Integration of Laser Scanning and Photogrammetry for Heritage Documentation

Von der Fakultät
Luft- und Raumfahrttechnik und Geodäsie
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

vorgelegt von

Yahya Alshawabkeh
aus Amman-Jordan

Hauptberichter:	Prof. Dr.-Ing. Dieter Fritsch
Mitberichter:	Prof. Dr. Sultan Maani

Tag der mündlichen Prüfung: 6. Dezember 2006

Institut für Photogrammetrie der Universität Stuttgart
2006

Abstract

The use of 3D laser scanner in documenting heritage sites has increased significantly over the past few years. This is mainly due to advances of such systems, which allow for the fast and reliable generation of millions of 3D points. Despite the considerable progress of these approaches, the highest possible degree in efficiency and flexibility of data collection will be possible, if other techniques are combined during data processing. Within the research the potential of combining terrestrial laser scanning and close range photogrammetry for the documentation of heritage sites is discussed. Besides improving the geometry of the model, the integration aims on supporting the visual quality of the linear features like edges and cracks in the historical scenes. Although the laser scanner gives very rich surface details, it does not provide sufficient data to construct outlines for all surface features of the scanned object, even though they are clearly defined in the reality. In our approach, information on edges and linear surface features is based on the analysis of the images. For this purpose an integrated segmentation process based on image data will support the extraction of object geometry information from the laser data. The approach applies image based semi-automated techniques in order to bridge gaps in the laser scanner data and add new details, which are required to build more realistic perception of the scene volume.

Cultural heritage applications frequently require data collection by terrestrial laser scanning in very complex structural environments. Thus, compared to similar applications in industrial environments, this requires more complex processing to generate geometric models of sufficient realism. In addition to geometric data collection, texture mapping is particular important in the area of cultural heritage to have more complete documentation. Photo-realistic texturing can for example add information about the structure condition, which is not present in the 3D model such as decay of the material. Additionally, color image information is also indispensable for features like frescos and mosaics. In addition to that, texture mapping considered as a requirement application for visualization and animation purposes. For this reason, some commercial 3D systems already provide model-registered color texture by capturing the RGB values of each LIDAR point using a camera already integrated in the system. However, these images frequently are not sufficient for high quality texturing, which is desired for documentation, since the ideal conditions for taking the images may not coincide with those for laser scanning In addition, laser scanning from many viewpoints, as it is required to capture complex structures, is still relatively time consuming. For outdoor applications these large time differences will result in varying light conditions and changing shadows, thus the recorded images will have different radiometric properties. Such problems may disturb the appearance of the resulting textured model. So it is therefore more useful to acquire geometry and texture by two independent processes and allow for an image collection at optimal position and time for texturing. This is especially true for the high requirements of realistic documentation of heritage sites. Thus, the thesis presents an approach for projective texture mapping from photographs onto triangulated surfaces from 3D laser scanning. By these means, the effort to generate photo-realistic models of complex shaped objects can be reduced considerably. The images are collected from multiple viewpoints, which do not necessarily correspond to the viewpoints of LIDAR data collection. In order to handle the resulting problem of occlusions, the visibility of the model areas in the respective images has to be established. For this purpose, a new visible surface algorithm has been developed, the algorithm works in both image and object space and efficiently detects ambient, back-face and view frustum occlusions. Occluding polygons are labelled and separated with their connectivity to texture them recursively using the optimal of the available images until the final textured model is produced. After this visibility processing, colour values will be correctly assigned from the photograph to the visible polygons. In order to gain a high quality texture, lens distortion and colour corrections are applied during processing.

The quality of the registration process, which aligns the laser scanner data with the imagery, is a crucial factor for the aspired combined processing. This registration can realized if correspondence coordinates are available in both systems. Since the accurate detection and measurement of point correspondences can be difficult especially for the point clouds from laser scanning, straight lines are measured between the image and the laser data as corresponding elements. The accuracy of the transformation depends on the accuracy with which the features have been extracted from the scans. For providing stable features of interest efficient segmentation algorithms have to be used to extract the features automatically for the co registration of data sets. In this thesis we have presented and discussed efficient edge detection algorithm that can detect the line features in both range and intensity images. In the proposed algorithm the distinguished points, which will comprise the edges, depend on the spatial analysis of the numerical description of the mean curvature values. The work was motivated by the fact that the optimality of edge detectors for range images has not been considered in the literature, some algorithms are limited to synthetic range images and will totally fail in the presence of noise, others which have been tested in real range images are complicated with large numbers of parame-

ters. Compared to known methods in literature, the proposed algorithm exhibits the following features: computational efficiency, high accuracy in the localization of the edge points, easy to implement, and Image noise does not degenerate its performance. Although the central task of the proposed edge detection algorithm is to reliable detect and locate edge points, a rich description of edge points give the ability to reliably detecting and characterizing the edge types as a crease and step edges, and then go further to classify the crease edges as concave or convex types. The algorithm was initially proposed for range image segmentation and has been extended to segment the intensity images with some improvements. The generality and robustness of the algorithm is illustrated on scene images with different available range sensors.

Zusammenfassung

Die Nutzung von 3D Laser-Scannern zur Dokumentation von Kulturdenkmälern hat in den vergangenen Jahren erheblich zugenommen. Dieses ist auf die Weiterentwicklung derartiger Systeme zurückzuführen, die eine schnelle und zuverlässige Generierung von Millionen 3D-Punkten ermöglichen. Trotz des beträchtlichen Fortschritts bei dieser Vorgehensweise werden höchste Leistungsfähigkeit und Flexibilität bei der Datenerfassung erst dann möglich, wenn sie bei der Datenauswertung mit anderen Techniken kombiniert werden. Die Arbeit beschäftigt sich mit dem Potenzial des kombinierten terrestrischen Laser-Scannings und der Nahbereichsphotogrammetrie für die Dokumentation von Kulturdenkmälern. Neben der Verbesserung des geometrischen Modells zielt die Integration auf die Verbesserung der visuellen Qualität linearer Merkmale wie Ecken und Kanten. Obwohl der Laser-Scanner sehr reichhaltige Oberflächendetails liefert, werden keine ausreichenden Daten zur Konstruktion der Konturen aller Oberflächenmerkmale des gescannten Objektes bereitgestellt, auch wenn sie in Wirklichkeit eindeutig definiert sind.

Bei dem entwickelten Ansatz werden Informationen über Kanten und linearen Oberflächenmerkmalen aus der Analyse von Distanz- und Intensitätsbildern bereitgestellt. Zu diesem Zweck wird ein integrierter Segmentierungsprozess für die Extraktion von objekt-geometrischen Informationen aus Laserdaten genutzt. Der Ansatz verwendet bildbezogene halbautomatische Techniken, um Lücken in den Laser-Scanner-Daten zu schließen und neue Details hinzu zufügen, die für den Aufbau realistischer Szenen benötigt werden.

Anwendungen im Rahmen der Dokumentation von Kulturdenkmälern erfordern die Datenerfassung mittels terrestrischem Laser-Scanning in sehr komplex strukturierten Umgebungen. Verglichen mit ähnlichen Anwendungen in Industrieumgebungen macht dies eine aufwendigere Verarbeitung nötig, um geometrische Modelle von ausreichend wirklichkeitsgetreuer Darstellung zu generieren. Neben der geometrischen Datenerfassung ist die Texturierung von Oberflächen für die lückenlose Dokumentation besonders wichtig. Photorealistische Texturen können zum Beispiel Informationen über die Strukturbeschaffenheiten, wie z. B. Materialverfall, hinzufügen, die in rein geometrischen 3D-Modellen nicht vorhanden sind. Zusätzlich sind Farbbildinformationen für Merkmale wie Fresken und Mosaiken unabdingbar. Ferner wird die Abbildung von Oberflächen für Visualisierungs- und Animationszwecke benötigt. Aus diesem Grund verfügen einige kommerzielle 3D-Systeme bereits über eine modellregistrierte Farbtextur, indem die RGB-Werte eines jeden LIDAR-Punktes durch eine im System integrierte Kamera erfasst werden. Diese Bilder sind jedoch häufig für eine qualitativ hochwertige Texturierung nicht ausreichend, da die idealen Vorraussetzungen für die Aufnahmen der Bilder nicht mit denen des Laser-Scannings übereinstimmen. Ferner ist das Laser-Scanning von mehreren Standpunkten aus wie es für die Erfassung komplexer Strukturen benötigt wird sehr zeitaufwendig. Bei Außenanwendungen resultiert dieses in verändernden Lichtbedingungen und wechselnden Schatten. Daher werden die aufgenommenen Bilder verschiedene radiometrische Eigenschaften aufweisen. Diese Probleme können das Erscheinungsbild des Texturmodells stören. Daher ist es sinnvoller, Geometrie und Textur durch zwei unabhängige Prozesse zu gewinnen und eine Datenerfassung in optimaler Position und Zeit für die Texturierung zu ermöglichen. Dies trifft besonders auf die hohen Anforderungen bei der Dokumentation von Kulturdenkmälern zu. Die vorliegende Arbeit stellt einen Ansatz für die projektive Texturabbildung von Photographien auf triangulierte Oberflächen aus 3D Laser-Scanning-Aufnahmen vor. Dadurch wird der Aufwand zur Generierung von photorealistischen Modellen komplexer Objekte beträchtlich reduziert. Die Bilderfassung erfolgt von mehreren Standpunkten aus, die nicht notwendigerweise mit den Standpunkten der LIDAR Datenerfassung übereinstimmen. Um das Problem der Verdeckung zu lösen, muss die Sichtbarkeit der Modellgebiete in den jeweiligen Aufnahmen gewährleistet sein. Zu diesem Zweck wurde ein neuer Algorithmus zur Bestimmung der sichtbaren Oberflächen entwickelt. Die im Bild verdeckten Polygone werden markiert, um sie in einem rekursiven Prozess durch die alternativen zur Verfügung stehenden Bilder zu texturieren.

Die Qualität des Registrierungsprozesses, der die Transformation zwischen Laser-Scanner-Daten und Bilddaten ermöglicht, ist ein äußerst wichtiger Faktor für die angestrebte kombinierte Prozessierung. Die Registrierung erfordert die Bestimmung korrespondierender Element in beiden Koordinatensystemen. Da die fehlerfreie Detektion und Messung identischer Punkte in Punktwolken schwierig ist, werden gerade Linien zwischen Bild und den Laserdaten als korrespondierende Element gemessen. Die Genauigkeit der Transformation hängt ab von der Genauigkeit, mit der diese Merkmale von den Scans extrahiert wurden. Um stabile Elemente bereitzustellen, müssen effiziente Segmentierungsalgorithmen für die automatische Extraktion von Merkmalen aus den Datensätzen verwendet werden. In dieser Dissertation wird ein Algorithmus zur Kantendetektion vorgestellt und diskutiert, der Linienmerkmale sowohl in Intensitäts- als auch Distanzbildern aufdeckt. In dem vorgestellten Algorithmus werden die gesuchten Kantenpunkte durch eine räumliche Analyse der mittleren Krümmung bestimmt. Die Arbeit wurde durch den Umstand motiviert, dass optimale Kantendetektoren für Distanzbilder noch nicht vorhanden sind. Einige der existierenden Algorithmen beschränken sich auf synthetische Distanzbilder während andere Verfahren, die auf reale Distanzbildern angewendet

werden, eine große Anzahl von Parametern besitzen und sehr komplex sind. Verglichen mit bekannten Methoden ist der vorgestellte Algorithmus effizient, genau, einfach zu implementieren und robust gegen Rauschen.

Der Algorithmus wurde ursprünglich für die Segmentierung von Distanzbildern entwickelt und anschließend mit einigen Verbesserungen zur Segmentierung von Intensitätsbildern erweitert. Die Allgemeingültigkeit und Robustheit des Algorithmus wird durch die Bearbeitung von Szenen nachgewiesen, die von unterschiedlichen Sensoren erfasst wurden.

Content

1 Introduction

The recovery and representation of 3-d geometric and photometric information of the real world is one of the most challenging and well-studied problems in computer visions as well as in digital photogrammetry. An example is the generation of highly realistic geometric models of historical architectures and heritage sites. The 3D architectural modeling archive is important from various points of view; recording of the monument/building is fundamental when reconstruction or restoration process is needed, for example in the case if they ever destroyed by fire, earth quick, war or erosion. It is also crucial for historical studies and archeological researches [Carbonell, 1989]. 3D models allow to visualize scenes from different viewpoints that may are diffecult in the real world due to size or surrounding objects. Additionally, it is considered as an efficient tool for virtual tourism and virtual museum [Georgopoulos et al, 2004]. There are many ways of creating a 3D model of the buildings, and many technologies can be employed to record the spatial and visual complexity of the real world environments. During the generation of these models, requirements such as high geometric accuracy, availability of all details, and photo-realism have to be met by the different approaches used for data collection [El-Hakim et al, 2002]. The fact that cultural heritage objects comprise a large variety of different nature, sizes, and complexity creates a challenge to find the suitable documentation tools and techniques for gathering 3D data of the architectures. In general the existing data collection tools can be grouped into two primary categories: Non-photographic and photographic techniques. Non-photographic or so-called graphic documentation techniques are based on conventional surveying to produce plans, elevations, and architectural details. The tools used in this technique range from the steel or fiberglass surveyor's tape to the optical theodolite or "total station" camera. The measuring is acceptable for gathering tens or even hundreds of coordinates on small artifacts. But it is slow and time consuming for modeling large architectural-scale objects, which requires thousands of measurements to create an accurate virtual representation. A second group of technologies has called photographical documentation techniques. These tools don't require contact with the object, offering clear advantages when working at landscape or monument scale. These camera tools include; projected light stripe digitizers, prototype automated stereo video photogrammetry systems, and an array of laser-based scanners; ranging from triangulation to phase shift and laser radar (LIDAR). Although these tools have their own limitations, they are generally much faster than contact systems, a key requirement for preparing complex 3D representations for virtualized reality. Choosing the suitable technique for metric surveying mainly depend on the object size and achievable representation of details. In additions, factors such the required accuracy, the needed resolution, and the accessibility of the object may influence the chosen of the optimal method [Boehler, 2005].

