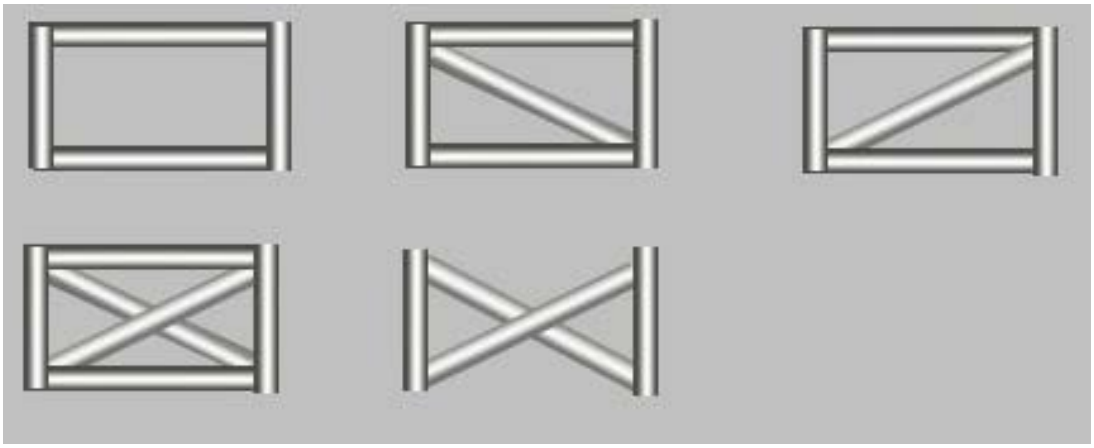

Figure 6.15 Typical patterns in a scaffolding structure

First, we establish a library of shape templates through investigating assembly rules and structures of the scaffolding. In the library, each shape template corresponds to a scaffolding pattern, which is represented with a binary image. One advantage of this is that the proposed approach is easily extended to different kinds of structures by adding new templates to the library.

The proposed matching approach utilises a Generalized Hough Transform (GHT) to detect scaffolding patterns from images (Ballard, 1981). GHT is a method for detecting an arbitrary non-analytic shape from an image. The GHT method takes a template of a shape and an image as inputs, and determines locations of objects matching the template in the image. Through iteratively detecting each template from a segmented scaffolding image, numbers of each shape in the library are estimated.

In the maintenance projects of process plant**s**, especially in Australia, the quantities of scaffolding are usually measured by weight, typically by tons. To do this, a model between numbers of scaffolding shapes and the respective weight is defined. The model between shapes and weights is related to specifications of different scaffolding, which means it varies with different scaffolding manufacturers and projects.

6.3 Scaffolding quantity estimation via linear feature analysis

After the scaffolding beams have been detected through the above process in Section 6.1, the next step is to use the collected information to estimate the actual quantity of the work performed in a scaffolding project. Since there is a variety in scaffolding design and in on-site practice, quantity survey methods should be based on specific circumstances depending on different working areas and scaffolding types. This research only suggests some common practices that are technically available and can be adopted regarding actual conditions.

The developed method aims to produce statistical indicators of scaffolding quantities which meet the practical requirements of project managers. In current practice in Australia, the total quantity of scaffolding is usually represented by the mass in terms of tons. According to our on-site survey of workers in LNG plants, the numbers of each basic element, such as beams, joints, and their total weight, are significant indicators used in current progress survey and material management. These numbers are manually surveyed at present. However, they can be automatically calculated from the proposed detection results. The method estimates four key indicators, including the number of beams, length of beams, number of joints and number of basic rectangle patterns. By further considering prior-knowledge on the scaffolding used in a project, the mass of the scaffolding can also be inferred from the four indicators. The indicators and their related estimations**,** based on scaffolding detection results, are presented in the following subsections.

A grid graph is used to model the scaffolding structure detected with the proposed method. The grid graph is built from the scaffolding grid lines model, where intersections of scaffolding grid lines are mapped as vertexes on a grid graph, and each scaffolding beam is represented using one or more connecting edges on the grid graph. Figure 6.16(b) shows actual examples of scaffolding grid graphs corresponding to Figure 6.16(a), where blue lines represent possible scaffolding grid lines, and thick red lines represent actual scaffolding beams which form a subset of the whole of the scaffolding grid lines.

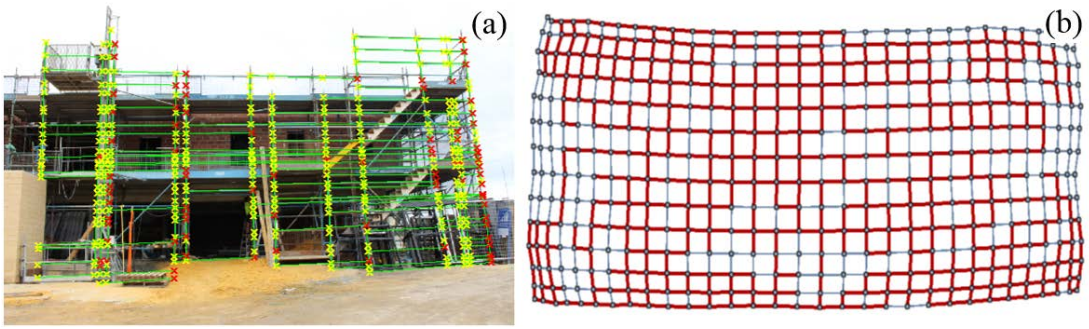

Figure 6.16. Examples of scaffolding grid lines an**d** corresponding graph

6.3.1 Tally by beams and joints

A straightforward way to carry out quantity estimation is to simply count the total number of each kind of element, e.g. beams and joints of the scaffolding. To tally the number, the relationship of the detected scaffolding is represented as a grid graph as described above. To illustrate the process of object counting, an example of a grid graph is shown in Figure 6.17, where the left image shows marked segments in scaffolding grid lines as green lines, and their intersection as yellow or red crosses. The grid graph corresponding to Figure 6.17(a) is displayed in Figure 6.17(b).

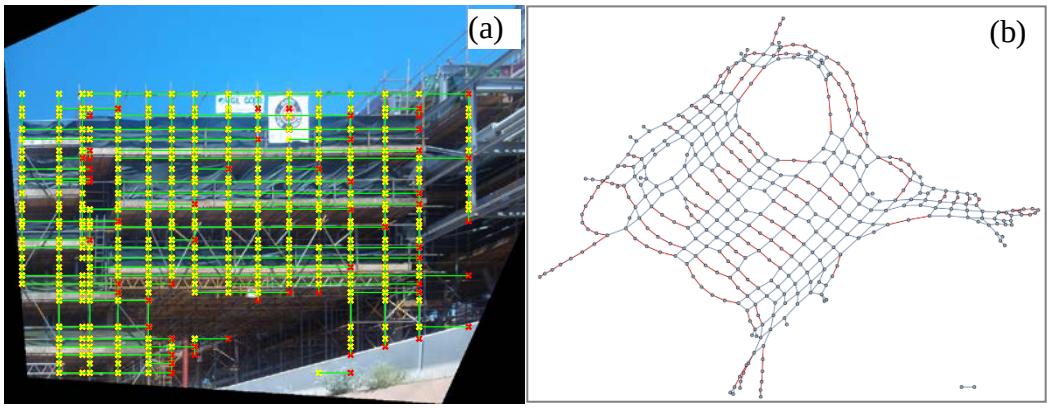

Figure 6.17. (a) Results of scaffolding beam detection; and (b) grid graph of scaffolding beams

A single scaffolding beam may go through several scaffolding grid intersections where the scaffolding does not follow a regular pattern, as illustrated in Fig 6.16. In Figure 6.17(b), each strip of connecting edges marked red in the graph goes through

several vertices, or scaffolding grid line intersections. In actuality, they represent a single scaffolding beam, and should be merged in the process of object counting. To perform the merging process the following steps must be undertaken:

1. Separate the graph into two subgraphs, representing horizontal scaffolding beam components and vertical scaffolding beam components separately.
2. Find vertices representing intersections of scaffolding beams. These vertices should either have a degree larger than 3, which means at least three edges connect to each one of them, or be the intersection of horizontal and vertical components. These vertices form a set S.
3. In each of two components, select a vertex which has a degree of 2 but not in set S. Beginning from that vertex, search the nearest vertices in set S towards a bilateral direction. Denote these two vertices as V_1 and V_2, and edges travelled in the searching process form a set D.
4. Create an edge directly linking vertex V_1 and V_2, and then delete all edges in set D.

A simplified graph corresponding to the example in Figure 6.17 is presented in Figure 6.18(b), with its corresponding illustration in Figure 6.18(a). In this simplified graph, each vertex represents an identified scaffolding joint, which is represented as a circle dot in the image. Each edge in the graph represents an identified scaffolding beam. By counting the number of vertices and edges in the graph, the tally of scaffolding beams and joints can be achieved.

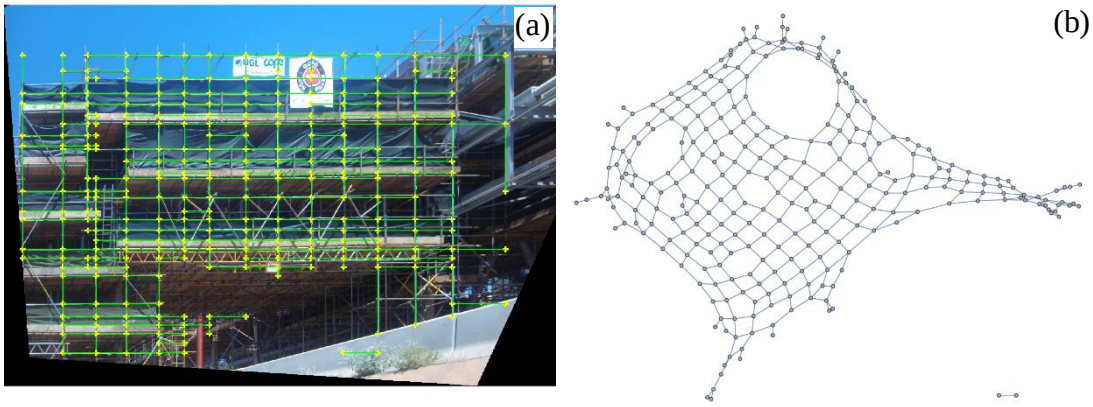

Figure 6.18. (a) Scaffolding beams after a graph simplification; and (b) the corresponding simplified graph

6.3.2 Tally by total length

The limitation of tally by number is that it cannot represent the differences between the sizes or weights of different scaffolding beams. Survey by length could somewhat represent the differences in weight, by supposing that the density of scaffolding is uniform. Here, we use the total length of all detected beams as another quantity indicator.

After the rectification, the scaffolding beams in the image are at the same scale and thus, the physical length of a beam can be obtained by multiplying a scale factor by its length in the image. The scale factor can be simply obtained by calculating the ratio of the length of an object in physical space to its length in the image. One possible solution is to use a camera fixed at the same position repetitively taking photos. Thus, the scale is fixed for all images and can be calibrated during the setup. However, a long-term fixed position for camera installation may be hard to find in practice. Prior information on the size of a rectangular scaffolding structure is used to assist in deciding the scale factor. Specifically, we choose a detected rectangle in the rectified image, and get its size from an as-designed document. The scale factor is the ratio between the designed size and its measured size in the image.

6.3.3 Tally by rectangle patterns

By taking some specific scaffolding patterns, including their topology and geometric characteristics into consideration, a more detailed result can be derived. As proposed at the beginning of this section, scaffolding is established by horizontal and vertical beams, which means rectangles represent one of the most common patterns for scaffolding.

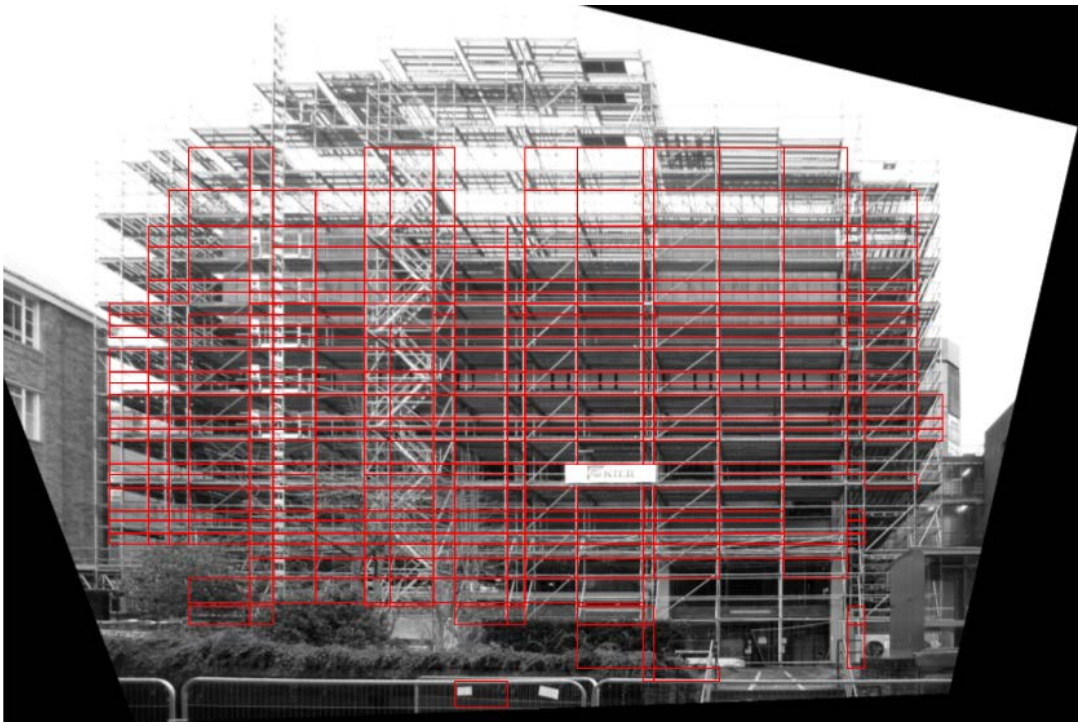

Figure 6.19. Rectangle components of scaffolding detected in an image

Only closed rectangles formed by four scaffolding beams would be regarded as a rectangle pattern. The detection and counting of the number of all these patterns is an indicator of the scaffolding quantity. Due to incompleteness in scaffolding beam detection, some rectangle patterns might be missing or falsely recognised in the detection. Figure 6.19 represents all detected rectangle patterns in a scaffolding image, where there are 455 rectangle patterns in total.

A more detailed result can be obtained if these scaffolding patterns are further distinguished. As observed in Figure 6.19, available distinguishing characteristics include rectangle area, rectangle perimeter and aspect ratio. By classifying rectangle

patterns with these characteristics, the difference between scaffolding beams can be counted in. With empirical equations derived from actual scaffolding designs, which use numbers of patterns in different categories to synthesise an indicator of project progress, the survey result can be closer to the actual situation.

6.4 Experiments and discussion

In this section, we present experiments to verify and evaluate the proposed methods for scaffolding detection and quantity estimation. The methods proposed in this research are implemented as a prototype using MATLAB and Mathematica 11. Different scenarios of scaffolding projects are selected to test the performance of the developed methods with photos captured on-site as inputs. The results of the scaffolding detection are firstly presented and examined intuitively. Quantity estimation results are then quantitatively evaluated with different indicators and compared with results of a manual inspection.

6.4.1 Data

The tested photos are captured in different scenarios**,** including a building construction site and a LNG plant.

In the building scenario, scaffolding is erected to assist in the construction of a two-story residential house. The scaffolding used in this scenario belongs to one class and is erected following a uniform pattern. Photos of the scaffolding around this house are captured with a consumable-level handheld camera. There are 55 photos covering different areas of the house.

In the LNG plant scenario, there are many scaffolding projects going on at the same time for maintenance purposes. These scaffolding projects use thousands of tons of scaffolding and cost millions of dollars per year. There are several different types of scaffolding on-site. Depending on the scaffolding classes and structures of different facilities, scaffolding is erected following different patterns and rules. Subject to the safety requirements in the LNG industry, a handheld camera, Gravity X from Pixavi (BARTEC PIXAVI 2017) is used to capture photos. The camera is intrinsically safe

and certified for use in hazardous environments such as LNG plants. The resolution of the camera is 2592× 1944 pixels. About 250 photos were taken in the plant, covering different scenarios of scaffolding structures.

6.4.2 Scaffolding detection

This subsection presents experiments of scaffolding detection with image segmentation methods, including the colour-based approach as proposed in subsection 6.1.1, and the depth-based approach as proposed in subsection 6.1.2, along with the linear-feature-based detection as proposed in subsection 6.1.3.

6.4.2.1 Test on colour-based scaffolding detection

Firstly, we conducted an experiment in detecting scaffolding with colour-based segmentation. The experiment inputs scaffolding photos along pipe racks in the LNG plant. Figure 6.20(a) shows the input photo. It shows that the colour of most bracing tubes is different to that of the background. Using a graph-cut with a GMM segmentation algorithm, the detection results are displayed in Figure 6.20(b), where the background is marked with black. It shows that the proposed approach can successfully detect the scaffolding for the input photo.

Figure 6.20(c) shows the corresponding manually labelled ground truth image, where the background is black. In Figure 6.20(d), the difference between the ground truth and detection results is shown, where coloured areas represent scaffolding pixels which failed to be detected. The difference shows that most missing pixels are on the edges of bracing tubes while the main areas of a tube are preserved in the results.

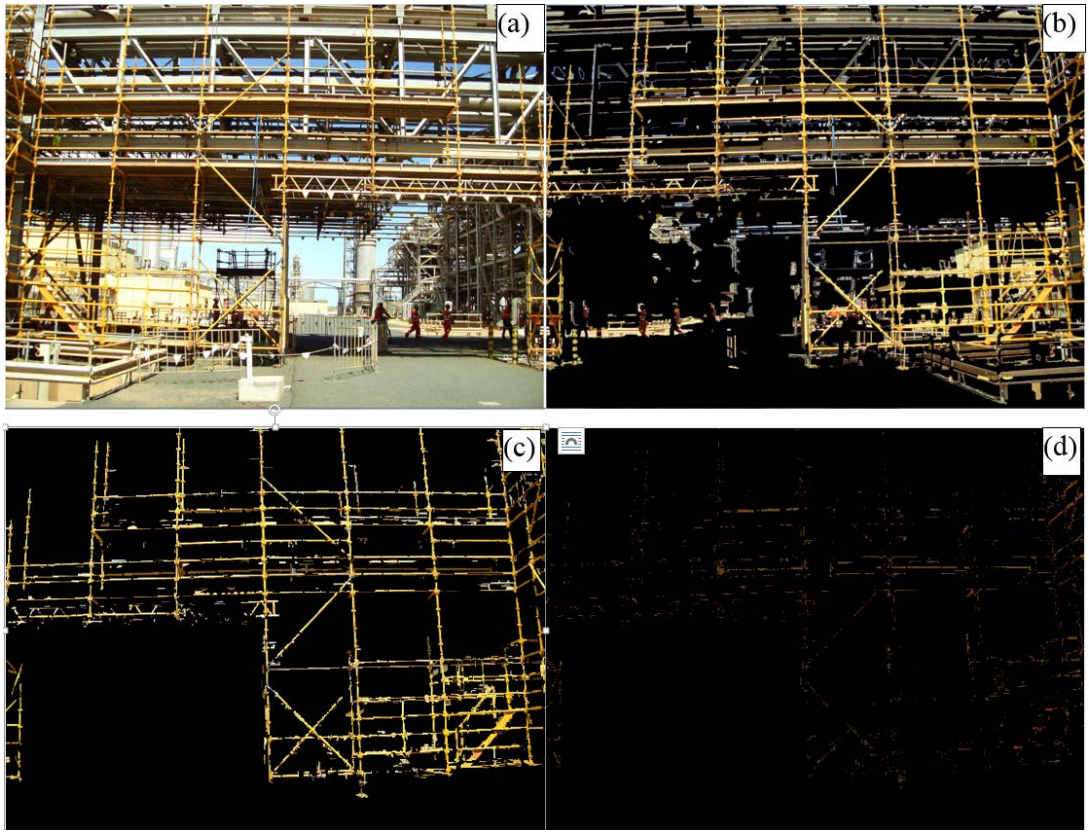

Figure 6.20. Results of scaffolding detection with colour-based segmentation: (a) a raw image, (b) results of the scaffolding detection, (c) the ground truth manually labelled and (d) undetected scaffolding pixels

6.4.2.2 Test on depth-based scaffolding detection

In addition, to evaluating the feasibility of the proposed depth-based detection approach, an experiment was conducted in a building construction site, as displayed in Figure 6.21(a). With photos of the building, the proposed approach generated a depth map of the building using multi-view stereo technologies. The depth map was then used to detect scaffolding with the discussed methods above.

The depth map of the building is displayed in Figure 6.21(a), where the intensity of a pixel represents relative depth. It can be seen from the image that the scaffolding is brighter than the backgrounds, due to their depth variation. The intensities of pixels belonging to the scaffolding and backgrounds show significant differences, which makes it easy to segment the scaffolding from the image with a simple thresholding

method. The binary image produced by thresholding is shown in Figure 6.21(b), where white represents the scaffolding. Even though there are some outliers and discontinuities, the thresholding produced a clear skeleton of the scaffolding.

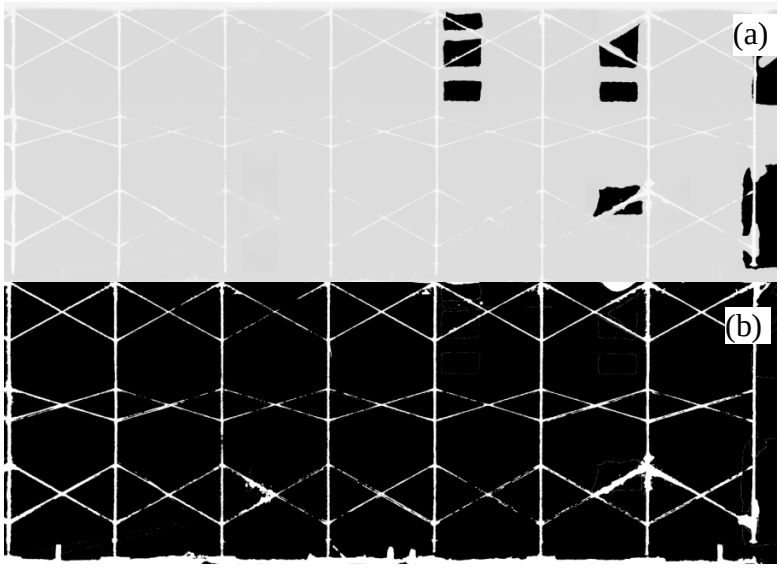

Figure 6.21. The results of scaffolding detection: (a) a depth map and (2) its binary image from thresholding

With the above depth map and the segmentation process, the results of the scaffolding detection are shown in Figure 6.22. Figure 6.22(b) shows detected scaffolding with the depth-based approach, where pixels not belonging to the scaffolding are black in colour. Figure 6.22(c) shows the corresponding manually lablled ground truth image. The differences between the ground truth and the detection results are shown in Figure 6.22(d), where coloured pixels represent actual scaffolding areas that are missing from detection results.