One well-accepted technique frequently applied in the context of heritage site documentation is close range photogrammetry. In the past decade, these traditional terrestrial approaches have also benefited from the fact that digital image collection is now feasible with of-the-shelf cameras. Thus the efficiency of photogrammetric data collection could be improved considerably by the integration of semi-automatic tools based on digital image processing. Recently, laser scanning has become a standard tool for 3D data collection for the generation of high quality 3D models of cultural heritage sites and historical building. These systems allow for the fast and reliable generation of millions of 3D points based on the run-time of reflected light pulses. This results in a very effective and dense measurement of surface geometry. Despite the considerable progress of these approaches, there are still some aspects that have to be considered comparing the two technologies:

- Acquisition time: The complete coverage of spatially complex objects like heritage sites can only be guaranteed, if data collection is realized from different viewpoints. Even though this is possible in most scenarios, problems can result from the facts that setting up and dismounting the complete laser system is relatively time consuming. In contrast to that, the effort and time to collect additional images with a standard digital camera can almost be neglected, but longer and sometimes-manual work needs to be done for extracting dense 3D measurements from stereoscopic pictures.
- Shadow and irregular surfaces: Laser scanning is able to capture data in areas where the photogrammetry technique could not, such as in the areas of shadows and the irregular surfaces. Comparatively with laser technique, image based modeling alone is difficult or even impractical for parts of surfaces contain irregular and unmarked geometrics details.
- Texture mapping: Laser scanner can provide black and white reflectance of the reflected laser beam. Some laser scanners have already an integrated camera with the system. The main task for the images is to support the semi automatic evaluation of the system such as visual detection of the tie points. However, these images

frequently are not sufficient for high quality texturing, which is desired for documentation. Thus, for the high requirements of realistic documentation of heritage sites, textured imagery should be collected in an independent process to allow for an image collection at optimal position and time for texturing [Alshawabkeh & Haala, 2005].

- Extraction of linear surface feature (i.e. edges, cracks, relief's and draw): Although the laser scanner gives very rich surface details, it does not provide sufficient data to construct the clearest outline of the object scanned [Borg & Cannataci, 2002]. Major edges could be modeled, but small items like the cracks, joints and reliefs could not. Typically, many cultural heritage applications show clearly that photogrammetry always leads to more complete and precise results when edges are sought, especially when those edges are mainly recognizable from texture contrast. But from another side, although digital photogrammetry is more accurate in outline rendition, it requires high-density correspondence networks for edge 3D modelling, and automated measurements often requires manual editing. Thus combined methods should be favored [Alshawabkeh & Haala, 2004].
- Depending on type of sensor, the resolution of laser scanner is generally limited. Object resolution is theoretically a function of the size of the laser spot itself on the object and the smallest possible increment of the angle between two adjacent points. On another hand, high accurate measurements could be achieved using high-resolution digital images.

From the above summary, it is abvious that there is no technique by it selfe can satisfy all the requiremmnets for high quality documnation of heritage sites. Thus the work in this thesis wants to underline the necessity to integrate the laser scanner and photogrammetry techniques to improve and optimize the geometry and visual quality of the collected 3D model of large outdoor structures. The main focus is the development of algorithms and tools for recovery of the geometric and photometric information from the collected range and intensity images.

1.1 Needs for a Combination of Images and Laser Data

This section describes in details the necessity for fusing laser and visuzal data in order to overcome the defects of either system. By using such hybrid approaches the effort to generate photo-realistic models of complex shaped objects can be reduced considerably.

1.1.1 High Quality Texture Mapping

Cultural heritage applications frequently require data collection by terrestrial laser scanning in very complex structural environments. Thus, compared to similar applications in industrial environments, this requires more complex processing to generate geometric models of sufficient realism. In addition to geometric data collection, texture mapping is particular important in the area of cultural heritage to have more complete documentation. Photo-realistic texturing can for example add information about the structure condition, which is not present in the 3D model such as decay of the material. Additionally, color image information is also indispensable for features like frescos and mosaics. In addition to that, texture mapping considered as a requirement application for visualization and animation purposes. For this reason, some commercial 3D systems already provide model-registered color texture by capturing the RGB values of each LIDAR point using a camera already integrated in the system. However, these images frequently are not sufficient for high quality texturing, which is desired for documentation, since the ideal conditions for taking the images may not coincide with those for laser scanning [El-Hakim et al, 2002]. In addition, laser scanning from many viewpoints, as it is required to capture complex structures, is still relatively time consuming. For outdoor applications these large time differences will result in varying light conditions and changing shadows, and the recorded images will have different radiometric properties. Such problems may disturb the appearance of the resulting textured model. So it is therefore more useful to acquire geometry and texture by two independent processes and allow for an image collection at optimal position and time for texturing.

In order to warp independent images onto the collected object surfaces using perspective or affine transformation projection, different approaches are available [Kada et al, 2003; Remondino et al, 2004]. However, such approaches can only be applied for simple objects with restricted number of surfaces with no occlusions. Otherwise, if direct projection transformation is applied without considering occlusions, the mapping between object geometry and image will be incorrect. Warping the image over the geometry associate each 3D point to a pixel in the color image, the tex-

ture pierces through the geometry and gets mapped onto all occluded polygons on the path of the projected ray. Thus, the geometry that is occluded in the image will receive incorrect texture coordinates instead of remaining in shadow. The occlusion problem can be avoided by manual texture extraction and mapping, however, such tedious task can take up to several days for good results. Fully and semi-automatic texture extraction and placement of the real scene have been presented in different works [Grammatikopulos et al, 2004; Sequeira & Goncalves, 2002; Neugebauer & Klein, 1999] but with more complex processing. Those approaches use automatic techniques for occlusion detection which could be calssified under two categories; image space approach such as Z buffer algorithm [Catmull, 1975] and object space approach such as painter's algorithm [Martens, 2003]. In general image space algorithm requires large memory and long computation time and show bad performance as the complexity of scene is increasing. On the other hand, object space visibility algorithms, which can be used alternatively, show high precision and good results in high complex scenes but they are computationally expensive and difficult to implement.

For this purpose the thesis will present an effective automatic approach to the problem of high-resolution photo-realistic texture mapping onto 3D complex model generated from range images. Our approach allows taking the images at different time from laser scanning and at whatever locations that will be the best for texturing. Through the approach we will introduce our visibility algorithm for occlusion detection and image fusing to have a realistic final model. The algorithm works in both image and object spaces. The main characteristics of this algorithm are the ability of detecting all types of occlusion problems; ambient, self and view frustum occlusions. In addition to that, it is flexible to allow the visibility detection of different resolution of geometric and 2D images.

1.1.2 Improving the Geometry and Visual Quality of the 3D Model

It was discussed in the first section that current laser scanners can produce large point clouds fast and reliably. But the resolution of this data is still be insufficient, especially if edges and linear surface features have to be collected. The visual appearance of these features is very hard to interpret. In the contrary, the digital photogrammetry is more accurate in outline rendition, especially if they are clearly defined in the reality. But image based modeling alone is difficult or even impractical for parts of surfaces, which contain irregular and unmarked geometrics details. Additionally, the identification of points to be measured, being manual or semi automatic, requires a long and tedious work, especially if a considerable number of points has to be captured. In addition to this problem, the complete coverage of spatially complex objects like heritage sites can only be guaranteed, if data collection is realized from different viewpoints. Even though this is possible in most scenarios, problems can result from the fact that setting up and dismounting the complete laser system is relatively time consuming. In contrast to that, the effort to collect additional images with a standard digital camera can almost be neglected. Compared to laser scanning there are fewer restrictions on the range of measurements during image collection, which simplifies the selection of different viewpoint in order to cover the complete structure of the object.

In our approach the integration helps to improve the geometry and visual quality of the collected 3D model. During data collection the information on edges and linear surface features like cracks is based on the analysis of the images, whereas information on object geometry is provided from the laser data. Additionally, areas, which are not accessible in the laser scanner data due to occlusions are added based on semiautomatic evaluation of the imagery. By these means, a complete 3D features for the scene can be generated with sufficient and clear details.

The quality of the registration process, which aligns the laser scanner data with the imagery, is a crucial factor for the combined processing. Since the accuracy of the transformation depends on the reliable extraction of primitives from the range and intensity images, efficient segmentation algorithms are considered as important task for a combined evaluation of this type of data.

1.2 Laser Data Segmentation and Pose Estimation

Registration, or alignment, of the local coordinate systems of the range and colour data sets is the first step in the fusion problem. To calculate the pose for arbitrary rotation and translation, we need to know the correspondences between features in the image and the geometry. However, the automatic extraction and matching of suitable primitives such as points, corners, and lines is extremely difficult especially in complex scenes [Ferencz, 2001]. An alternative way is the use of artificial targets like spheres or signals, which can be detected and identified more easily in both laser range data and intensity images [Yu et al, 2000]. However, this requires additional effort during data collection and

management. Additionally, time-consuming high-resolution scanning is required to accurately measure the position of the target center. Moreover, these artificial objects may occlude important parts of the intensity images, which are required for texture mapping [Lensch et al, 2000]. Alternatively, user supplied corresponding primitives like 3D points or lines and their 2D projections on the image plane can be used. Many solutions have been developed for pose estimation based on such 2D-3D correspondences. A number of traditional photogrammetric approaches use point matches, which allow for direct solutions based on three, four or six corresponding points [Haralick et al, 1994]. However, the perception of such point structures is limited within range data and can result an accuracy of manual measurement, which is not appropriate for registration [Liu & Stamos, 2005]. For this reason, the use of linear features is advantageous. From a practical point of view, line detection is more accurate and more reliable than points. Line matching is also more robust than point matching with respect to partial occlusions [Christy & Horaud, 1999]. Those features can also help to increase the reliability of the adjustment during spatial resection. In addition to the fact that linear features frequently occur within images of man made environments, it is easier to automatically extract such edge structures compared to a measurement of distinct points [Horaud et al, 1997, Lee & Habib, 2002]. In order to automatically provide stable features of interest, efficient segmentation algorithms are required. Most of these segmentation algorithms are based on range images instead of unordered 3D point clouds [Yu & Ferencz, 2001]. For such 2.5D raster grids neighborhood relations are available implicitly and tools from image processing can be adopted. Thus, the implementation of segmentation algorithms is simplified considerably.

Algorithms developed for the segmentation of intensity image have been discussed extensively in literature. Well known examples for the real time segmentation of intensity images are [Palmer et al, 1996; Canny, 1986]. In the other hand, ready-made solutions for range image segmentation are not available to a comparable extend [Gächter, 2005]. While in the past range data collection was mainly applied for industrial scenes captured at close distances, nowadays long-range laser scanners are available for many users. By these means detailed data sets of complex outdoor scenes are collected, which pose much more serious challenges for range image analysis than the traditional polyhedral world. The difficulties result from the fact that range data of natural scenes are relatively noisy. These measurement errors affect the approximation of the surfaces during segmentation. In addition, the natural scenes are complex since lots of individual objects or irregular surfaces occur. Some existing segmentation algorithms are limited to high quality range images and will fail in the presence of noise. Others are complicated and have a large numbers of parameters while generic and efficient edge detectors for range images are still missing. This was our motivation for the development of an edge detection algorithm for range images. By these means, stable features of interest are provided. They allow for an accurate alignment of the range and image data, which is required for the generation of high quality and photo-realistic 3D models.

1.3 Objectives of the Thesis

Within the thesis the potential of combining terrestrial laser scanning and close range photogrammetry for the documentation of heritage sites is discussed. In the thesis we will use each technique to its strength, using laser scanner for object models and augmenting the main object model with photogrammetry technique, especially in the areas of high details. Besides improving the geometry of the model, the integration aims on supporting the visual quality of the linear surface features and improve interpretation and detail in complex scan. Thus, the work is aimed towards achieving high quality photorealistic visualization of the generated 3D models. In this regard, one can classify the contributions of this thesis into the following major contributions:

- Developing of automatic multi-image photo texturing approach of complex 3D scene. For this purpose efficient algorithms addressing the image fusion and visibility are developed.
- Increasing the visual quality of the linear surface features (i.e. cracks, edges and reliefs) extracted from laser scanner model through photogram metric measures.
- Complete the 3D features of laser model using semiautomatic imagery approach.
- Developing a new edge detection algorithim for range image segmentation. The algorithim is used to provide stable features of interest toward automation of laser data and imagery co-registration.