For further evaluation, we also give quantitative indicators of the scaffolding detection in Table 6.1, where the precision and recall of detection rates are presented. The indicators were obtained based on numbers of pixels. For the colour-based segmentation, the presented test achieved a recall of 67.0% and a precision of 37.7%. For the depth-based method, the recall was 68.9% while the precision was 59.7%.

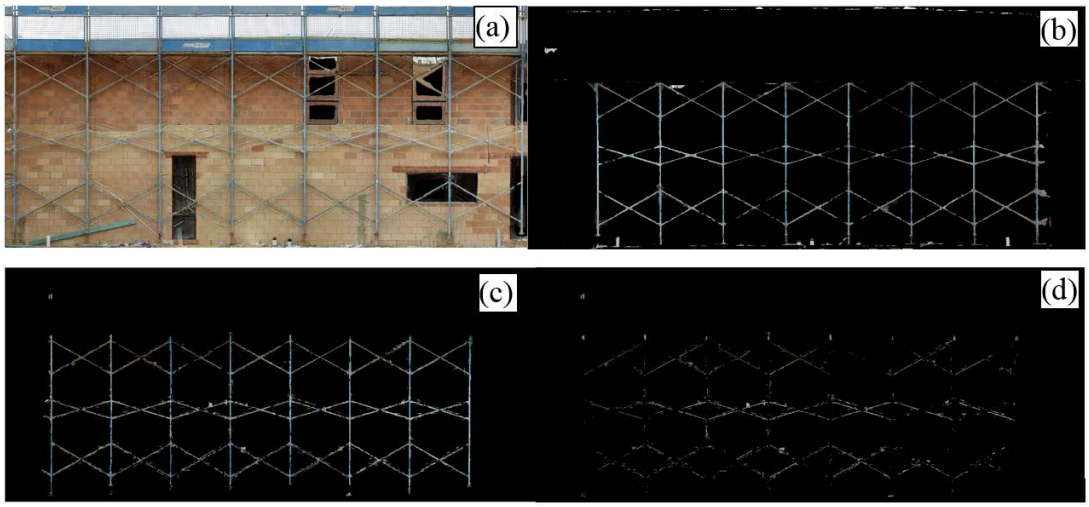

Figure 6.22: Results of scaffolding detection with depth-based segmentation: (a) a raw image, (b) results of the scaffolding detection, (c) the ground truth manually labelled and (d) undetected scaffolding pixels

Table 6.1 Precision and recall of scaffolding detection

Method	**Colour**	**Depth map**
Precision (%)	37.7	59.7
Recall (%)	67.0	68.9

The experiments show that both colour-based and depth-based detection approaches can detect the main skeleton of bracing tubes. The successful detection of the scaffolding skeleton makes the following scaffolding quantity estimation possible. Although the recall rate is not high, the results show that the main components of most scaffolding tubes are successfully detected. The evidence that most missing pixels shown in Figure 6.20(d) and Figure 6.22(d) are edge pixels of a bracing tube also supports this. However, this does not impact on the quantity estimation of the scaffolding, since the following quantity estimation process does not rely on the width of detected scaffolding.

6.4.2.3 Test on linear-feature-based scaffolding detection

The proposed linear feature analysis method for scaffolding detection was also tested with data from the LNG plant. From the data, we chose two typical scenarios in the plant and test results are presented in the following section.

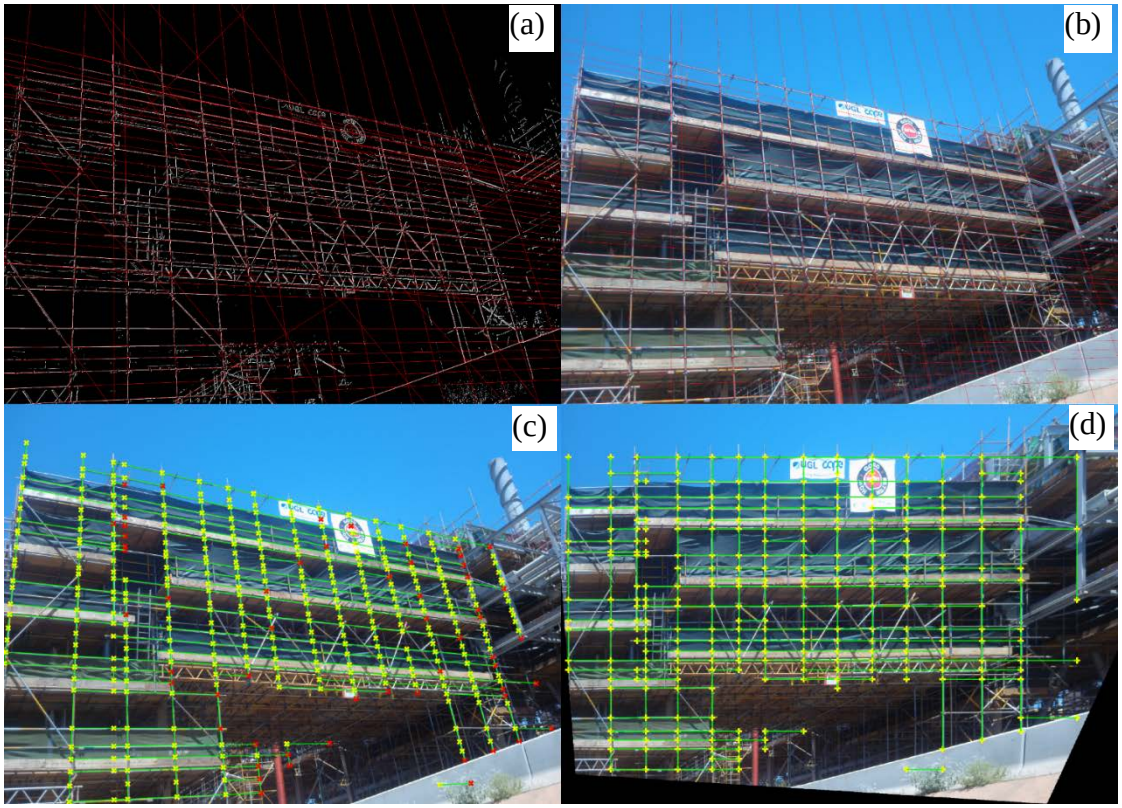

Figure 6.23. Example of scaffolding detection from an image: (a) Line detection results; (b) detected lines after the filtering; (c) detected scaffolding on the raw image; and (d) detected scaffolding beams on the rectified image

Figure 6.23 shows the results of interval steps of scaffolding detection with the proposed linear feature analysis approach. Starting from Image Binarisation and Hough Line Detection, the result of the initial step is shown in Figure 6.23(a). As shown in the figure, many questionable lines exist in the line set, including those crossing diagonally through the image, and those clustering near the valid scaffolding line. After a least square fitting and Graph-based filtering in Line Validity Test, a much better result can be achieved, which clearly represents the grid of the scaffolding structure as presented in Figure 6.23(b). Figure 6.24 shows the detailed filtering process on the least square fitting graph and intersection graph. Each dot represents a line detected in a photo of the scaffolding, and those red dots represent cancelled lines with a sequence of being cancelled marked near the dot. Figure 6.23(c) represents the remaining grid segments after Segment Confirmation.

The last Figure 6.23(d) represents the detected scaffolding beams in corrected prospective after merging non-meaningful grid intersections in Figure 6.23(c).

The test results show that scaffolding beams have been correctly detected from the photo. The test also shows that debris containment netting leads to an easier detection of the scaffolding as those mesh materials could either occlude other linear objects in the background or produce an appearance difference between the scaffolding and the background.

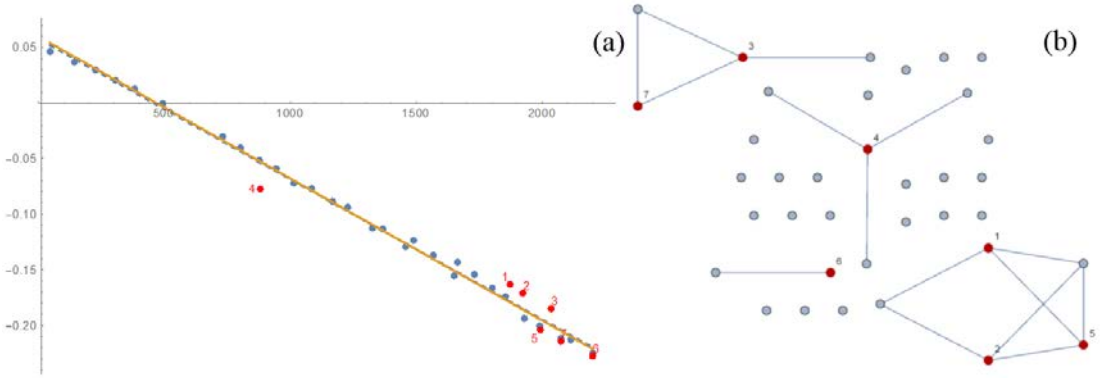

Figure 6.24. Filtering of line segments on the example shown in Figure 6.23: (a) linear regression results; and (b) the graph of the line intersection relationship

The following figure shows another more complicated example of scaffolding detection with the proposed linear feature analysis approach, given that there are more irrelevant construction materials laying on the ground. Figure 6.25(a) shows the result after Image Binarisation and Hough Line Detection, where a cluster of unclear lines occur due to complicated construction objects other than the scaffolding laying on the ground. Figure 6.25(b) shows a detected scaffolding grid after the Line Validity Test, where most questionable lines are deleted, though detection results are still influenced by these complicated objects. Figure 6.25(c) represents the remaining grid segments after Segment Confirmation. Figure 6.25(d) represents the detected scaffolding beams from a corrected perspective. Figure 6.26 shows the detailed filtering process on the least square fitting graph and intersection graph. Similarly, each dot represents a line detected in a photo of the scaffolding,

and the red dots represent cancelled lines with a sequence of being cancelled marked near the dot. Specifically, in Figure 6.24 (a), the dotted line represents the fitted linear model before filtering and the solid line represents the fitted model after filtering.