1.4 Thesis Outline

The input to our pipeline, depicted in figure 1.1, is a set of range images and independent photographs. The developed segmentation algorithm is used to simplify the dense datasets and provide stable features of interest in both range and intensity images, which are used to recover the positions of the 2D cameras with respect to the geometric model. Then automatic hybrid approaches for linear surface extraction and high photo-realistic 3D model are presented. An overview of the thesis chapters is given in this section.

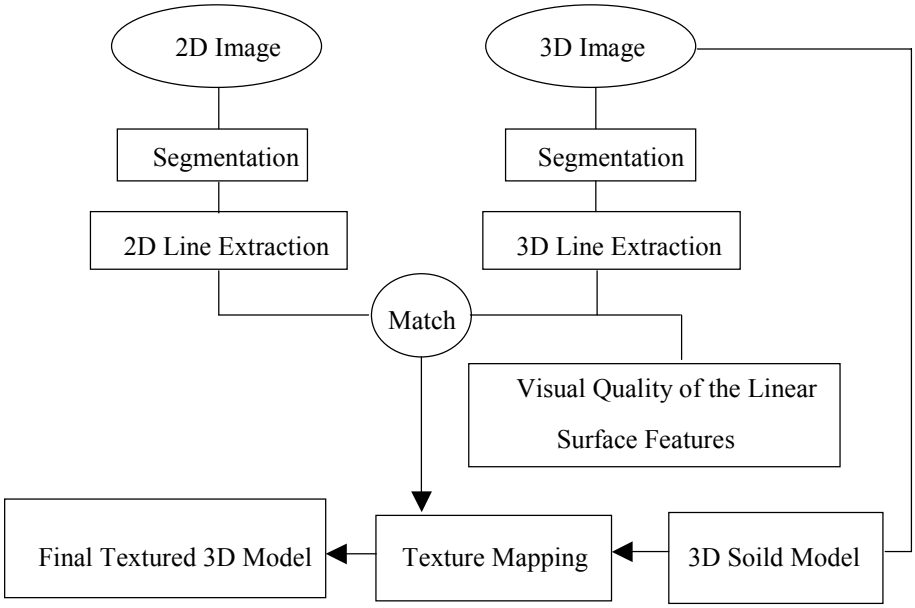


Figure 1.1: The workflow of the thesis

Chapter one: the introduction will discuss the problem statement, the motivation of the study, and the proposed methods for solving the tasks.

Chapter two: presents a survey covering previous works in the integration of laser scanner and photogrammetry technologies, and literatures in linear feature extraction, texture mapping and range image segmentation.

Chapter three: summarizes data acquisition and pre-processing in order to generate the 3d models for the treasury monument in ancient Petra city and the north theatre in ancient Jerash city using terrestrial laser scanner.

Chapter four: proposes new edge detection algorithms for range and intensity images, the chapter presents the mathematical properties of the algorithms, algorithms description and characteristics, and new edge-linking algorithm. In the experiments results, the generality and robustness of the algorithms are illustrated on scene images with different available range sensors.

Chapter five: In this chapter we will discuss an effective automatic approach to the problem of high-resolution photo-realistic texture mapping from separated images onto 3D complex model generated from range images. Through the approach we will introduce our visibility algorithm for occlusion detection and image fusing.

Chapter six: Presents a hybrid approach combining digital photogrammetry and laser scanner technologies in order to improve the visual quality of the linear surface features like edges, relief's and cracks in the historical scenes. The approach applies image based semi-automated techniques in order to bridge gaps in the laser scanner data, and add new details required to build more realistic perception of the scene volume.

Chapter seven: summarizes the contribution, draws our conclusions of this research work and indicates the direction for future researches.

2 Overview of 3D Construction Techniques

2.1 Introduction

This chapter reviews some of the recent developments techniques in the areas needed to build a complete photorealistic- 3D model object. In particular, we will present work in the areas of image-based modeling and range image modeling. In this literature those techniques are presented from different aspects and criteria, which are influcing the chain and the outcome of the 3D reconstruction of the object. Currently efforts are directed towards increasing the realistc and the quality of the produced 3D model by integration intensity and range images. Several approachs are discussed from different points of view.

2.2 Image Based Modelling

Image Based Modelling (IBM) also called photometric modelling is the technique of creating 3d models from photographs and then mapping the images onto the surfaces of the model. The geometry is created by identifying sets of common points from two or more source photos. the topological relationship between those points has to taken into consideration to form the edges and the surfaces of the object. Creating models directly from photographs has received increased interest in both computer vision and in computer graphics under the title of image-based modeling and rendering [Nishihara, 1984; Leveau & Faugeras, 1994]. Since real images are used as input, such an image-based system has an advantage in producing photorealistic renderings as output. However, these systems are only as strong as the underlying stereo algorithms. In such systems the photographs need to have a similar viewpoint for realible results. Because of this, current image-based techniques must use many closely spaced images, and in some cases employ significant amounts of user input for each image pair to supervise the stereo algorithm. Efforts to increase the level of automation became essential in order to widen the use of the technology. [Hartley & Zisserman, 2000; McMillan & Bishop, 1995] work toward automatic construction of graphical models of scenes when the input is a sequence of closely spaced 2-D images. [Neim, 1994] presented an algorithm for the robust and fast automatic construction of a 3D model of any real object using images from multiple views. The images are taken from a real object rotating in front of a stationary calibrated CCD TV camera. [Pollefeys et al, 2000] has used sequences of images as input and does pairwise dense matching between neighboring images in the sequence. The partial results are fused into one dense depth map, which is used for reconstructing a 3D surface. The system presented retrieves a 3D surface model from a sequence of images taken with off-the-shelf consumer cameras. The user acquires the images by freely moving the camera around the object. The obtained 3D model is a scaled version of the original object and the surface texture is obtained from the image sequence as well. The system uses full perspective cameras and does not require prior models nor calibration. The complete system combines state-of-the-art algorithms to solve the different sub problems: projective reconstruction, self-calibration and dense depth estimation.

The most known system in image modeling technique is the façade technique developed by [Debevec et al, 1996]. The method's main goal is the realistic creation of 3D models of architectures from small number of photographs. The basic geometric shape of the structure is first recovered interactively using models of polyhedral elements. In this step, the actual size of the elements and camera pose are captured assuming that the camera intrinsic parameters are known. The second step is an automated matching procedure, constrained by the now known basic model, to add geometric details. The approach proved to be effective in creating geometrically accurate and realistic models of architectures. For producing renderings, the approach presents view-dependent texture mapping, a method of compositing multiple views of a scene that better simulates geometric detail on basic models. The drawback is the high level of interaction, a preliminary model must be available and the geometry should not be too complex. [El-Hakim et al, 2002] follow a similar philosophy; replace basic shapes with a small number of seed points to achieve more flexibility and higher level of details. In addition, the camera poses and 3D coordinates are determined without any assumption of the shapes but instead by a full bundle adjustment, with or without self-calibration depending on the given configuration. This achieves higher geometric accuracy independent from the shape of the object. [Werner & Zisserman, 2002] proposed a fully automated Façade-like approach. They investigate a strategy for reconstruction of buildings from multiple (uncalibrated) images. In a similar manner to the Facade approach they first generate a coarse piecewise planar model of the principal scene planes and their delineations, and then use these facets to guide the search for indentations and protrusions such as windows and doors. However, unlike the Facade approach, which involves manual selection and

alignment of the geometric primitives, the strategy here is fully automatic. Since this is a fully automated approach, it requires feature detection, closely spaced images matching and camera pose estimation using projective geometry.

In general, for image modeling approaches, closely spaced images are required for automatic robust matching, but the alternative of improving stereo correspondence by using images taken from nearby locations has the disadvantage that computing depth becomes very sensitive to noise in image measurements. Thus, methods based on images alone are not able to capture highly detailed architectural environment [Grün et al, 2002]. In addition, three-dimensional measurement from images requires that interest points or edges be visible in the image, which is not always possible. Since images may be affected by the illumination or ambient light problems. For tourists where general visualization is enough, this approach may be adequate, but not for documentation requirements, which have been mentioned in the introduction chapter.

2.3 Range Image Modeling

Instead of the approach of using multiple images to reconstruct scene structure, an alternative technique is to use range-imaging sensors to directly measure depth to various points in the scene. In the last years 3D laser scanners have been used more and more in the 3D reconstruction and reproduction, because precise 3D data can be measured in very short time interval. The scan appears as a consecutive series of columns of sequential points that quickly form a three dimensional image. There are two main types of range sensors. The first is triangulation-based that consists of a transmitting device, sending a laser beam at a defined, incrementally changed angle from one end of mechanical base. A CCD camera at the other end of the base detects the returned laser spot. The triangulation system principle is depicted in figure 2.1.a. The accuracy of measurements depends on the triangle base relative to its height. Since, for practical reasons, the triangle base is rather short, triangulation-based systems have a limited range of less than 10 meters [Boehler et al, 2001]. The second sensors type is based on the time-of-flight. Those measure the delay between emission and detection of the light reflected by the surface. The principle of this type of sensors in recording the 3D information is illustrated in figure 2.1.b. The main advantage is that long-range imaging is possible, allowing for recording large structures. How ever, the achievable accuracy is relatively low and commonly resulting in centimeter range.

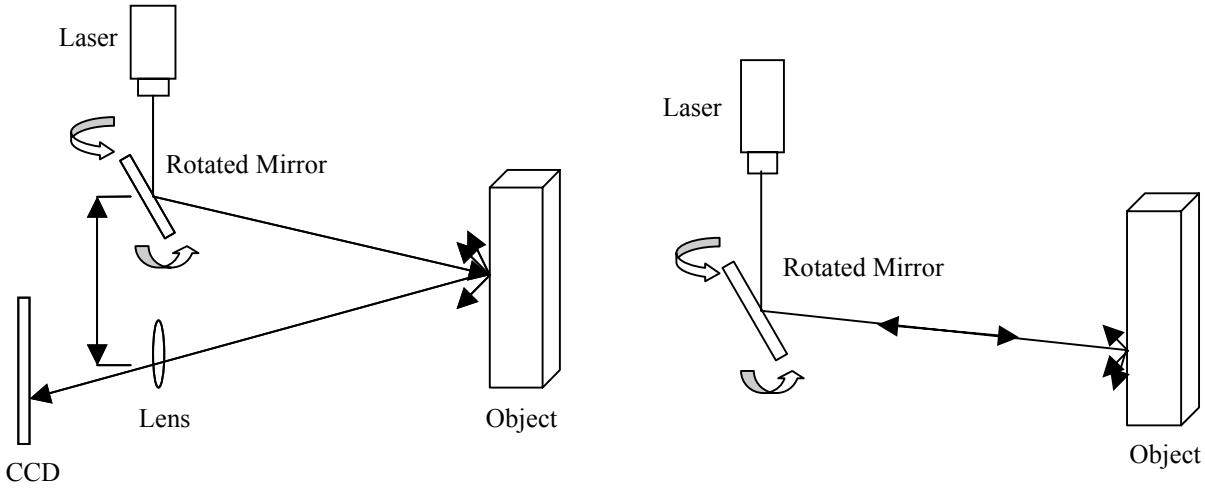


Figure 2.1: The main types of range sensors; (a) Triangulation principle and (b) Time of flight principle

Even though the previous adavantages, range cameras may produce erroneous results when sudden changes in surface height occur on surfaces with large reflectance variations and on rough surfaces. [El-Hakim et al, 1995] explained this fact by showing that the laser ray projected onto a scene is not infinitesimal in diameter. Hence, when the laser spot crosses a transition, the imaged spot will lose most of its symmetry, and the centroid of the light distribution for the laser spot will shift to indicate erroneous range data as can be depicted in the figure 2.2 below.

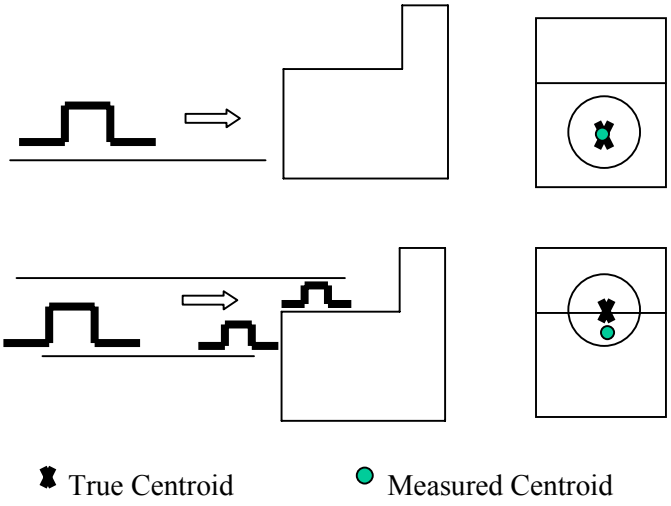


Figure 2.2: The behavior of laser spot on a surface height transition

2.3.1 Range Images

Range image is a large matrix of distance measurements from the scanner sensor coordinate system to surface points on objects in scene. Each matrix element, or pixel, has a value representing depth value (z). Generally it is useful to view this large matrix as a digital surface, because the sensed value at each pixel can be considered as noisy samples of an analog surface z= f (x, y). The arrangement of the laser beam that sample the scene is shown in figure 2.3. The angle β between the successive rows and the angle α between the successive columns is constant. The laser beams are indexed on a 2D rectangular grid, where it is possible to define the connectivity relationship in order to generate an interpolated range image.

In surface based approach for range image segmentation, the digital surface considered as a noisy, quantized, discretely sampled version of piecewise-smooth graph surface and represented using the following equation:

$$\tilde{g}(i,j) = g(i,j) + \tilde{n}(x,y)$$

Where the noise $\tilde{n}(x,y)$ is added to the real input $g(i,j)$. The sampling is the spacing of discrete values in the domain of the image. This is usually described in terms of some sampling rate –how many samples are taken per unit of each dimension (dots per inch). In the other hand the quantization represent the spacing of discrete values in the range of an image (number of bits per pixel). An example is the black and white image has one bit per pixel. Sampling and quantization are independent, and each plays a significant role in the resulting signal. Range images encode at each point in the image the distance to the nearest object at that point, nearer objects are brighter, farther objects are darker. Figure 2.4 displays a range image for Al-khasneh monument in Petra city.

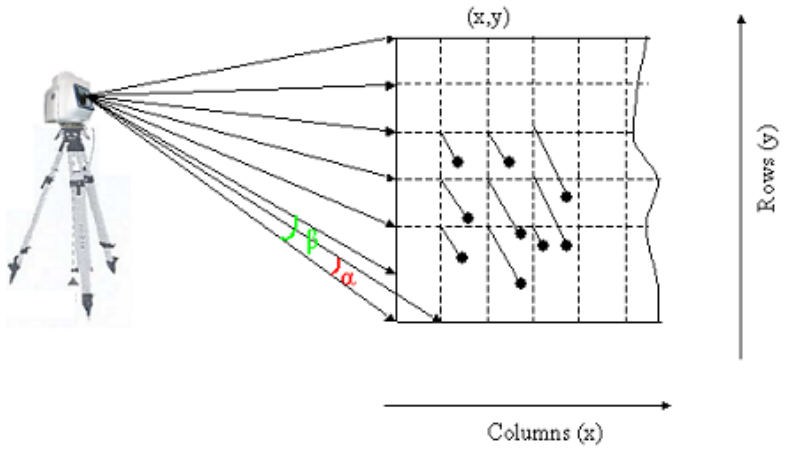


Figure 2.3: Range image representation

Figure 2.4: Range image for Al-khasneh /Petra city

2.3.2 3D Model Reconstruction

In general, the whole process for generating the 3D model can be subdivided into three main steps: meshing or triangulation process, registration and georeferencing process and merging process. Triangulation converts the set of the collected raw 3D points into a triangulated surface. The main aim of triangulation (meshing) process is to converts the point-based data into a visually more intuitive representation. Registration refers to computing a rigid transformation that brings the meshed points of one range image into alignment with the portions of a surface that is shares with another range image. A crucial step preceding the data fusion is the geometric alghinemt of the separate data set into a common reference frame. This is known as georeferencing. Conventionally, registration and georeferencing are carried out in a combined procedure. Registrartion of laserdata is performed by measuring at least three tie points in every scan. Georeferencing requires at least three known control points in all scans. The final step is the merging process; it is the process of creating a single and non-redundant surface representation from the overlapping meshes. The following sections discussed in details those pre-processing steps.

Triangulation Process

Once the the points are colleted in a single coordinate system, the next step is to reduce the large number of 3D points into a triangular mesh that preserves the geometric details and at the same time suitable for fast rendering. A common method for the reconstruction of a geometric figure given a set of sample points is the use of a triangulation algorithm to connect the points and find the convex hull. There are a number of triangulation algorithms that may be advocated, the more popular algorithm is [Watson, 1981] algorithm, which implements Delaunay triangulation. Given a set of points, the plane can be split in domains for which the first point is closest; the second point is closest, etc. Such a partition is called a Voronoi diagram. If one draws a line between any two points whose Voronoi domains touch, a set of triangles is obtained, known as the Delaunay triangulation. Generally, this triangulation is unique. One of its properties is that the out circle of every triangle does not contain any other data point, the circle is depicted in figure 2.5. using red-colored line The Delaunay triangulation can be constructed using an appropriate ″circle criterion″ in 2D dimension or the spherical circumscribed circle criterion in 3D dimension [Crippa et al, 1998]. [Bernardini & Bajaj, 1997] present a method for build an initial triangle mesh that interpolates all data points, the approach is based on (regularized) alpha-shapes, and is capable of automatically selecting an optimal alpha value and improving the resulting mesh in areas of insufficient sampling. [Amenta et al, 1999] presented a new volumetric reconstruction algorithm

based on 3D Voronoi diagram. A subset of Voronoi vertices (poles) is taken as an approximation of the medial axis of the object. To each pole a polar ball is assigned and the finite union of the polar balls gives the approximation of the objects surface.

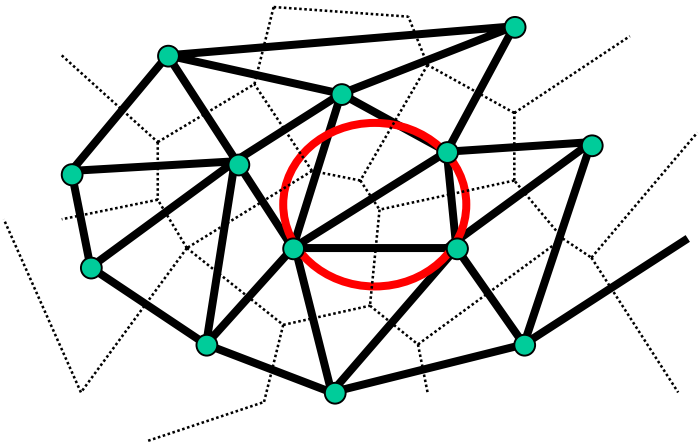


Figure 2.5: Delauney triangles (thick lines) and associated Voroni tessellations (thin lines), Delauncy circle criteria (red circle)

Registration of Range Data

A single range image is usually not sufficient to cover an object or a structure. The amount of necessary images depends on the shape of the object, amount of self-occlusion and obstacles, and the object size compared to the sensor range. Thus for object has a complex shape, a series of scans must be performed. In this case each scan has available in a local coordinate system. The reconstruction of the 3D model of the surveyed object requires the registration of the scans in a single local reference system. This phase can be performed in an interactive environment through the identification of the homologous points in adjacent and overlapping scans. Once the corresponding points are collected (at least 3), the 6 parameters transformation can be estimated and all the points of the scan can be changed into the coordinate system that has been assumed as the reference system. Thus, tie and control points, which are recognized in the point cloud, are needed to accomplish this task. These points may just be features of the object such as corners or special targets such as spheres or planer targets with high reflectivity.

In the case of using the targets, each pair of scans needs to contain at least three targets. Laser scanner devices record the X, Y, Z point coordinates and the average reflectivity of the impact area of the laser beam. Buildings are usually made of poor reflectivity material, thus if some reflectivity targets are superimposed onto the object, they can easily be found. Their 3D position is computed in the respective local coordinate system and the matching positions are used to compute the transformation between the two scans. The accuracy of the resulting transformation depends on the precision with which the targets are scanned and localised in the scans. Some scanners support automatic detection and high-resolution scanning of the targets [Bornaz et al, 2002]. The main drawback of using artificial targets is that they need physically to be placed in the scene; often this is difficult or even impossible to place the targets especially for scanning historical buildings or monuments. Moreover, these artificial targets may occlude important parts of the intensity images, which are required for texture mapping [Lensch et al, 2000]. Another procedure for registering two range images is to identify two sets of corresponding 3D points, and then further corresponding searching is done to optimize the matching process. [Horn, 1987] presented a quaternion-based non-linear optimization method for this purpose. The automation of the process of establishing point corresponding was presented first in the Iterative Closest Point (ICP) algorithm [Besl & Mckay, 1992]. It requires that an estimation of the relative position between two scans is known a-priori, which is usually provided interactively by the user. Based on this estimation the algorithm searches for each point in one scan the closest point in the other scan and uses the corresponding point pairs to compute a new relative position between the scans. This process is repeated iteratively until the relative position of the scans converges. The ICP algorithm works well if the scans contain relatively large areas of continuous surfaces and have sufficient overlap. But several problems still remain. From a computational expense point of view it is highly time consuming due to the exhaustive search for the nearest point [Sequeira, et al, 1999]. In additiuon, the accuracy of registration obtained using the ICP algorithm depends on the surface shape. If insufficient shape information is available then inaccurate or incorrect registration may occur [Vanico & Brunnett, 2002]. Variants of this method include the work by

[Turk & Levoy, 1994] and [Rusinkiewicz et al, 2002]. Several improvements on the ICP method have been shown in [Masuda & Yokoya, 1995; Bergevin et al, 1996]. The iterative closest compatible point (ICCP) algorithm has been proposed in order to reduce the search space of the ICP algorithm [Godin & Boulanger, 1995; Godin et al, 2001]. In the ICCP algorithm, the distance minimization is performed only between the pairs of points considered compatible on basis of their viewpoint invariant attributes such as curvature, color, and normal vector.

Nevertheless, the overall success of ICP-type methods is dependent on random parameters such as the starting position. Therefore newer approaches try to extract features from the range data that can be used for matching. [Sappa et al, 2001] propose to use ICP algorithm considering only a subset of points, the more pertinent ones. This subset is defined by the edge points-points where depth or orientation discontinuities appear. The proposed technique first extracts the edges from the two range images to be registered. Then, only over those extracted points the ICP algorithm is computed. This strategy is advantageous not only in terms of processing speed-up, but also in terms of the sensibility to the initial estimated transformation used to start ICP algorithm. [Wünstel & Röfer, 2003] present an approach that first detects certain global features of the room and then uses their semantic and spatial information to register the range images. [Stamos & Leordeanu, 2003] present an automated registration method for the accurate and efficient registration of a large number of complex ranges scans. The algorithm automatically computes pairwise registrations between individual scans, builds a topological graph, and places the scans in the same frame of reference. There can be found many other feature-based ICP approaches in the literature [Chua & Jarvis, 1996; Feldmar & Ayache, 1996; Wyngaerd, et al, 1999]. Other algorithms for combining multiple range images from different viewpoints are widely discussed in [Chen & Medioni, 1992; Zhang, 1994; Soucy & Lauendeau, 1995; Huber & Hebert, 2001; Grün & Akca, 2004].

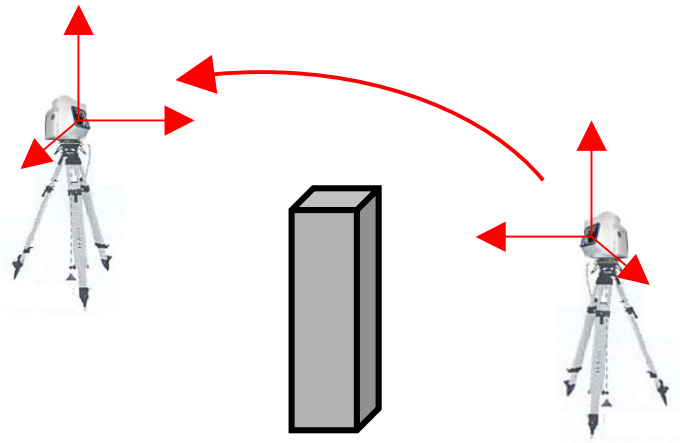


Figure 2.6: Multi scans for large and complex object

Merging Process

In this process, the areas where the images overlap must be integrated to create a non-redundant mesh. Two different approaches have been proposed to perform a surface-based integration of a set of multiple registered range images. The first approach compute surface models on sets of unorganized 3D points [Hoppe et al, 1992; Szeliski et al, 1993; Müller et al, 1998]. The methods work as following: (a) given a point p, its n nearest surface neighbors can be located by finding its n nearest 3D neighbors, (b) the density of data points should be quite uniform over the surface, and (c) the points are measured with the same accuracy. Thus computing a surface model from only range data points would lead to a model having the accuracy of the worst measured data point. These assumptions are too restrictive for using these methods to integrate sets of multiple range images [Soucy & Laurendeau, 1995]. The other approach for integration assumes that a parametric surface description can be derived from a single range image. An integrated surface model can then be obtained by merging multiple surface descriptions using an appropriate technique such as triangulation algorithms. The construction of a 3D range surface starting from a 2.5D range image is quite simple and easily feasible also in the cases of a nonuniform data distribution. [Soucy & Laurendeau, 1992] use a structured integration technique to combine multiple range images. Given n range images of an object, they first partition the points into a number of sets that are called common surface sets. The range points in one set are then used to create a grid of triangles whose positions are guided by a weighted average of the points in the set. Subsets of these grids are stitched together by a constrained Delaunay triangulation in one of n projections onto a plane. [Turk & Levoy, 1994] merge overlapping triangulated meshes using a zippering approach. The task of integration starts by creating a mesh of the object. Then, refining the vertex positions of the mesh by averaging the geometric detail that is present in all scans.