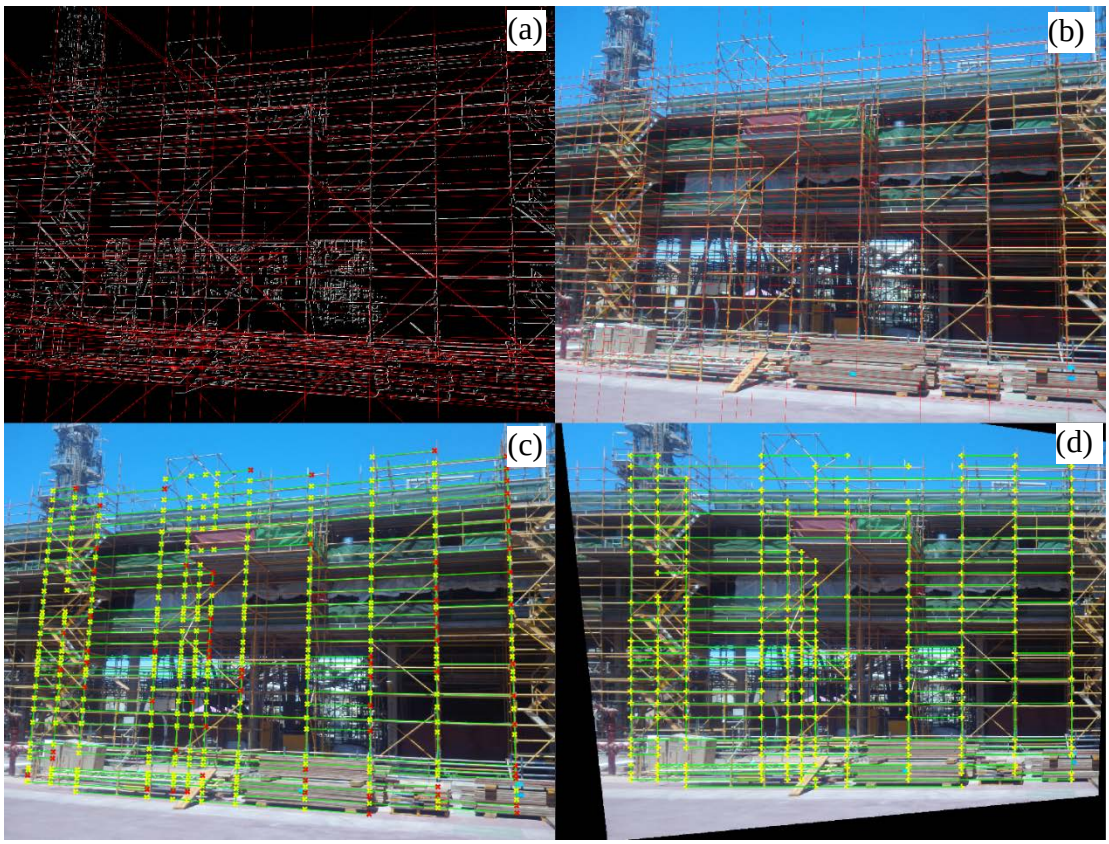

Figure 6.25. Another example of scaffolding detection from an image: (a) Line detection results; (b) detected lines after filtering; (c) detected scaffolding on the raw image; and (d) detected scaffolding beams on the rectified image

Both two tests show that the proposed linear feature analysis approach can detect horizontal and vertical scaffolding beams from a photo. It successfully recognises the scaffolding following the two basic assumptions and enables automatic scaffolding monitoring. However, the results also show that the approach suffers from the interference of other linear objects, such as thin pipes with similar orientation. For instance, in the second test the background is not fully covered by the debris containment netting and there are many other pipes included. Given the similar

appearance and spatial layout between the scaffolding and other pipes, it is difficult to remove these errors.

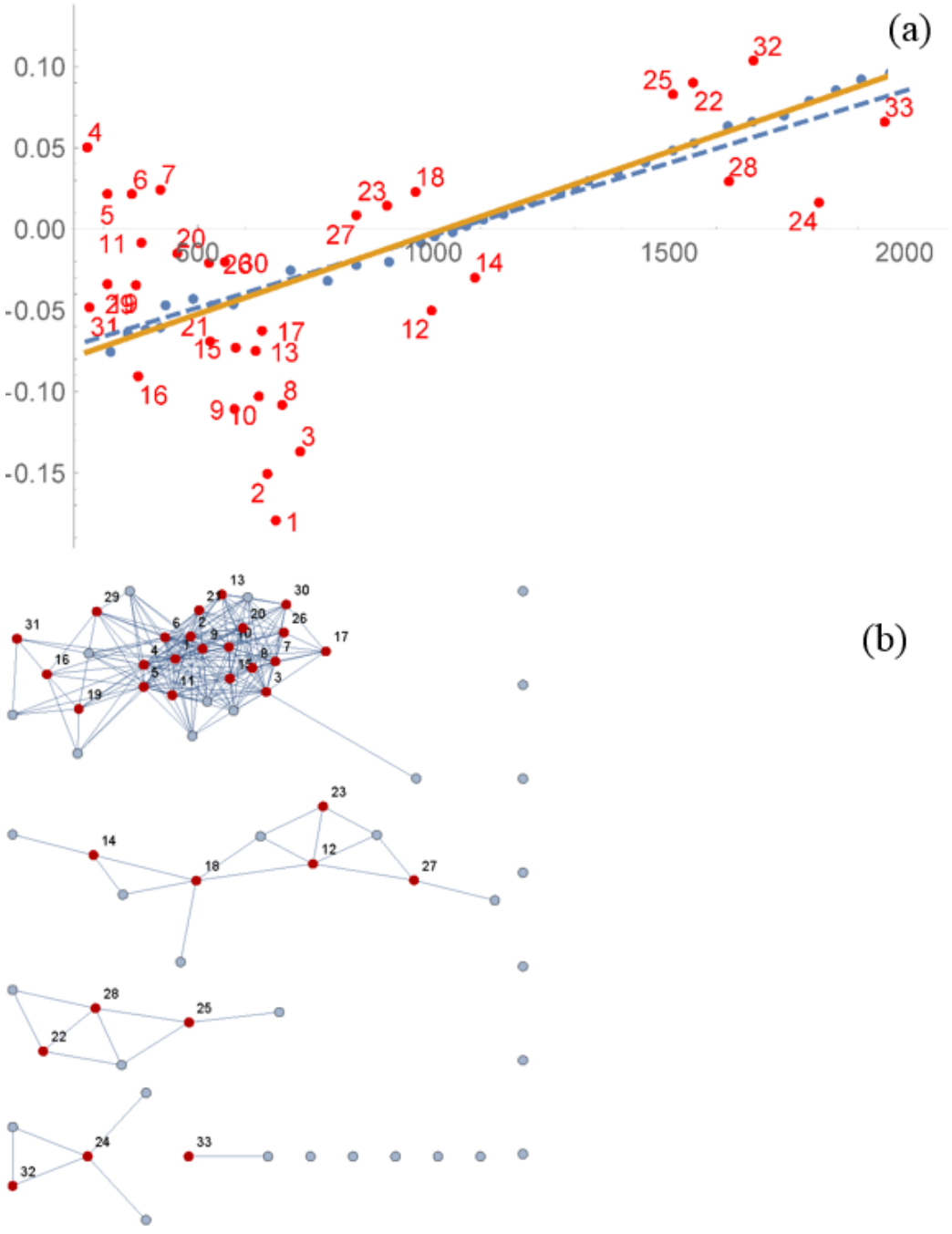

Figure 6.26. Filtering of line segments on the example shown in Figure 6.25: (a) linear regression results; and (b) the graph of the line intersection relationship

From the above examples, it can be shown that variance factors of the environment affect detection results to some extent. However, most of the detection results could be confirmed as valid by intuitive examinations. The detection system proposed is effective and robust in most cases.

6.4.3 Quantity estimation with template matching

In this subsection, an experiment of scaffolding quantity estimation with the template matching approach proposed in Section 6.2 is presented.

The input in this experiment is the scaffolding detection result with the depth-based approach from the preceding subsection, as shown in Figure 6.21(b). As can be seen, the scaffolding in this photo is composed of the same kind of repetitive pattern as displayed in Figure 6.15. A binary image representing this template was created and is displayed in Figure 6.27.

To estimate occurrences and locations of this pattern, we applied the GHT method to the segmented image as described in Section 6.2. The results of the GHT are displayed in Figure 6.30, where detected patterns are coloured and overlaid on the original image. The numbers shown on the image are used to label the order of the detection. As shown in the image, in this case, all 21 patterns were successfully detected.

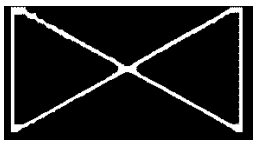

Figure 6.27. A template of a scaffolding pattern

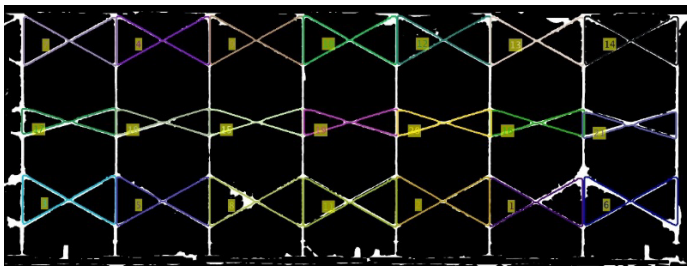

Figure 6.30. Results of scaffolding structure detection with the GHT method

In this experiment, the proposed template matching approach successfully detected all basic patterns in the photos of the scaffolding. The performance in this case demonstrated that our proposed template matching approach is feasible in monitoring scaffolding and its quantity during construction.

6.4.4 Quantity estimation with linear feature analysis

In this subsection, we describe experiments for estimating scaffolding quantities with a linear feature analysis.

Firstly, we used the results of the colour-based detection as shown in Figure 6.20(b) as inputs. The existence of various irregular patterns in this image makes it difficult to estimate quantities with template matching. The proposed linear feature analysis approach was used to further refine the detection results and to count the numbers of scaffolding elements.

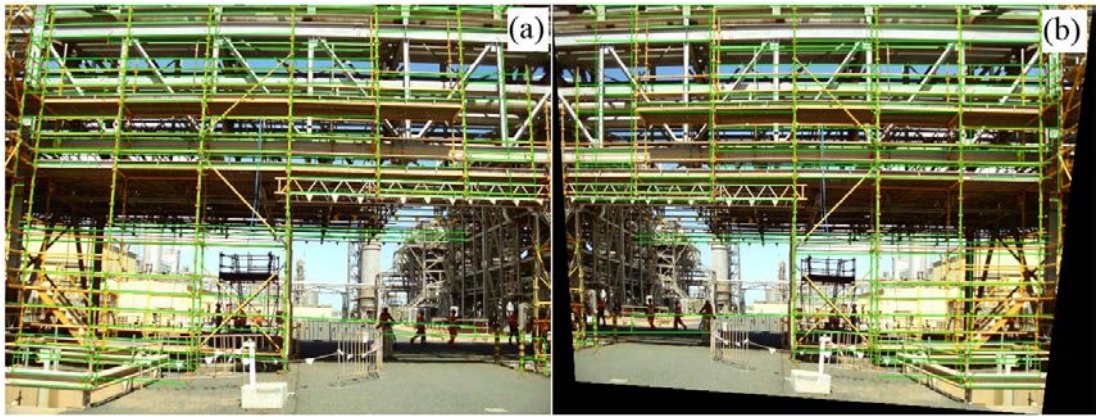

Figure 6.31. Results of scaffolding detection with colour-based segmentation and linear feature analysis: (a) results before the prospective rectification and (b) results after the prospective rectification

The final results of the detected scaffolding are displayed in Figure 6.31, where green lines represent finally detected scaffolding bracing tubes, and Figure 6.31 (a) and 6.31 (b) represent results before and after the rectification respectively. The results show that the bracing tubes are mainly marked correctly in the results even

though there are some missing or incorrect lines. The mistakes occur at the edges of the photo where there are similar linear features such as scaffolding tubes, while the missing lines around access ladders with short tubes differ from the main scaffolding. The figure also shows that the rectification process removes perspective distortions with horizontal and vertical beams perpendicular to each other.