Merging begins by converting two meshes that may have considerable overlap into a pair of meshes that just barely overlap along portions of their boundaries. This is done by simultaneously eating back the boundaries of each mesh that lie directly on top of the other mesh. Next, the meshes are zippered together: the triangles of one mesh are clipped to the boundary of the other mesh and the vertices on the boundary are shared. Once all the meshes have been combined, they allow all of the scans to contribute to the surface detail by finding the consensus geometry. The final position of a vertex is found by taking an average of nearby positions from each of the original range images. [Häusler & Karbacher, 1997] perform vertex insertion; gap bridging, and surface growth operations to merge the input meshes. Merging is followed by edge swap operations to obtain more balanced triangulations. Other methods and algorithms could be found in [Ratishauser et al, 1994; Rocchini et al, 2004].

2.4 Combining Laser Data with Image Based Techniques

Frequently a combination of terrestrial LIDAR and image data is applied during 3D reconstruction of complex terrestrial scenes in the context of cultural heritage applications. A successful integration of both data sources will support tasks like surface reconstruction and will facilitate subsequent processing activities such as the generation of 3D textured models. The combination of the different data sets requires an exact co-registration, which has to be solved by a 2D-3D pose estimation algorithm. Since the accuracy of the transformation depends on the reliable extraction of primitives from the range and intensity images, image segmentation algorithims are considered in order to extract of registration primitives. In this research we present an efficient edge detection algorithm for the extraction of linear features in both range and intensity image data. The purpose is to simplify the dense datasets and to provide stable features of interest, which are used to recover the positions of the 2D cameras with respect to the geometric model for tasks such as texture mapping. Chosen of line feature as the registration primitives is motivated by the fact that such primitive can be reliably, accurately and automatically extracted from photgrammetric and lidar data sets.

2.4.1 Range Data Segmentation

Range images provide geometric information about the object independent of the position, direction and intensity of light sources illuminating the scene or of reflectance properties of the object. For these reasons it has played a very important role in image understanding, three-dimensional object reconstruction, autonomous navigation, medical diagnosis, etc. Segmenting range images into homogeneous regions is the first step to recognise 3D objects. There are many definitions of segmentation [Lee et al, 1998], but simply, segmentation is to break the image into some meaningful non-overlapping homogeneous regions whose union is the entire image. Segmentation of range images has long been considered an important and difficult problem and continues to attract the attention of researchers in computer vision. However, ready-made solutions for range image segmentation are not available as it is the case for the intensity images [Gächter, 2005]. 3D segmentation techniques may be categorized under three main classes; region-based, edge based and hybrid approaches using region and edge methods. Region-based segmentation methods group range pixels into connected regions based on some homogeneity measure. Typically, they grow regions around seed points for which fitting a local surface gives a reliable estimate. In the edge-based segmentation approaches discontinuity points are first extracted from the range image and then used to guide the segmentation process. Both approaches have their strength and weaknesses and need an extensive post-processing phase in order to obtain the final segmentation. In hybrid technique both region and edge information are used to guide the segmentation process.

Region Based Approaches

Region growing ideas are used commonly in range image analysis. [Miligram & Bjorklund, 1980] presented an approach to planar surface extraction from range images created by an imaging laser data. [Henderson, 1983] developed a method for finding planer faces using different range data. Several papers have discussed range image techniques for detecting the cylinders in range data [Popplestone et al, 1975; Nevatia & Binford, 1973]. [Bolles & Fisher, 1981] present the Random Sample Consensus (Ransac) to finding ellipse and cylinders in range images from an active triangulation light strip range imaging sensors. [Herbert & Ponce, 1982] proposed a method for segmenting the range images into plan, cylindrical and conical primitives. Some of the researches above are limited to searching for particular structures in the data. For general range image segmentation algorithm, the concepts and techniques from differential geometry have been useful in describing the shape of arbitrary smooth surfaces arising from range images [Vemuri et al, 1986; Ferrie & Levine, 1985; Hoffmamn & Jain, 1987]. In this field the work of [Besl and Jain, 1988] is the most

important. The proposed algorithm depends on the differential geometry quantities to have initial classifying of the points. The method starts by approximating of the regions with polynomial surfaces, the smooth surfaces are locally characterized according to eight fundamental forms obtained by using the sign of the Gaussian and mean curvatures, as depicted in figure 2.7. These curvatures are chosen because they are invariant to rotation and translation of the data, and that points of the same curvature are likely to belong to the same surface. This coarse segmentation provides connected regions of coarsely classified points. Then a seed point in each region is chosen and a planar or quadratic surface fitted. Neighbouring points are incrementally added until the goodness of fit of the surface is below a certain threshold. When that happens, either the process stops or a surface of higher order is used. The advantage of this method is the efficient use of differential geometric quantities and the classification of points using fundamental forms. The weaknesses lie in the complexity and the large number of parameters used. In addition the approach needs a considerable calculation time necessary to grow the seeds, and controlling the overlapping of the grown regions seems a difficult task [Djebali et al, 2002].

Parallel to the Besl work, [Hoffman & Jain, 1987] calculate surface normals at each point and use clustering techniques to group points into surface patches. The clustering uses a six dimensional point feature; three spatial dimensions for surface coherence and detection of discontinuities, and three for the surface normal to connect surfaces of smooth curvature. A series of tests are applied to classify the regions as planar, convex or concave. Adjacent patches of similar classification can then be candidates for possible merger. [Sabata, 1990] continue works in using the homogeneity of normal vectors and their three projections onto the xy-plane, the yz-plane, and the zx-plane. After the initial clustering, the method proceeds to refine the clustering iteratively using a pyramidal algorithm. Four independent segmentations are made to form an over-segmented image. These are then merged by higher-level routines.

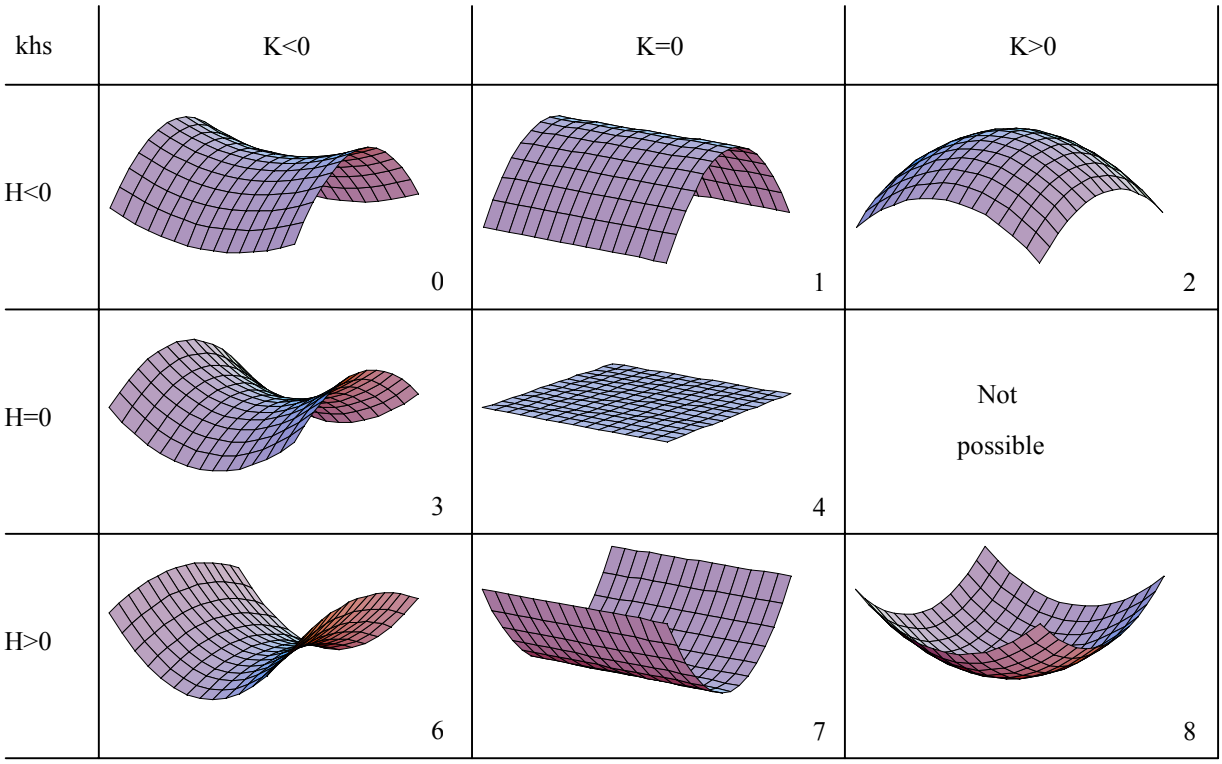


Figure 2.7: Fundamental surface types

Recently, [Lee et al, 1998] propose a novel range image segmentation technique using robust, adaptive least kth order squares (ALKS) estimator .The algorithm defines the region to be processed as the largest connected component of unlabeled pixels. Then the adaptive least kth order squares (ALKS) estimator is applied to minimize the kth order statistics of the squared of residuals. The optimal value of k is determined from the data, and the procedure detects the homogeneous surface patch representing the relative majority of the pixels. The ALKS procedure is used to the selected region and discriminate the inliers. After labeling the largest connected component of inliers as the delineated homogeneous patch, the model parameter estimates is refined by a least-squares fit to the inliers. The process will be repeated until the size of the largest connected component is less than a threshold. Finally, the isolated outliers sur-

rounded by inliers are eliminated, and the unlabeled pixel is allocated to the class of the majority of its labeled four-connected neighbors.

In general, most of range data segmentation algorithms mentioned above need pre-processing steps such as application of smoothing and median filters, prior to calculating any surface characteristics. For algorithms depend on partial derivatives estimation the sensitivity of these methods to noise in the data (especially outliers) may also lead to mis-classification and poor results [Marshall et al, 2001]. Some of them they are failing totally [Böhm, 2003]. In addition, the positions of the initial seeds greatly affect the performance of most region-based methods. When the seeds are placed on a boundary or a noise corrupted part of the image, the results will break down. Another general problem of region growing approaches is that they are computationally expensive [Hatger & Brenner, 2003].

Edge Based Approaches

Comparatively to the much interest in region segmentation, few segmentation works are based on edge based techniques only. Most of the researches above are focused on simple polyhedral objects and they are limited to searching for particular structure in the data such as straight and circles edges for matching purposes. [Mitiche & Aggarwal, 1983] developed an edge detector based on a probabilistic model that accounts for range measurement errors. Step edges are extracted first using standard techniques. Extraction of roof edges is important for many applications such as city modeling. Detecting roof edges usually need more difficult process since the information is usually embedded in noise. Thus, in theire approach, roof edges are hypothesized for each discrete image direction (usually four directions applied) at each pixel. Two planes are fitted to the immediate neighborhood of the pixel. The dihedral angels between theses plans are computed. If the dihedral angles are less than a threshold, the corresponding pixels will be neglected. A Bayesian likelihood ration is then computed and the most likely edge direction is chosen. Finally all the remaining pixels are passed through a non-maxima suppression algorithm that theoretically leaves only the desired edge pixels. [Hermann, 1985] extracted a details polyhedral object description of range images using edge approach. Edges are computed to create a line drawing using a Hough transformation. [Fan et al, 1987] computed directional surface curvature in the four main directions (0,45,90,135) by convolution masks, for the each of the zero crossing and local extrema are computed, local extrema is defend as a point whose absolute value is strictly larger than the absolute value of one of its neighbors and zero crossing is given by zero surrounded by nonzero numbers of opposite sign on the tow sides or by sequence of two numbers of opposite sign. To increase the confidence in localization of the features, scale space tracking is used. extrema and zero crossing are tracked independently. Then the four tracked images are merged into a single image by logical or operation. Because it depends on the values of curvature in the classification, the approach is very sensitive to the noisy and the smoothing process will crucially influence the results.

While in the past range data collection was mainly applied for industrial scenes captured at close distances, nowadays long-range laser scanners are available for many users. By these means detailed data sets of complex outdoor scenes are collected, which pose much more serious challenges for range image analysis than the traditional polyhedral world. The difficulties result from the fact that range data of natural scenes are relatively noisy. These measurement errors affect the approximation of the surfaces during segmentation. In addition, the natural scenes are complex since lots of individual objects or irregular surfaces occur. For segmentation of this type of data [Sappa et al, 2001] propose a two-step approach. The first step generates a binary edge map based on a scan line approximation technique as e.g. proposed by [Jiang & Bunke, 1999]. The second step aims on contour extraction by a weighted graph. A minimum spanning tree (MST) is computed to obtain the shortest path, which links all the edge points. One of the main drawbacks of this algorithm is the fact that during the MST filtering many edges are eliminated. Recently [Han et al, 2004] presented a stochastic jump-diffusion algorithm for the segmentation of range images in a Bayesian framework. The algorithm can be used for processing of complex real-world scenes. Although it is considered as the most advanced algorithm for complex scene segmentation, some drawbacks such as computational complexity and the large number of required parameters are still mentioned. In addition, suitable a priori assumptions are required.

Some existing algorithms are limited to high quality range images and will fail in the presence of noise. Others are complicated and have a large numbers of parameters while generic and efficient edge detectors for range images are still missing. This was our motivation for the development of an edge detection algorithm for range images.