Further evaluation of the results is presented with quantitative indicators in Table 6.2. Performance is evaluated with two indicators: precision and recall. The precision and recall are defined in Equation 5.10. In this case, true positive (TF) is the number of correctly detected scaffolding components; false positive (FP) represents the number of segments falsely detected as scaffolding components; and false negative (FN) denotes components that were not detected. As shown in the figure, the total number detected by the algorithm was 311 while that of manual inspection was 305. The precision and recall for beam detection are 83.92% and 85.57% respectively.

Table 6.2. Precision and recall of scaffolding beams detection with colour-based segmentation and linear feature analysis

Images	**Number of beams**		**Performance**	
	Manual	Automatic	Precision (%)	Recall (%)
Figure 6.31(a)	305	311	83.92	85.57

In addition, the results of the scaffolding detection were used with the linear feature analysis as input. The four key indicators, including the number of beams, length of beams, number of joints and number of basic rectangle patterns were calculated based on the proposed approach. The visual results of the two examples are displayed in Figure 6.32 and Figure 6.33 respectively, where the left image shows the automated detection result and the right one shows the manual inspection result. The numbers in the left image denote the sequence of the manual inspection.

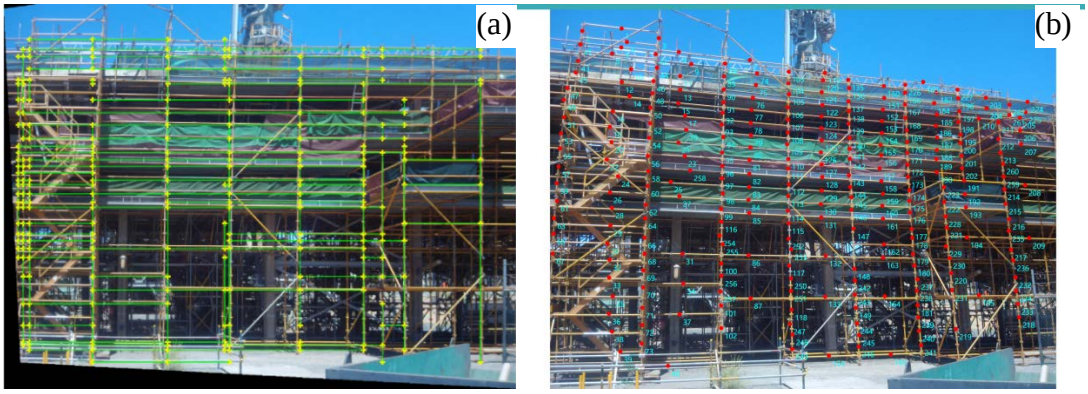

Figure 6.32. Quantity estimation of the scaffolding in Image 1

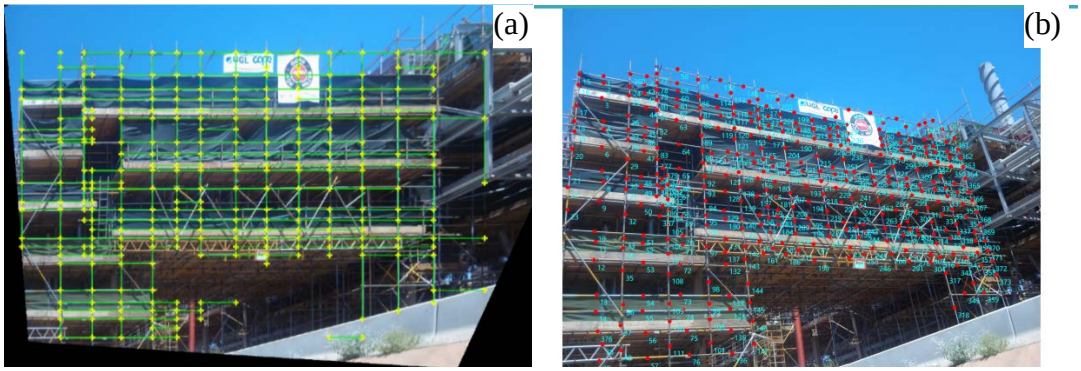

Figure 6.33. Quantity estimation of the scaffolding in Image 2

The detailed quantity evaluation results of the above two photos are present in Table 6.3 and Table 6.4. The results of the developed approach are compared against those of a manual inspection. As can be seen in the table, the precision and recall rates of scaffolding tubes estimation for Image 1 are 80.65 % and 81.74% respectively, while those for Image 2 are 88.56% and 85.48% respectively. Table 6.4 shows the results of other quantity indicators with numbers of manual inspections compared, including numbers of joints and rectangle patterns. In addition, the total length of scaffolding beams is also presented in pixels.

Table 6.3. Performance of the developed approach for beam counting

Images	**Number of beams**		**Performance**	
	Manual	Automatic	Precision (%)	Recall (%)
Image 1	367	372	80.65	81.74
Image 2	489	472	88.56	85.48

The results demonstrate that the linear feature analysis method can be used to estimate the quantity of the scaffolding. The estimation accuracy relies on the input image of the scaffolding detection. Similar to the first test shown in Figure 6.31, while most scaffolding beams are correctly detected from photos, most errors occur in areas where the scaffolding shows a different pattern, including tower ladders. The reason is that scaffolding ladders include not only horizontal and vertical tubes but also inclined components which enable workers' access. This conflicts with the assumption that only horizontal and vertical tubes exist. In addition, scaffolding ladders do not usually lie on the same physical plane as other scaffolding tubes, causing errors in line detection and perspective rectification.

Table 6.4. A comparison between the developed approach and the manual process

Components	**Number of Joints**		**Number of Rectangles**		**Total Length (in Pixels)**	
Methods	Manual	Automatic	Manual	Automatic	Horizontal	Vertical
Sample 1	227	252	132	118	48896	15605
Sample 2	256	288	204	183	28567	14069

6.5 Summary

In this chapter, to effectively monitor scaffolding construction and dismantling process, a hierarchical approach for monitoring scaffolding with cameras in LNG plants was presented. The proposed approach used different scaffolding detection and quantity estimation methods in different scenarios. For the scaffolding that was different to the background in colour, a colour-based segmentation method was used

to detect the scaffolding from a photo. For the scaffolding that showed a depth difference from other objects, a depth map of the scene was produced and the scaffolding was detected based on depth differences. For other complex scenarios, a linear feature analysis approach was developed, based on the assumption that scaffolding is mainly composed of horizontal and vertical tubes lying on the same physical plane. After the scaffolding detection, two methods of estimating scaffolding quantity were developed. One method was pattern matching with a predefined library, and this is suitable for scaffolding composed of regular patterns. The other was linear feature clustering which is suitable for scaffolding with irregular patterns.

A prototype was also developed to implement the developed approaches. The prototype system was tested in different environments, including building and LNG facilities. With both recall and precision rates over 80%, the results show that the approach is feasible and effective in detecting scaffolding and estimating its quantity. The developed approach is promising for promoting efficient and effective monitoring of scaffolding activities.

7 SUMMARY AND CONCLUSION

Australia has seen significant developments in the LNG industry, with huge investments into numerous projects and great increases in production capacity. Despite the LNG boom in past decades, the LNG industry in Australia suffers from big challenges, including low productivity and high cost during plant construction and operation phases. This research investigated methods for monitoring LNG facilities with vision-based techniques to offer timely feedback of as-built information and progress status.

7.1 Summary

It was proposed that more semantic information could be automatically retrieved from vision data using novel methods, and that retrieved as-built information could be used to support construction management tasks and increase productivity in a LNG plant. To achieve this, a framework for monitoring LNG facilities with vision data was proposed. The research focused on two types of structures: one being permanent components in LNG facilities, and the other temporary scaffolding supporting construction and maintenance of LNG plants. Both vision-based

collection methods for the two kinds of structures and related semantic interpretation approaches were investigated.

Current vision-based techniques and the wild deployment of vision sensors in the construction industry enable efficient as-built data collection, which is promising for the monitoring of construction projects. This research studied the performance of the collection of 3D data with vision techniques in a LNG plant. Given the structural complexity and safety requirements of LNG plants, this research focused on 3D reconstruction with commercial intrinsically safe cameras. In addition, the 3D reconstruction of scaffolding widely used in LNG projects was also studied. The research showed that the resolution of intrinsically safe cameras has an impact on the accuracy and completeness of 3D modelling of both LNG components and scaffolding. While laser scanning is more accurate and efficient, it requires a shutdown of LNG operations, which leads to low productivity and high costs. Thus, laser scanning could be used during construction and shutdown maintenance phases, while for the daily operation phase, intrinsically safe cameras are more cost-effective in the LNG industry. The research presented the performance of 3D reconstruction with intrinsically safe cameras and it can serve as a reference for other users to decide whether it is applicable to other management tasks.

To enable the automatic as-built modelling and progress monitoring of LNG facilities, this research studied point cloud interpretation methods using automatic classification and recognition methods. The research implemented a semantic segmentation method based on the Markov Random Field model, which enables a class-level interpretation. Before using MRF, the research first trained a SVM classifier for various classes of LNG facilities and then produced an initial classification results. Point cloud supervoxelisation was integrated to enhance the robustness and reduce computation costs. Quantitative results were presented by testing the proposed method with different datasets. The results showed high recall and precision rates for most classes. Although point clouds from a terrestrial laser scanner were used in this test, the approach may also be combined with LIDAR (Light Detection and Ranging) sensors on UAV, which enables a rapid and

autonomous 3D mapping of construction sites. This can save a great deal of mapping time and labour costs.

Further object-level interpretation was achieved by matching classes with a pre-defined object library using local feature matching methods. A high recognition rate was achieved on different industrial components in the presented experiments. Object-level recognition makes fully automated 3D geometric modelling of industrial facilities possible by automating the labelling of different objects, which are currently interactively labelled by users. The developed methods can contribute to a more efficient processing of large volume of vision data on-site and to some extent tackle the challenge of data overload in construction management.

Existing scaffolding monitoring processes are mainly manual and subjective. To enable effective and automatic scaffolding monitoring in LNG plants, this research proposed a hierarchical approach towards monitoring the scaffolding erection and dismantling process by using photos. Depending on the types of scaffolding and the characteristics of LNG plants, different methods were developed to detect scaffolding tubes from photos and to estimate the quantity of the scaffolding in various scenarios. These developed methods included colour-based, depth-based scaffolding segmentation methods and linear feature analysis detection method. A prototype system was also developed to integrate and test the proposed methods. The experiments on various scenarios presented a quantitative evaluation of the proposed methods and showed accuracy and feasibility in scaffolding quantity estimation. The developed prototype system offers novel choices for current construction projects involving scaffolding. The system is expected to improve the productivity of many management tasks, including material management and progress control, and to be superior to conventional manual approaches.

7.2 Contribution

The developed approaches on point cloud and photo interpretations enable efficient as-built information modelling in industrial plants, and they contribute to many

aspects of on-site monitoring. The following paragraphs list the contributions of this research:

Guidelines for 3D data collection in intrinsically safe plants

By conducting a series of tests, the research demonstrates the feasibility of 3D scaffolding reconstruction from photos. By evaluating the performance of using intrinsically safe cameras in 3D reconstruction, this research gives guidelines to the user of the equipment, along with method selection when performing similar tasks. The conducted tests provide an example of how to produce 3D data on LNG plants by integrating laser scanning and photos from intrinsically safe cameras.