Hybrid Approaches

Both region and edge segmentation techniques have some desirable strength. Region-based techniques are generally quite robust in noisy environment, whereas the edge-based techniques respond strongly at the boundaries between adjacent regions. In addition, both of those techniques also have their own drawbacks. Most region growing-based

techniques highly depend on the selection of seed regions for their success. For some clustering-based methods if the specified number of clusters is less than the actual number of clusters in a range image then pixels belonging to different regions would be grouped as a single cluster. If the number of clusters is over specified then there is a chance of over segmentation and pixels belonging to the same physical region would be partitioned into several physically adjacent clusters. On the other hand, edge-based techniques suffer from the fact that extensive empirical post processing may be needed to provide final segmentation. Moreover, the edge maps are in general not closed and incorrect edge points may be detected. Some of the researchers solve such problems of the both techniques by means of hybrid approaches. The information provided by the edge detection method is used to estimate the number of the regions necessary to initialize the clustering algorithm. So the edge detection technique is carried out to guide and improve the clustering technique applied to the region-growing algorithm. [Yokoya & Levine, 1989] propose a hybrid approach to the problem of range image segmentation. The initial region-based segmentation is equivalent to the method of [Besl and Jain, 1988], except for the use of the selective local surface fit. The method employs a selective surface fit based on the computation of first and second partial derivatives determined by locally approximating object surfaces using biquadratic polynomials. By computing the Gaussian and mean curvatures and examining their signs. Two initial edge-based segmentations are also computed from the partial derivatives and depth values. One detects jump edges by computing differences in depth, while the other highlights roof edges by differences in surface normals. The three initial image maps are then combined to produce the final range image segmentation. Major problems can be stated to the different values of the thresholds, which yield different KH-sign maps. And subsequently different region maps.

Recently, [Stamos & Leordeanu, 2004] developed a segmentation algorithm to extract 3D linear features from planar intersections. They group the 3D points from the range scans into clusters of neighboring points, which correspond to the same planar surface. In the point classification phase a plane is fit to the point's neighborhood. The normal of the computed plane is computed corresponds to the smallest eigenvector of the x by x matrix. The smallest eigenvalue expresses the deviation of the points from the fitted plane. If the deviation is below a user specified threshold the center of the neighborhood is classified as locally planar point. The next step is to merge the initial list of clusters and to create a minimum number of clusters of maximum size. After clustering process, the intersection of the planar regions provides three-dimensional lines. Extraction of 3-D lines is done in three stages: Intersection of neighboring 3D planes to produce 3D lines of infinite extent. Verification of the infinite 3D lines, this step involves the computation of the distance between the bounded surfaces and the produced 3D line. Generation of 3D linear segments out of the infinite 3D lines, this is done by keeping the parts of the infinite 3D lines which are verified by the range data-set.

2.4.2 3D-2D Registration Process

Pose estimation or external camera calibration is the problem of estimating the location (position and orientation) of a 2D camera with respect to a 3D model. Most systems depending on combining of range and image sensing solve the 3D range to 2D image registration problem by fixing the relative position and orientation of the camera with respect to the range sensor. The two sensors are rigidly attached on the same platform [Sequira et al, 1999; Pulli, 1998; Zhao & Shibasaki, 2003]. The static arrangement of sensors leads to a lack of 2D sensing flexibility, the 2D camera cannot adjust by changing its focal length and position to the requirement of each particular sense. Also, the light condition at the time of 3D range capturing may not suitable for the 2D images. Additionally, for many laser scanner systems nowadays available camera has low resolution that gives poor color information and poor texture, which may not desired for documentation purposes.

In another side, using a free camera, one can choose higher image resolution for texture mapping. Moreover, it allows taking pictures from different points of view, so to handle occlusions and fill the gaps in the 3D laser model by means of convergent photogrammetry [Deveau et al, 2004]. Using freely 2D camera independently of 3D range data needs to find translation (three parameters) and rotation (another three parameters) that describe the exact pose of the camera when the image was taken. A number of traditional photogrammetric approaches use point matches and the advantageous to utilize linear features in photogrammetric applicationsc have been already discussed in section 1.2. However, linear feature contains more geometric and semantic information than point. It is also easier to be extracted from the photogrammetric images.

For line landmarks three correspondences are enough for estimating the pose parameters [Kumar & Hanson, 1994; Christy & Horaud, 1996]. However, due to noisy measurements the minimum number of correspondences is not enough for a robust solution [Liu et al, 1990]. [Ansar & Danilidis, 2002] propose a novel set of algorithms for pose estimation from lines. The solutions are developed from a general procedure for linearizing quadratic systems of a

specific type. If a unique solution for the pose problem exists, then the algorithms are guaranteed to return it. A non-linear solution is possible if at least three line correspondences are given. [Chen, 1991] showed that with three line correspondences the solutions are given by an eight-degree equation in one unknown. [DeMenthons & Davis, 1995] method is among one of the first attempts to link linear techniques with non-linear techniques. The method consists of iteratively improving the pose computed with a weak perspective camera model to converge, at the limit, to a pose estimation computed with a perspective camera model. In [Kurazume et al, 2002] the traditional ICP algorithm has been extended for use in projective registration. The edges are first extracted from both reflectance and intensity images. The 3D- 2D correspondences between the extracted edge points are then established using the traditional ICP algorithm. Another technique proposed in [Stamous & Allen, 2003] achieves image to range registration based on vanishing lines and orthogonality conditions but the method is only reliable in scenes with strong geometric contents.

The previous methods are able to solve the correspondence problem, but for simplification they are using a weak perspective. For exact co-registration and georeferencing a further iteration is necessary using the spatial resection approach [Hrabacek & van den Heuvel, 2000; Heuel and Förstner, 2001]. Recently, [Klinec & Fritsch, 2003] present a robust spatial resection approach to overlay extracted straight lines in the image and corresponding edges of the 3D model to estimate the parameters of the exterior orientation. For the feature based spatial resection initially the well-known collinearity equations are utilized. But as they are using straight lines as tie-information, the commonly used collinearity equations are modified by the parameterizations. To solve the spatial resection problem, a least squares algorithm using the Gauss-Markov-Model was implemented. In a least squares approach (as used for solving the spatial resection) the non-linear equations cannot be used without initial values. To overcome this dilemma and for estimation of rough (initial) exterior orientation parameters the feature based direct linear transformation (DLT) is used.

2.4.3 Texture Mapping

Texture mapping is a relatively efficient means to create the appearance of complexity without the tedium of modeling and rendering every 3-D detail of a surface. The technique has been applied to a number of surface attributes: surface color, surface normal [Blinn, 1978], transparency [Gardner, 1985], specularity, illumination, and surface displacement [Carey et al, 1985]. In this chapter the approaches of mapping colored images will be discussed. In general there are two tasks common to all texture mapping approaches: the mapping of the texture function from texture space to screen space and the filtering that is necessary in order to avoid aliasing. The mapping from texture space to screen space is split into two transformation phases. One transformation is between object space and texture (input) space and the other is between object space and screen (output) space, as depicted in figure 2.8. Texture space is labeled (u, v), object space is labeled (x_o, y_o, z_o), and screen space is labeled (x, y). Often the object–texture and object– screen transformations are concatenated in order to save computations. Typically, this transformation is done using either affine or perspective projection [Heckbert, 1986].

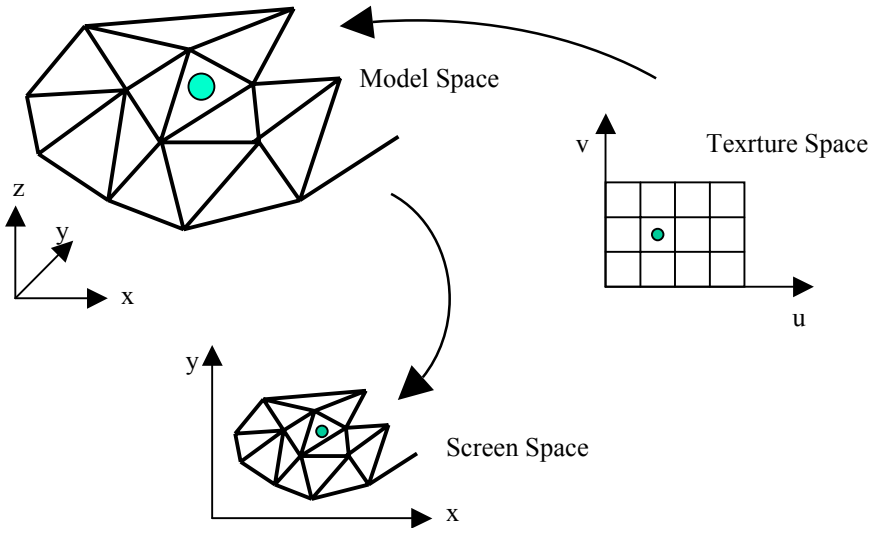


Figure 2.8: Transformation relationship between texture, model and screen space.

In affine mapping [Paeth, 1986; Nehlig & Ghazanfarpour, 1992; Remondino & Niederoest, 2004], the texture coordinates of a polygon are linearly interpolated across the polygon surface. The warping combines of scales, rotations, translations, and shears transformations. This technique does not account for perspective and therefore produces distortion texture. Thus, it works well when there is a small angle between the view direction and the surface normal of the object being rendered, since in these cases the z distance varies little across the face. Despite its obvious problems, affine mapping is a very useful when combined with other rendering techniques to speed up the rendering process. The other transformation technique is the perspective mapping, also known as the projective or homogenous transformation, in this approach the projection is done from one plane through a point onto another plane. The important difference is the use of photogrammetry techniques to represent the texture imagery acquisition geometry. The co linearity condition in photogrammetry is the basis for characterizing the acquisition geometry. This type of projection preserves lines at all orientations and sacrifice equal spacing. The techniques used in the perspective transformation of images have been discussed in [Heckbert, 1983; Lavalle´e & Szeliski 1995]. The choice of a suitable transformation approach depends on the model and the position of the camera taken for texture acquisition. Projective mapping ensures the proper perspective foreshortening while the linear parameterization of the affine mapping leads to errors. Those errors depend on ; triangle size, the location of texture space coordinates of each triangle vertex, the angle between camera viewing direction and the triangle normal [Niem & Broszio, 1995].

After the mapping is computed and the texture is warped, the image must be resampled on the screen grid. This process is called filtering. The cheapest texture filtering method is point sampling, where the texture value nearest the desired sample point is used. Some of the more sophisticated techniques that have been used to anti-alias texture-mapped images are; trilinear interpolation on hierarchical resolution texture patterns [Tanimoto & Pavlidis, 1975; Burt, 1981] and MIP maps [Williams, 1983], elliptically weighted averaging (EWA) [Greene & Heckbert, 1986]. [Heckbert, 1986] presented a review and comparison of many of these anti-aliasing techniques.

For the texture mapping of building facades, a large number of approaches have been presented [Brenner & Haala, 1998; Wang et al, 2003; Debevec et al, 1998; Boehm, 2004; Ortin & Remondino, 2005]. One great problem in texturing buildings façades is the presence of obstacles in front of the façade. Direct warping independent images onto the collected object surfaces using perspective or affine transformation projection can only be applied for simple objects with restricted number of surfaces with no occlusions [Ofek et al, 1997; Visnovcova et al, 2001; Kada et al, 2003]. Otherwise, if direct projection transformation is applied without considering occlusions, the mapping between object geometry and image will be incorrect. Warping the image over the geometry associated each 3D point to a pixel in the color image, the texture pierce through the geometry and gets mapped onto all occluded polygons on the path of the projected ray. So the geometry that is occluded in the image will receive incorrect texture coordinates instead of remaining in shadow. Figure 2.9 shows this problem and defines the three types of occlusion problems as they are known in computer graphics applications; ambient, self and view frustum occlusions.

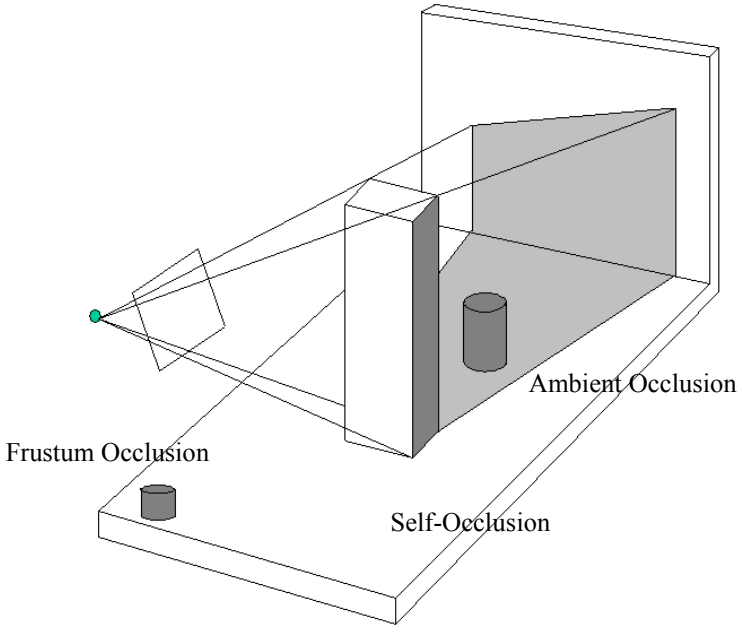


Figure 2.9: Types of occlusion

The occlusion problem can be avoided by manual texture extraction and mapping, however, such tedious task can take up to several days for good results. Fully and semi-automatic texture extraction and placement of the real scene have been presented in different works [Grammatikopulos et al, 2004; Sequeira & Goncalves, 2002; Neugebauer & Klein, 1999] but with more complex processing. Those approaches use automatic techniques for occlusion detection. The goal of the visible surface algorithms is to detect and draw only the surfaces that are visible from given viewpoint as can be depicted in figure 2.10. Two main types of approaches for occlusion detection are presented: the first one working in image space and compute the visible surface at image resolution. Where as the other working in object space and compute the visible surface at object precision. An example of image space occlusion method is the scan-line algorithm [Romney, 1970]. It generates the graphic image left to right, top to bottom in the same way as the picture is scanned on a TV screen. The algorithm limited its computation on each scan line to the polygon edges and faces intersect with the scan line.