Semantic information modelling methods of LNG plants

The proposed segmentation and recognition methods are promising in the automatic as-built modelling of industrial plants. By combining automatic geometric modelling methods, it is possible to produce a fully automated modelling process. In addition, the recognised class and industrial component information can enrich geometric models to create semantically rich as-built models. By reducing time and cost for the manual interpretation of laser scanning data, the implemented methods enable a rapid mapping of construction sites. They can also be used to offer timely information updating for off-site visualisation and decision-making.

Automatic algorithms for detecting scaffolding from photos

The developed algorithms can detect scaffolding beams from photos automatically and effectively in several scenarios. The novel algorithms can be practically applied to LNG plants by considering the existing commercial scaffolding types and actual on-site environments. To the best of our knowledge, this research is the first work devoted to automatically detecting and monitoring scaffolding with photos. The research also contributes a novel approach to estimating actual scaffolding quantities from photos. By estimating the quantities of the erected scaffolding, the novel approach allows an automatic and effective usage of the data from on-site cameras, and enables remote scaffolding erection monitoring. In addition, the developed

algorithms can be applied to other tasks related to scaffolding, such as vision-based scaffolding inspection and health monitoring.

A prototype system for automatic scaffolding erection monitoring based on vision data

The developed prototype system is promising in promoting the overall processes of scaffolding construction management. By producing timely and objective information on construction progress, the system can contribute to removing current manual daily inspection by on-site workers, and effectively reduces labour costs and safety risks related to on-site inspections. This contribution promotes the monitoring of temporary structures in construction sites. The real time feedback from the prototype system can also be used to analyse working patterns and discover waste during scaffolding construction. In addition, as-used quantity reports from the system can help improve productivity in scaffolding erection, dismantling and supply chain management.

7.3 Limitations and future work

Despite these contributions, this research also has some limitations and gives rise to some questions in need of further research. For the semantic interpretation, this research proposed methods for semantic classification and further object-level identification in LNG plants. It is also possible to create the as-built information with a higher Level-of-Details (LODs) which could meet the demands of greater construction management tasks. By integrating vision-based inspection methods, the recognised object could be enriched with as-is conditions information. LNG facilities suffer from severe corrosion issues and need effective approaches to monitor the health of structures. Vision-based inspection is promising for those inspection tasks, especially in the LNG industry. In addition, material classification using photos could enable operation-level progress monitoring.

This research achieved an overall accuracy of 70% in scaffolding monitoring. According to feedback from industry partners, the accuracy is sufficient for some management applications. However, to satisfy the requirements of more complex

management, a higher level of accuracy is required. Further improvements to the scaffolding detection algorithm could be another future research direction. Rather than detection based on merely a single photo, periodic photos at the same location and scaffolding detection methods with change detection between time-lapsed photos could be potential choices.

In this research, we used both laser scanning data and photos as inputs. However, they can be used individually as inputs for different tasks. Point clouds are semantically segmented based only on 3D features without considering corresponding photos. Photos are either directly processed for scaffolding detection or used to produce a 3D point cloud. Integrated segmentation of point clouds and photos is believed to produce more robust and accurate results as there are additional features available.

This research only focuses on the technical part of vision-based monitoring in LNG plants, which includes data collection techniques and approaches for automatic visual data interpretation. Further research is also necessary on how to incorporate the developed approaches into current workflows of a LNG plant construction or maintenance projects, or how to design a novel workflow efficiently using the proposed monitoring approaches. For example, an integration effort could be devoted to scaffolding monitoring methods with BIM platforms to enable efficient management of scaffolding construction and dismantling tasks. Even though the benefits of automatic scaffolding monitoring with vision-based techniques are obvious with regard to saving labour costs and improving monitoring efficiency, it is still worth quantitatively comparing the proposed novel approach against the existing methods so as to clarify its influence on productivity. In addition, rather than the LNG industry, it would also be useful to investigate applications of the developed approaches in other areas, such as building and infrastructure construction and maintenance.

8 References

Abeid, J. & Arditi, D., 2002. Time-lapse digital photography applied to project management, *Journal of Construction Engineering and Management*, 128(6), pp. 530-535.

Abudayyeh, O. Y., 1997. A multimedia construction delay management system, *Microcomputers in Civil Engineering*, 12(3), pp. 183-192.

Agarwal, A., Jawahar, C. V. & Narayanan, P.J., 2005. A survey of planar homography estimation techniques. *Tech. Rep. IIIT/TR/2005/12*, (3), pp.1–25. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.102.321 [Accessed June 7, 2017].

Agarwal, S. & Snavely, N., 2009. Building rome in a day. … *Vision, 2009 IEEE …*, pp.105–112. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5459148 [Accessed September 29, 2013].

Agisoft., 2015. Agisoft PhotoScan. Available at: http://www.agisoft.com/ [Accessed April 3, 2017].

Aijazi, A.K., Checchin, P. and Trassoudaine, L., 2013. Segmentation based classification of 3D urban point clouds: A super-voxel based approach with evaluation. *Remote Sensing*, 5(4), pp.1624-1650.

Anil, E.B., Tang, P., Akinci, B. and Huber, D., 2013. Deviation analysis method for the assessment of the quality of the as-is Building Information Models generated from point cloud data. *Automation in Construction*, 35, pp.507-516.

Anon, 2010. E57 Data Examples. Available at: http://www.libe57.org/data.html [Accessed April 12, 2017].

Armeni, I., Sax, S., Zamir, A.R. and Savarese, S., 2017. Joint 2D-3D-Semantic Data for Indoor Scene Understanding. arXiv preprint arXiv:1702.01105.

Barazzetti, L., 2016. Parametric as-built model generation of complex shapes from point clouds. *Advanced Engineering Informatics*, 30(3), pp.298–311.

Bariya, P. & Nishino, K., 2010. Scale-hierarchical 3D object recognition in cluttered scenes. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, pp. 1657–1664. Available at: http://ieeexplore.ieee.org/document/5539774/ [Accessed June 18, 2017].

BARTEC PIXAVI, 2017. Intrinsically Safe Camera. Available at: https://www.pixavi.com/product/mobile-devices/intrinsically-safe-camera/ [Accessed May 23, 2017].

Bayramoglu, N. & Alatan, A.A., 2015. Comparison of 3D local and global descriptors for similarity retrieval of range data. *Neurocomputing*, 184, pp.13–27. Available at: http://www.sciencedirect.com/science/article/pii/S092523121501841X [Accessed June 18, 2017].

Besl, P.J. & McKay, N.D., 1992. A Method for Registration of 3-D Shapes. *SPIE*, 1611, pp.586–606. Available at: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=981454 [Accessed September 18, 2016].

Bhatla, A. et al., 2012. Evaluation of accuracy of as-built 3D modeling from photos taken by handheld digital cameras. *Automation in Construction*, 28, pp.116–127.

Bosché, F., 2010. Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. *Advanced engineering informatics*. Available at: http://www.sciencedirect.com/science/article/pii/S1474034609000482 [Accessed July 8, 2016].

Bosche, F., Haas, C.T. and Akinci, B., 2009. Automated recognition of 3D CAD objects in site laser scans for project 3D status visualization and performance control. *Journal of Computing in Civil Engineering*, 23(6), pp.311-318.

Bosché, F. et al., 2015. The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction*, 49, pp.201–213.

Bosché, F., Guillemet, A., Turkan, Y., Haas, C.T. and Haas, R., 2013. Tracking the built status of MEP works: Assessing the value of a Scan-vs-BIM system. *Journal of Computing in Civil Engineering*, 28(4), p.05014004.

Brilakis, I.K., Soibelman, L. and Shinagawa, Y., 2006. Construction site image retrieval based on material cluster recognition. *Advanced Engineering Informatics*, 20(4), pp.443-452.

Bügler, M. et al., 2014. A comprehensive methodology for vision-based progress and activity estimation of excavation processes for productivity assessment. *Proceedings of the*. Available at: http://www.cms.bgu.tum.de/publications/2014_Buegler_EG-ICE.pdf [Accessed July 10, 2016].

Castellani, U. et al., 2008. Sparse points matching by combining 3D mesh saliency with statistical descriptors. In *Computer Graphics Forum*. pp. 643–652. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8659.2008.01162.x/full [Accessed June 18, 2017].

Chai, J., Wu, C., et al., 2016. A VISION-BASED APPROACH FOR SCAFFOLDING MONITORING. In *Proceedings of the 16th International Conference on Construction Applications of Virtual Reality*. Hong Kong, pp. 506–513.

Chai, J., Chi, H.-L., et al., 2016. Automatic as-built modeling for concurrent progress tracking of plant construction based on laser scanning. *Concurrent Engineering*, 24(4), pp.369–380. Available at: http://journals.sagepub.com/doi/10.1177/1063293X16670449 [Accessed April 11, 2017].

Chai, J. et al., 2017. Reference tag supported RFID tracking using robust support vector regression and Kalman filter. *Advanced Engineering Informatics*, 32, pp.1–10. Available at: http://www.sciencedirect.com/science/article/pii/S1474034616301367 [Accessed April 11, 2017].

Chai, J., Chi, H. & Wang, X., 2015. A Novel Automatic Process for Construction Progress Tracking Based on Laser Scanning for Industrial Plants. *International Conference on Cooperative Design*. Available at: http://link.springer.com/chapter/10.1007/978-3-319-24132-6_32 [Accessed July 12, 2016].

Chang, C.-C. & Lin, C.-J., 2011. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), p.27:1-27:27. Available at: http://dl.acm.org/citation.cfm?doid=1961189.1961199.

Chi, H.L. et al., 2016. The Development of a BIM-enabled Inspection Management System for Maintenance Diagnoses of Oil and Gas Plants. Available at: http://search.proquest.com/docview/1823095396?pq-origsite=gscholar [Accessed April 11, 2017].

Cimpoi, M., Maji, S. & Vedaldi, A., 2015. Deep filter banks for texture recognition and segmentation. In *2015 IEEE Conference on Computer Vision and Pattern*

Recognition (CVPR). IEEE, pp. 3828–3836. Available at: http://ieeexplore.ieee.org/document/7299007/ [Accessed April 12, 2017].

Dai, F. et al., 2013. Comparison of Image-Based and Time-of-Flight-Based Technologies for Three-Dimensional Reconstruction of Infrastructure. *Journal of Construction Engineering and Management*, 139(1), pp.69–79. Available at: http://ascelibrary.org/doi/10.1061/%28ASCE%29CO.1943-7862.0000565 [Accessed April 11, 2017].

Dall'Asta, E. & Roncella, R., 2014. A comparison of semiglobal and local dense matching algorithms for surface reconstruction. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. pp. 187–194. Available at: http://search.proquest.com/docview/1757013612?pq-origsite=gscholar [Accessed April 20, 2017].

Delong, A. et al., 2012. Fast approximate energy minimization with label costs. *International Journal of Computer Vision*, 96(1), pp.1–27.

Dimitrov, A. & Golparvar-Fard, M., 2015. Segmentation of building point cloud models including detailed architectural/structural features and MEP systems. *Automation in Construction*, 51, pp.32–45.

Ding, K. et al., 2016. RFID-Enabled Physical Object Tracking in Process Flow Based on an Enhanced Graphical Deduction Modeling Method. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp.1–13. Available at: http://ieeexplore.ieee.org/document/7466820/ [Accessed April 11, 2017].

Du, J. & Teng, H., 2007. 3D laser scanning and GPS technology for landslide earthwork volume estimation. *Automation in Construction*. Available at: http://www.sciencedirect.com/science/article/pii/S0926580506001191 [Accessed July 8, 2016].

Duda, R.O. & Hart, P.E., 1972. Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1), pp.11–15. Available at: http://portal.acm.org/citation.cfm?doid=361237.361242 [Accessed June 7, 2017].

Economidou, M. et al., 2011. Europe ' S Buildings Under the Microscope: A country-by-country review of the energy performance of buildings, Bruxelles, Belgium.

El-Omari, S. & Moselhi, O., 2008. Integrating 3D laser scanning and photogrammetry for progress measurement of construction work. *Automation in construction*. Available at: http://www.sciencedirect.com/science/article/pii/S0926580508000836 [Accessed July 8, 2016].

El-Omari, S. & Moselhi, O., 2008. Integrating 3D laser scanning and photogrammetry for progress measurement of construction work. *Automation in Construction*, 18(1), pp.1–9. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0926580508000836 [Accessed June 17, 2017].

El-Omari, S. & Moselhi, O., 2011. Integrating automated data acquisition technologies for progress reporting of construction projects. *Automation in Construction*, 20(6), pp.699–705. Available at: http://www.sciencedirect.com/science/article/pii/S0926580510002098 [Accessed April 11, 2017].

Fischler, M. a. & Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), pp.381–395. Available at: http://dl.acm.org/citation.cfm?id=358692 [Accessed November 21, 2016].

Funkhouser, T. et al., 2003. A search engine for 3D models. *ACM Transactions on Graphics*, 22(1), pp.83–105. Available at: http://portal.acm.org/citation.cfm?doid=588272.588279 [Accessed June 18, 2017].

Galland, O. et al., 2016. Application of open-source photogrammetric software MicMac for monitoring surface deformation in laboratory models. *Journal of Geophysical Research: Solid Earth*, 121(4), pp.2852–2872. Available at: http://doi.wiley.com/10.1002/2015JB012564 [Accessed April 3, 2017].

Ganesan, P., Rajini, V. & Rajkumar, R.I., 2010. Segmentation and edge detection of color images using CIELAB color space and edge detectors. In *INTERACT-2010*. IEEE, pp. 393–397. Available at: http://ieeexplore.ieee.org/document/5706186/ [Accessed April 27, 2017].

Gikas, V., 2012. Three-dimensional laser scanning for geometry documentation and construction management of highway tunnels during excavation. *Sensors*. Available at: http://www.mdpi.com/1424-8220/12/8/11249/htm [Accessed July 8, 2016].

Golovinskiy, A., Kim, V.G. and Funkhouser, T., 2009, September. Shape-based recognition of 3D point clouds in urban environments. In *Computer Vision, 2009 IEEE 12th International Conference on* (pp. 2154-2161). IEEE.

Golparvar-Fard, M., Pena-Mora, F. & Savarese, S., 2009. D4AR - a 4 dimensional augmented reality model for automation construction progress monitoring data collection, processing and communication. *Journal of Information Technology in Construction*, 14(June), pp.129–153. Available at: https://pdfs.semanticscholar.org/338e/04a889d11a6c6d46827dada3e67dda1833ce.pdf [Accessed September 18, 2016].

Golparvar-Fard, M., Peña-Mora, F. and Savarese, S., 2012. Automated progress monitoring using unordered daily construction photographs and IFC-based building information models. *Journal of Computing in Civil Engineering*, 29(1), p.04014025.

Gong, J., Yu, Y. & Roda, A., 2014. 3D Laser Scanning for Quality Control and Assurance in Bridge Deck Construction. Available at: https://trid.trb.org/view.aspx?id=1342958 [Accessed July 8, 2016].

Guo, Y. et al., 2014. 3D object recognition in cluttered scenes with local surface features: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*,

36(11), pp.2270–2287. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6787078 [Accessed September 16, 2016].

Hackmann, G. et al., 2014. Cyber-physical codesign of distributed structural health monitoring with wireless sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 25(1), pp.63–72. Available at: http://ieeexplore.ieee.org/abstract/document/6463394/ [Accessed April 11, 2017].

Han, K.K. & Golparvar-Fard, M., 2015. Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs. *Automation in Construction*, 53, pp.44–57.

Hartley, R. & Zisserman, A., 2004. *Multiple View Geometry in Computer Vision*, Available at: http://ebooks.cambridge.org/ref/id/CBO9780511811685.

Hmida, H. et al., 2013. Knowledge base approach for 3d objects detection in point clouds using 3d processing and specialists knowledge. *arXiv preprint arXiv:1301.4991*. Available at: http://arxiv.org/abs/1301.4991 [Accessed July 11, 2016].

Huang, J. & You, S., 2014. Segmentation and matching: Towards a robust object detection system. In *2014 IEEE Winter Conference on Applications of Computer Vision, WACV 2014*. pp. 325–332. Available at: http://ieeexplore.ieee.org/abstract/document/6836082/ [Accessed June 28, 2017].

Iqbal, M., Naeem, M., Anpalagan, A., Ahmed, A. and Azam, M., 2015. Wireless sensor network optimization: multi-objective paradigm. *Sensors*, 15(7), pp.17572-17620.

Kato, Z. & Pong, T.C., 2006. A Markov random field image segmentation model for color textured images. *Image and Vision Computing*, 24(10), pp.1103–1114. Available at: http://www.sciencedirect.com/science/article/pii/S0262885606001223 [Accessed April 28, 2017].

Khoury, H.M. & Kamat, V.R., 2009. Evaluation of position tracking technologies for user localization in indoor construction environments. *Automation in Construction*, 18(4), pp.444–457.

Kim, C., Kim, B. & Kim, H., 2013. 4D CAD model updating using image processing-based construction progress monitoring. *Automation in Construction*. Available at: http://www.sciencedirect.com/science/article/pii/S0926580513000332 [Accessed July 8, 2016].

Kim, K. & Teizer, J., 2014. Automatic design and planning of scaffolding systems using building information modeling. *Advanced Engineering Informatics*, 28(1), pp.66–80. Available at: http://www.sciencedirect.com/science/article/pii/S1474034613000979 [Accessed April 11, 2017].

Klein, L., Li, N. & Becerik-Gerber, B., 2012. Image-based verification of as-built documentation of operational buildings. *Automation in Construction*, 21, pp.161–171.

Knopp, J. et al., 2010. Hough transform and 3D SURF for robust three dimensional classification. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Berlin, Heidelberg, pp. 589–602. Available at: http://link.springer.com/10.1007/978-3-642-15567-3_43 [Accessed June 18, 2017].

Koppula, H.S. et al., 2011. Semantic Labeling of 3D Point Clouds for Indoor Scenes. *Neural Information Processing Systems*, pp.1–9. Available at: http://papers.nips.cc/paper/4226-semantic-labeling-of-3d-point-clouds-for-indoor-scenes [Accessed November 18, 2016].

Kršák, B. et al., 2016. Use of low-cost UAV photogrammetry to analyze the accuracy of a digital elevation model in a case study. *Measurement*, 91, pp.276–287. Available at: http://www.sciencedirect.com/science/article/pii/S0263224116301749 [Accessed April 11, 2017].

Lehtomäki, M. et al., 2016. Object Classification and Recognition From Mobile Laser Scanning Point Clouds in a Road Environment. *IEEE Transactions on Geoscience and Remote Sensing*, 54(2), pp.1226–1239. Available at: http://ieeexplore.ieee.org/document/7287763/ [Accessed November 18, 2016].

Lu, Y. & Rasmussen, C., 2012. Simplified markov random fields for efficient semantic labeling of 3D point clouds. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 2690–2697. Available at: http://ieeexplore.ieee.org/document/6386039/ [Accessed November 18, 2016].

Luo, X., O'Brien, W.J. & Julien, C.L., 2011. Comparative evaluation of Received Signal-Strength Index (RSSI) based indoor localization techniques for construction jobsites. *Advanced Engineering Informatics*, 25(2), pp.355–363.

Maria, F. Di et al., 2016. Quality assessment for recycling aggregates from construction and demolition waste: An image-based approach for particle size estimation. *Waste Management*. Available at: http://www.sciencedirect.com/science/article/pii/S0956053X15302397 [Accessed July 8, 2016].

Marks, E., Cheng, T. & Teizer, J., 2013. Laser scanning for safe equipment design that increases operator visibility by measuring blind spots. *Journal of Construction Engineering*. Available at: http://ascelibrary.org/doi/abs/10.1061/(ASCE)CO.1943-7862.0000690 [Accessed July 8, 2016].

Marzouk, M.M. & Zaher, M.M., 2015. Tracking construction projects progress using mobile hand-held devices.

Mian, A.S., Bennamoun, M. and Owens, R., 2006. Three-dimensional model-based object recognition and segmentation in cluttered scenes. *IEEE transactions on pattern analysis and machine intelligence*, 28(10), pp.1584-1601.

Mitra, N.J., Nguyen, A. & Guibas, L., 2004. Estimating Surface Normals in Noisy Point Cloud Data. *International Journal of Computational Geometry and*

Applications, 14(4–5), pp.261–276. Available at: http://dl.acm.org/citation.cfm?id=777840 [Accessed April 27, 2017].

Moon, S., Choi, B. & Yang, B., 2011. USN-Based Data Acquisition for Increasing Safety in the Concrete Formwork Operation. *Journal of Computing in Civil Engineering*. Available at: http://ascelibrary.org/doi/abs/10.1061/(ASCE)CP.1943-5487.0000132 [Accessed April 11, 2017].

Do Nascimento, E.R. et al., 2013. On the development of a robust, fast and lightweight keypoint descriptor. *Neurocomputing*, 120, pp.141–155. Available at: http://www.sciencedirect.com/science/article/pii/S092523121300307X [Accessed June 18, 2017].

Olsen, M., Kuester, F. & Chang, B., 2009. Terrestrial laser scanning-based structural damage assessment. *Journal of Computing in*. Available at: http://ascelibrary.org/doi/abs/10.1061/(ASCE)CP.1943-5487.0000028 [Accessed July 8, 2016].

Omar, T. & Nehdi, M.L., 2016. Data acquisition technologies for construction progress tracking. *Automation in Construction*, 70, pp.143–155. Available at: http://www.sciencedirect.com/science/article/pii/S0926580516301376 [Accessed April 11, 2017].

Papon, J. et al., 2013. Voxel cloud connectivity segmentation - Supervoxels for point clouds. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. pp. 2027–2034.

Park, C. et al., 2013. A framework for proactive construction defect management using BIM, augmented reality and ontology-based data collection template. *Automation in Construction*. Available at: http://www.sciencedirect.com/science/article/pii/S0926580512001598 [Accessed July 8, 2016].

Park, C. & Kim, H., 2013. A framework for construction safety management and visualization system. *Automation in Construction*. Available at:

http://www.sciencedirect.com/science/article/pii/S0926580512001616 [Accessed July 8, 2016].

Park, H. et al., 2007. A new approach for health monitoring of structures: terrestrial laser scanning. *Computer-Aided Civil and*. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8667.2006.00466.x/full [Accessed July 8, 2016].

Park, J., Kim, K. & Cho, Y.K., 2016. Framework of Automated Construction-Safety Monitoring Using Cloud-Enabled BIM and BLE Mobile Tracking Sensors. *Journal of Construction Engineering and Management*, p.5016019. Available at: http://ascelibrary.org/doi/10.1061/%28ASCE%29CO.1943-7862.0001223 [Accessed November 9, 2016].