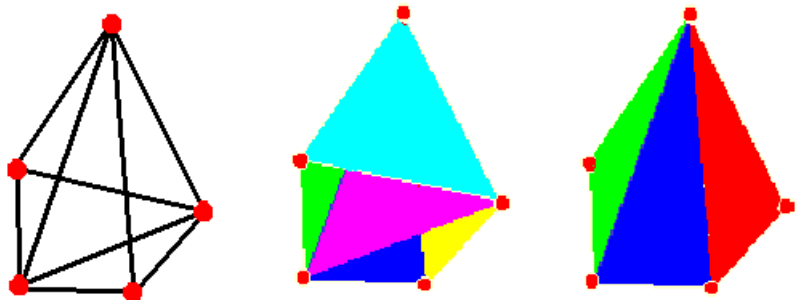


Figure 2.10: The goal of visible surface algorithms: detect and draw only the surfaces that are visible from given viewpoint

[Catmull, 1975] presents more simple and easy implement Z buffer algorithm works in image space. In the algorithm the visibility is computed by comparing the surface depths at each pixel position on the projection plane. If the depth value is greater than the current value for the 3D point, the point associated to the previous depth value will be labeled as occluded, see figure 2.11. Compared to other occlusion algorithms, it is easy to implement but requires large memory and long computation time. [Grammatikopulos et al, 2004] speed up the searching process of the Z buffer algorithm by tessellating the textured image area into a rectangular grid with cells larger than those of the original image. In general, the Z_buffer algorithm shows quantization errors and bad performance as the complexity of scene is increasing [Sutherland et al, 1974]. This algorithm has the obvious disadvantage that it can handle a maximum 180-degree shadow frustum per buffer. If you placed a light source inside the scene, then you might need six such buffers to handle all shadow cases. Thus, the algorithm deals most effectively with spotlights and directional lights. In addition, it has aliasing problems due to discretized depth map cells and orientation of the shadow z-buffer. [Hourcade and Nicolas, 1985] attempted to improve the aliasing problems by applying stochastic sampling and some pre-filtering to the z-buffer approach. Other occlusion culling algorithms calculated in image space discussed also in [Zhang et al, 1997; Greene & Kass, 1993].

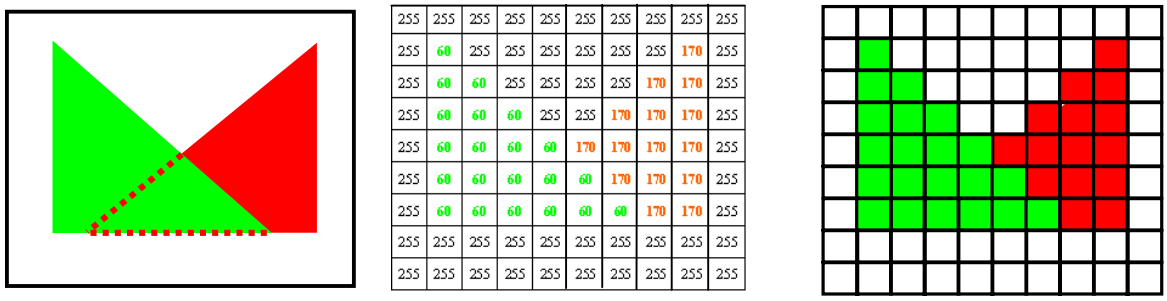
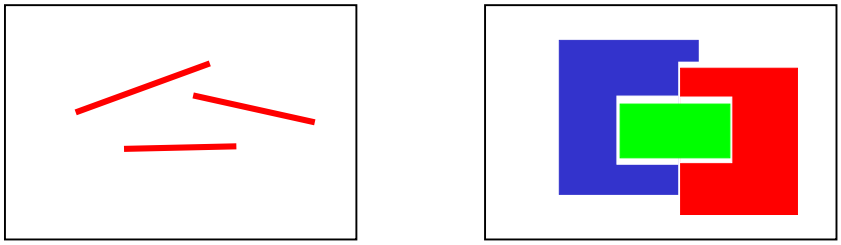


Figure 2.11: Example using integer z-buffer value with near = 0, far =255 .The green triangle depth value = 60, the red triangle depth value =120

The algorithm of [Appel, 1967] exemplifies a class of hidden-line algorithms that compute line drawings in object space. The basic idea derives from the fact that the visibility of a edge can only change at the point where the projec-

tion of the edge cross the projection of some contour edge on the viewing plane. Appel defines the quantitative invisibility of a point as the number of relevant faces that lie between the point and the viewpoint. The solution to the hidden-line problem requires computing the quantitative invisibility of every point on each relevant edge. Another example, which is very widely used, is the painter's algorithm. First, the algorithm sorts polygons by z-coordinates in object space. If the polygon's extents overlap, the overlapped triangles need to be spitted, then the polygons are sorted and rendered in back to front order as can be depicted in figure 2.12. [Woo et al 1990; Sutherland et al, 1974] present a valuable survey of the occlusion algorithms.



Top view Side view

Figure 2.12: Painter's algorithm

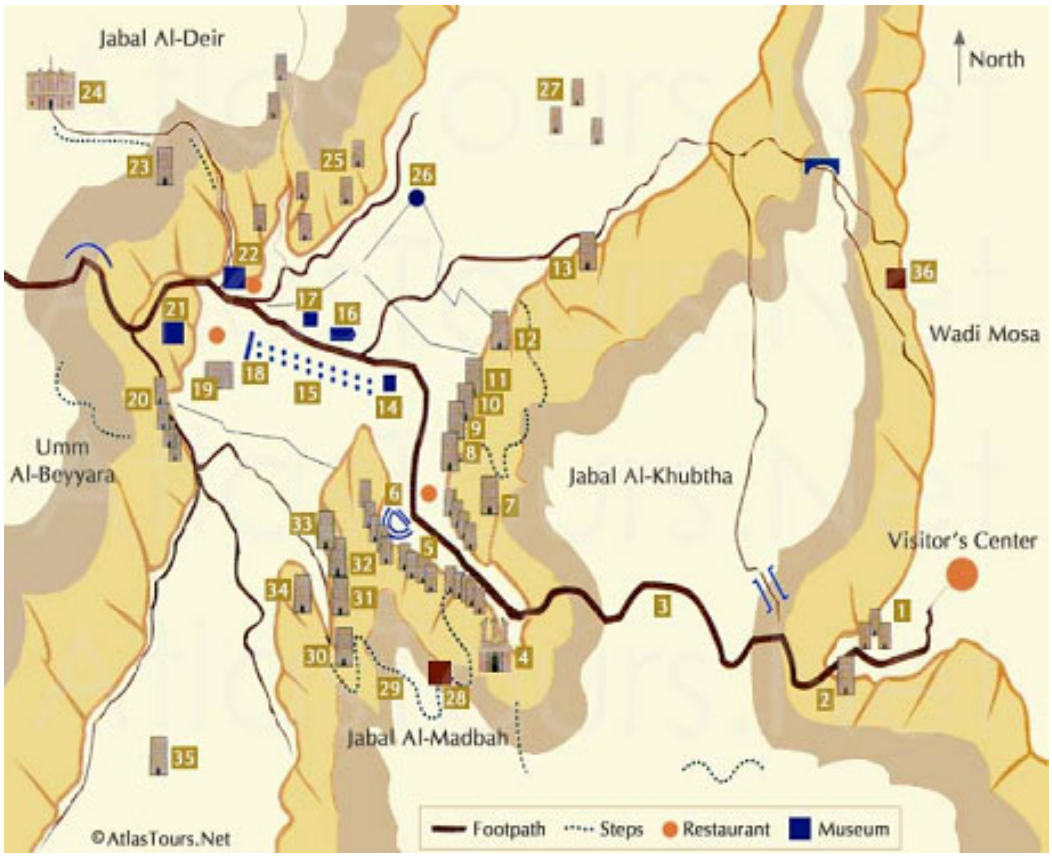
In general image space algorithms are easy to implement comparatively with the object space algorithms but they require large memory and long computation time. In addition they show bad performance as the complexity of the scene is increasing. On the other hand, object space visibility algorithms, which can be used alternatively, show high precision and good results in high complex scenes but they are computationally expensive and difficult to implement [Luebke et al, 2003]. In this research, we presented a new automatic technique to enhance the 3D geometric model of real world objects with texture reconstructed from separate sets of photographs. For this purpose an efficient algorithm addressing the image fusion and visibility is used. One advantage of our algorithm that it is easily allows the visibility detection of different resolution of 3D geometric models and 2D images. .

3 Data Collection and Pre-processing

The collection of the data, which has been used for our investigations, was performed in cooperation between Institute of Photogrammetry (ifp), Stuttgart University and Queen Rania Institute of Tourism and Heritage, Hashemite University of Jordan. One of the project goals is the generation of a 3D documentation of the Al-Khasneh monument in Petra city, and the north theatre of ancient Jerash city. The resulting accurate 3D models with rigorous metric values will support archaeology, history and culture researchers with an accurate geometric recording resource, also will help architects for further studying and evaluating the restoration process and maintenance plan work. Additionally, the resulting virtual presentations of the historical buildings in these heritage sites can be used for promotion purposes in tourism. From photogrammetric point of view, architecture of the ancient Nabataean city of Petra are carved into rose-colored sandstone cliffs and present a typical example of 2.5 model that could be accessd from one main viewpoint whereas the architectures in ancient jerash city spread in a wide open area and could be accessed from different viewpoints.

3.1 Treasury Monument of Ancient Petra City

The ancient Nabataean city of Petra has often been called the eighth wonder of the ancient world. Petra city in southwestern Jordan prospered as the capital of the Nabataean empire from 400 B.C. to A.D. 106. The Nabatu migrated from western Arabia and settled in the area. By the fourth century BC they occupied the region and made it the capital of their kingdom. From its origins as a fortress city, Petra became a wealthy commercial crossroads between the Arabian, Assyrian, Egyptian, Greek and Roman cultures. Control of this crucial trade route between the upland areas of Jordan, the Red Sea, Damascus and southern Arabia was the lifeblood of the Nabatean Empire and brought Petra its fortune. The riches the Nabateans accrued allowed them to carve the monumental temples, tombs and administrative that we see today. Besides trade, the Nabataens created one of the most advanced hydraulic systems at that time in the world, developing water conservation systems and a system of dams to divert the rush of swollen winter waters that created flash floods in the area. Sophisticated ceramic pipelines, reservoirs, gravity feeds and cisterns served the urban environment, while outside the city, dams closed wadis in order to collect water during the rainy season, stone terraces retarded runoff and trapped topsoil and irrigation lines fed crops. After the Roman conquest of Syria and Palestine in 63 BC, the Romans gradually extended their control into the south of Jordan, and in the year 106 AD, Petra and Nabataea then became part of the Roman province known as Arabia Petra with its capital at Petra. In 313 AD, the Romans recognized Christianity as the state religion, and in 330 AD Constantine established the Eastern Roman Empire with its capital at Constantinople. With this, the Byzantine influence began to invaded Petra. However, with the changing trade routes, Petra's commercial decline was inevitable and its demise was further impacted by a severe earthquake in 551 AD. It continued to decline with the Muslim invasion of the 7th century. Petra was revealed to the western world in 1812 for the first time since the Crusades when it was re-discovered by the Swiss explorer Johann Ludwig Burckhardt. Today Petra is considered to be one of the most unique cities in the world. Nominated by UNESCO to be the 8th wonder of the world. Petra's temples, tombs, theaters and other buildings are scattered over 400 square miles, Petra map is depicted in figure 3.1. These architectures are carved into rose-colored sandstone cliffs. After a visitor enters Petra via Al-Siq, a two-kilometres impressive crack in the mountain, the first facade to be seen is Al-Khasneh, which is depicted in figures 3.2. It is considered as the best-known monuments in Petra city. The Al-Khasneh facade is 40m high and remarkably well preserved, probably because the confined space in which it was built has protected it from the effects of erosion. The name Al-Khasneh, as the Arabs call it, means treasury or tax house for passing camel caravans, while others have proposed that the Al-Khasneh Monument was a tomb. Behind the impressive facade of Al-Khasneh, large square rooms have been carved out of the rock. As one walks further into Petra, one can see the remains of different tombs, depicted in figures 3.3, 3.4. However, to reach the inspiring monastery (Al-Deir) one must follow the Roman Cardo to the end of the valley, and climb the staircase (approximately 900 steps). The monastery is depicted in figure 3.5 [Sedlaczek, 2000].