Park, M.W., Makhmalbaf, A. and Brilakis, I., 2011. Comparative study of vision tracking methods for tracking of construction site resources. *Automation in Construction*, 20(7), pp.905-915.

Permuter, H., Francos, J. & Jermyn, I.H., 2003. Gaussian mixture models of texture and colour for image database retrieval. In *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03)*. IEEE, p. III-569-72. Available at: http://ieeexplore.ieee.org/document/1199538/ [Accessed April 18, 2017].

Poudel, D., 2013. 3D Point Cloud Segmentation. In *2013 6th IEEE Conference on Robotics,*. pp. 1–6. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6758588 [Accessed November 21, 2016].

Pradhananga, N. & Teizer, J., 2013. Automatic spatio-temporal analysis of construction site equipment operations using GPS data. *Automation in Construction*, 29, pp.107–122. Available at: http://www.sciencedirect.com/science/article/pii/S0926580512001537 [Accessed April 14, 2017].

Pu, S. & Vosselman, G., 2009. Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(6), pp.575–584.

Qi, C.R., Su, H., Mo, K. and Guibas, L.J., 2016. Pointnet: Deep learning on point sets for 3d classification and segmentation. arXiv preprint arXiv:1612.00593.

Rose, K.H., 2013. A Guide to the Project Management Body of Knowledge (PMBOK® Guide)-Fifth Edition. *Project Management Journal*, 44(3), pp.e1–e1. Available at: http://doi.wiley.com/10.1002/pmj.21345 [Accessed April 11, 2017].

Ruiz-Sarmiento, J.R., Galindo, C. and Gonzalez-Jimenez, J., 2017. A survey on learning approaches for Undirected Graphical Models. Application to scene object recognition. *International Journal of Approximate Reasoning*, 83, pp.434-451.

Rukkanchanunt, T. (2014). *Model-based metrics for assessing completeness and accuracy for 3D image-based reconstruction methods*. Master of Science. University of Illinois at Urbana-Champaign.

Rusu, R. et al., 2009. Fast geometric point labeling using conditional random fields. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*. IEEE, pp. 7–12. Available at: http://ieeexplore.ieee.org/document/5354763/ [Accessed November 21, 2016].

Rusu, R.B. et al., 2008. Aligning Point Cloud Views using Persistent Feature Histograms. *Intelligent Robots and*. Available at: http://ieeexplore.ieee.org/abstract/document/4650967/ [Accessed June 28, 2017].

Rusu, R.B., Blodow, N. & Beetz, M., 2009. Fast Point Feature Histograms (FPFH) for 3D registration. In *IEEE International Conference on Robotics and Automation*. IEEE, pp. 3212–3217. Available at: http://ieeexplore.ieee.org/document/5152473/ [Accessed April 27, 2017].

Rusu, R.B. & Cousins, S., 2011. 3D is here: point cloud library. *IEEE International Conference on Robotics and Automation*, pp.1–4. Available at:

http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5980567 [Accessed September 18, 2016].

Schindler, K., 2012. An overview and comparison of smooth labeling methods for land-cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, 50(11 PART1), pp.4534–4545. Available at: http://ieeexplore.ieee.org/abstract/document/6198886/ [Accessed June 14, 2017].

Schnabel, R., Wahl, R. & Klein, R., 2007. Efficient RANSAC for point-cloud shape detection. *Computer Graphics Forum*, 26(2), pp.214–226. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8659.2007.01016.x/full [Accessed November 21, 2016].

Scovanner, P., Ali, S. & Shah, M., 2007. A 3-Dimensional SIFT Descriptor and Its Application to Action Recognition. *Proceedings of the 15th international conference on Multimedia*, pp.357–360. Available at: http://dl.acm.org/citation.cfm?id=1291311 [Accessed September 18, 2016].

Sezgin, M., 2004. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*. Available at: http://electronicimaging.spiedigitallibrary.org/article.aspx?articleid=1098183 [Accessed May 2, 2017].

Shapovalov, R., Velizhev, A. & Barinova, O., 2010. Non-Associative Markov Networks for 3D Point Cloud Classification. *ISPRS*, XXXVIII-3A, pp.103–108.

Shrivakshan, G.T., 2012. A Comparison of various Edge Detection Techniques used in Image Processing. , 9(5), pp.269–276. Available at: https://www.researchgate.net/publication/303142762_A_Comparison_of_various_Edge_Detection_Techniques_used_in_Image_Processing [Accessed June 7, 2017].

Son, H. et al., 2014. Classification of major construction materials in construction environments using ensemble classifiers. *Advanced Engineering Informatics*, 28(1), pp.1–10. Available at:

http://www.sciencedirect.com/science/article/pii/S1474034613000773 [Accessed April 11, 2017].

Son, S., Park, H. & Lee, K., 2002. Automated laser scanning system for reverse engineering and inspection. *International Journal of Machine Tools and*. Available at: http://www.sciencedirect.com/science/article/pii/S0890695502000305 [Accessed July 8, 2016].

Su, Y., Hashash, Y. & Liu, L., 2006. Integration of construction as-built data via laser scanning with geotechnical monitoring of urban excavation. *Journal of construction engineering*. Available at: http://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-9364(2006)132:12(1234) [Accessed July 8, 2016].

Szeliski, R. et al., 2008. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(6), pp.1068–1080. Available at: http://ieeexplore.ieee.org/document/4420084/ [Accessed April 19, 2017].

Tang, P., Huber, D. and Akinci, B., 2010. Characterization of laser scanners and algorithms for detecting flatness defects on concrete surfaces. *Journal of Computing in Civil Engineering*, 25(1), pp.31-42.

Tang, P. et al., 2010. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in construction*. Available at: http://www.sciencedirect.com/science/article/pii/S0926580510000907 [Accessed July 8, 2016].

Turkan, Y., Bosche, F., Haas, C.T. and Haas, R., 2012. Automated progress tracking using 4D schedule and 3D sensing technologies. *Automation in Construction*, 22, pp.414-421.

Volk, R., Stengel, J. & Schultmann, F., 2014. Building Information Modeling (BIM) for existing buildings—Literature review and future needs. *Automation in*

construction. Available at: http://www.sciencedirect.com/science/article/pii/S092658051300191X [Accessed July 8, 2016].

Vosselman, G. and Dijkman, S., 2001. 3D building model reconstruction from point clouds and ground plans. *International archives of photogrammetry remote sensing and spatial information sciences*, 34(3/W4), pp.37-44.

Wang, C., Cho, Y.K. and Kim, C., 2015. Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Automation in Construction*, 56, pp.1-13.

Wang, J. et al., 2016. BIM-Enabled Collaborative Scaffolding Scoping and Design. In Springer, Cham, pp. 43–50. Available at: http://link.springer.com/10.1007/978-3-319-46771-9_6 [Accessed April 11, 2017].

Weinmann, M. et al., 2015. CONTEXTUAL CLASSIFICATION OF POINT CLOUD DATA BY EXPLOITING INDIVIDUAL 3D NEIGBOURHOODS. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3/W4, pp.271–278. Available at: http://search.proquest.com/openview/68bddd48cbbd9eea6a63612619a638d3/1?pq-origsite=gscholar&cbl=2037681 [Accessed November 18, 2016].

Weinmann, M., Jutzi, B. & Mallet, C., 2013. Feature relevance assessment for the semantic interpretation of 3D point cloud data. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W2(1), pp.313–318. Available at: http://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/II-5-W2/313/2013/isprsannals-II-5-W2-313-2013.html [Accessed November 18, 2016].

Wong, J.K.W. et al., 2014. Virtual Prototyping for Construction Site Co2 Emissions and Hazard Detection. *International Journal of Advanced Robotic Systems*, 11(8), p.130. Available at: http://journals.sagepub.com/doi/10.5772/58439 [Accessed April 14, 2017].

Wu, X., Zhu, X., Wu, G.Q. and Ding, W., 2014. Data mining with big data. *IEEE transactions on knowledge and data engineering*, 26(1), pp.97-107.

Xiong, X., Adan, A., Akinci, B. and Huber, D., 2013. Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in Construction*, 31, pp.325-337.

Xiong, X. and Huber, D., 2010, September. Using Context to Create Semantic 3D Models of Indoor Environments. In *Proceedings of the British Machine Vision Conference* (pp. 1-11).

Xu, Y. et al., 2016. Classification of photogrammetric point clouds of scaffolds for construction site monitoring using subspace clustering and PCA. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*. pp. 725–732. Available at: http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLI-B3/725/2016/isprs-archives-XLI-B3-725-2016.pdf [Accessed July 10, 2017].

Xu, Y. et al., 2015. Reconstruction of Scaffolding Components from Photogrammetric Point Clouds of a Construction Site. *ISPRS Annals*. Available at: http://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/II-3-W5/401/2015/isprsannals-II-3-W5-401-2015.pdf [Accessed April 11, 2017].

Yang, J. et al., 2015. Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future. *Advanced Engineering*. Available at: http://www.sciencedirect.com/science/article/pii/S1474034615000233 [Accessed July 8, 2016].

Yang, J., Shi, Z. & Wu, Z., 2016. Vision-based action recognition of construction workers using dense trajectories. *Advanced Engineering Informatics*, 30(3), pp.327–336.

Yang, Q., 2015. Stereo matching using tree filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(4), pp.834–846. Available at: http://ieeexplore.ieee.org/document/6888475/ [Accessed April 19, 2017].

Yuan, X., Anumba, C.J. & Parfitt, M.K., 2016. Cyber-physical systems for temporary structure monitoring. *Automation in Construction*, 66, pp.1–14. Available at: http://www.sciencedirect.com/science/article/pii/S0926580516300279 [Accessed April 11, 2017].

Zhang, L., Zhang, L. & Du, B., 2016. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2), pp.22–40. Available at: http://ieeexplore.ieee.org/document/7486259/ [Accessed April 12, 2017].

Zhang, S., Teizer, J. & Pradhananga, N., 2015. Workforce location tracking to model, visualize and analyze workspace requirements in building information models for construction safety planning. *in Construction*. Available at: http://www.sciencedirect.com/science/article/pii/S0926580515002022 [Accessed July 8, 2016].

Zhou, G. & Yi, T., 2013. Recent developments on wireless sensor networks technology for bridge health monitoring. *Mathematical Problems in Engineering*. Available at: https://www.hindawi.com/journals/mpe/2013/947867/abs/ [Accessed April 11, 2017].

Zhu, Z. and Brilakis, I., 2010. Concrete column recognition in images and videos. *Journal of computing in civil engineering*, 24(6), pp.478-487.

Zhu, Z. and Brilakis, I., 2010, June. Intelligent Learning and Detection of Concrete Regions in Construction Site Images. In *Proceedings of the 6th International Conference on Innovation in Architecture, Engineering and Construction (AEC)* (pp. 9-11).

Zhu, Z., German, S. and Brilakis, I., 2011. Visual retrieval of concrete crack properties for automated post-earthquake structural safety evaluation. *Automation in Construction*, 20(7), pp.874-883.

Zollmann, S. et al., 2014. Augmented reality for construction site monitoring and documentation. In *Proceedings of the IEEE*. pp. 137–154. Available at: http://ieeexplore.ieee.org/document/6705607/ [Accessed November 9, 2016].

APPENDIX

LIST OF FIGURES

LIST OF TABLES