1 Djin Blocks	10 Corinthian Tomb	19 Qasr Al-Bint	28 High Palace of Sacrifice
2 Obelisk Tomb	11 Palace Tomb	20 Unfinished Tomb	29 Lion Monument
3 Al-Siq	12 Sextus Florentinus Tomb	21 Al-Habees Museum	30 Garden Temple Complex
4 The Treasury	13 House of Dorotheos	22 Petra Archeological Museum	31 Triclinium
5 Street of Facades	14 The Nymphaeum	23 Lion Triclinium	32 Renaissance Tomb
6 The Theater	15 Colonnaded Street	24 Al-Deir	33 Broken Pediment Tomb
7 Aneisho Tomb	16 Byzantine Church	25 Turkmanian Tomb	34 Roman Soldier Tomb
8 Urn Tomb	17 Winged Lion Temple	26 Conway Tower	35 Snake Monument
9 Silk Tomb	18 The Arched Gate	27 Moghar Annassara	36 Crusader Fort

Figure 3.1: Petra detailed map (Atlas)

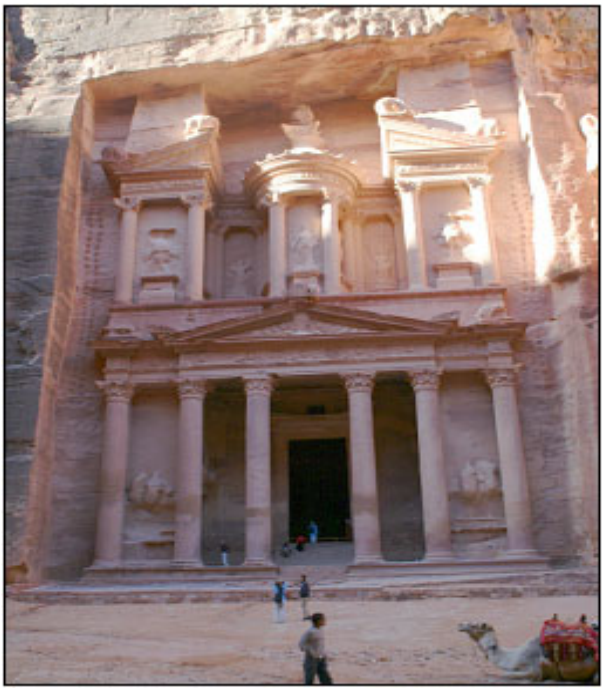


Figure 3.2: Al-Khasneh facade, Petra

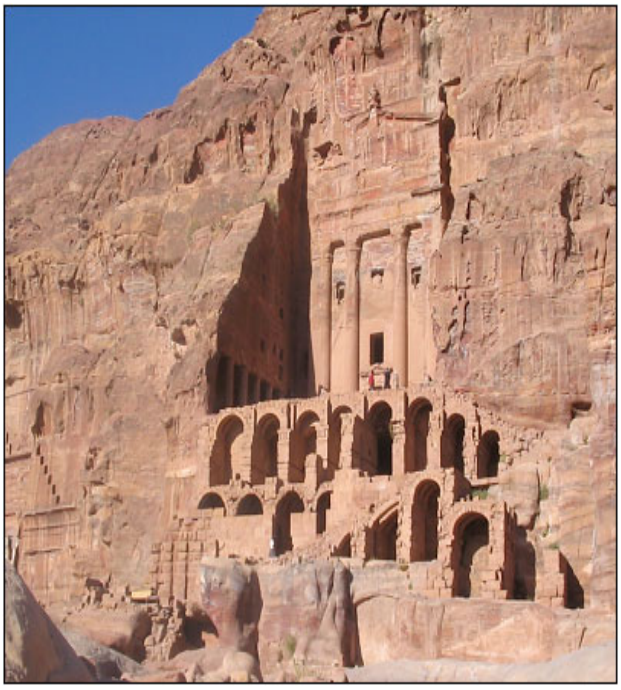


Figure 3.3: Royal tombs

Figure 3.4: Urn tomb

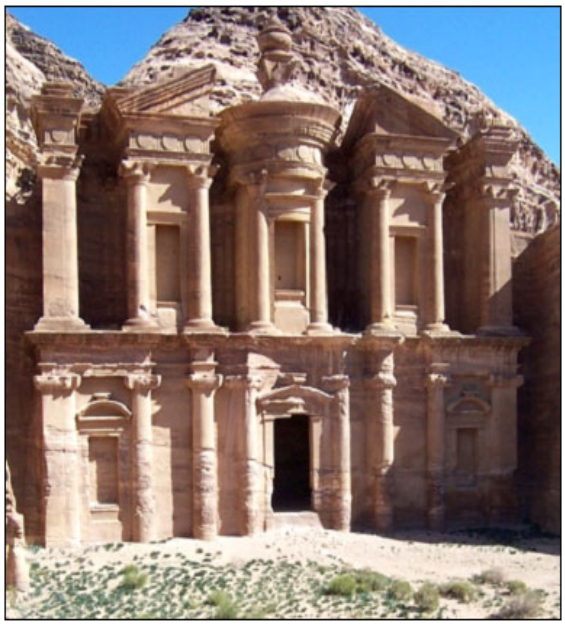
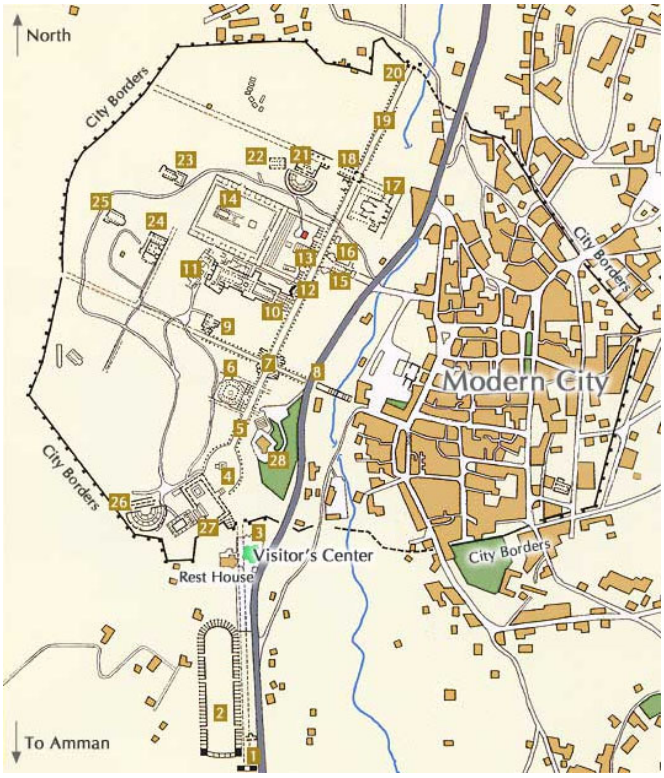


Figure 3.5: The Monastery

3.2 Ancient Jerash City

Jerash is located in the north of the Hashemite Kingdom of Jordan, at some 45 kilometres away from Amman the capital. It is 585 meters above the sea level; the map of the old city is depicted in figure 3.6. Alexander the Great founded Jerash in 332 BC, this is done during the spreading of the Greek civilization. Then commander Ptolemy II (285-246 B.C) transformed it from small village to metro pole. Jerash became a great city highly influenced by the traditional Greek Civilization, till the beginning of second century B.C. In 63 B.C Jerash was conquered by the Roman General Pompey who named it "Gerasa". At the beginning of the Romanian presence, Jerash joined the alliance of the ten free cities, known as the Decapolis. This alliance was set up by the Romans in order to protect their trade in Syria. After the expansion of the Arab Nabataean Kingdome in the south, Jerash expand its trade with the Nabataean, from which it adopted the currency, religious believes and architecture, which was perfectly mastered by the Nabataeans. The Nabatean left in Jerash some inscriptions glorifying the Nabataean god "Dushara" on which stands now the ruins of a Christian Cathedral. The Roman Emperor Trajan occupied the Nabataean capital "Petra" in 106 A.D. To celebrate his victory, the people of Jerash built the North Gate at the end of the Cardo or Colonnade Street, depicted in figure 3.7. Trajan (53-117 A.D) left beautiful buildings, temples, theatres and baths. In 395 A.D Christianity became the dominant religion in Jerash. This was reflected on the churches in Jerash whose number rose to twelve, particularly during the reign of emperor Justinian (527-565 A.D), the entrance of one of those churches is depicted in figure 3.8. The churches had walls covered with marble and mosaic. In 614 A.D Jerash was conquered by the Persians, who destroyed most of its monuments. In 627A.D the Roman Emperor Heraclius recoverd it and the city becaome Chassanide under their rule. In 636 A.D Jerash was conuered by the Muslims. It is flourished during the Islamic reign and it was famous for its metal and pottery industry. When the trade routes turned away from Jerash, its importance diminished. During the regin of Ottomans, Jerash became part of Ajlun's district. The Ottomans took god care of the city and buil a Mosque. In 1927, Jerash became a subdistrict in the Emirate Tranjordan under the British mandate till the independence in 1946. From then Jerash became one of the most important archaeological sites in Jordan. Gerasa is considered to be one of the best-preserved and most complete Roman cities in the world. There are a large number of still striking monuments: Gates, Arches, temples, colonnaded street, forum, theatres, baths, Byzantine churches and an almost complete circuit of city walls. Figure 3.9 shows the south theatre, which has been built during reign of Emperor Domitian, between 90-92 A.D, theatre seats more than 3000 spectators. Figure 3.10 shows the Nymphaeum, It is a splendid example of an ornamental fountain from the late 2nd century A.D. The North theatre, which we used for exemplarily data collection is depicted in figure 3.11. In front is a colonnaded plaza where a staircase led up to the entrance, depicted in figure 3.12. It was inaugurated in 165 A.D. The theater originally had only 14 rows of seats, and was used as a performance stage as well as the city council chamber. In 235 AD, the theater was doubled in size to its present capacity of 1600. Two vaulted passages formed the entrance to the orchestra, and spectators entered through passages between the upper rows of seats. The theater fell into disuse in the 5th century, and in later centuries, many of its stones were taken for use in other buildings. [Belloni, 2000;Ulama, 1996].



1- Hadrian's Arch
2- Hippodrome
3- The South Gate
4- Oval Plaza
5- Colonnaded Street"
6- Macellum
7- South Tetrapylon
8- South Bridge
9- Umayyad Houses
10- The Cathedral
11- Church
12- Nymphaeum
13- Propylaeum
14- Temple Esplanade
15- Propylaeum Church
16- Umayyad Mosque
17- West Baths
18- North Tetrapylon
19- North Colonnaded Street
20- North Gate
21- North Theater
22- Church of Bishop Isaiah
23- Temple of Artemis
24- Three Churches
25- Church of St. Genesius
26- South Theater
27- Temple of Zeus
28- The Museum

Figure 3.6: Detailed map of ancient Jerash city (Atlas)

Figure 3.7: The Columned street

Figure 3.8: The entrance of the Cathedral

Figure 3.9: South theatre of Jerash

Figure 3.10: Nymphaeum

Figure 3.11: North theatre of Jerash

Figure 3.12: The entrance of the theatre

3.3 Sensor Applied and Measurement Configuration

In our project, the 3D laser scanning system GS100, manufactured by Mensi S.A., France was applied. The scanner features a field of view of 360° in the horizontal and 60° in the vertical direction, enabling the collection of full panoramic views. The distance measurement is realized by the time of flight measurement principle based on a green laser at 532 nm. The scanning range of the system allows distance measurements between 2 and 100 meters. The scanner's spot size is 3 mm at a distance of 50 meters; the standard deviation of the distance measurement is 6 mm for a single shot. The system is able to measure 5000 points per second and it is suitable for our application. During data collection a calibrated video snapshot of 768x576 pixel resolution is additionally captured, which is automatically mapped to the corresponding point measurements. Depending on the application and the complexity of the environment, achieving high geometric and realism 3D model may require a large number of 3D and 2D images. Because it is not possible to have a complete 3D coverage for the Al-Khasneh facade based on data collected from a single station, three different viewpoints with five scans were done to resolve the occlusions. The problem to choose the viewpoint positions represents an important phase of the survey for such a monument since the mountainous environment surrounding Al-Khasneh restricts potential sensor stations. For Al-khasneh façade, three positions have been selected, from the en-

trance area of the monument, from the left of the monument, and one scan was collected from an elevated viewpoint. Since the vertical field of view of the laser scanner from these positions could not cover all the facade from one scan, the left and top scanning were done using 2 scans from the same position, taking into consideration sufficient overlapping regions to allow for a subsequent integration. In total, the five scans resulted in almost 5 million collected points. In additional to the outer survey, a 360-degree scanning from a station in the interior of the Al-Khasneh had been taken, which resulted in 19 million points. Figure 3.13 shows the point cloud of some of these scans for the façade with colour information overlaid. The North theater of Jerash is not confined by the environment as the Al-Khasneh, and its architectures are scattered in more wide area. Thus, seven viewpoints were selected to cover the theatre and the area surroundings in order to solve the occlusion; the setup positions are illustrated in figure 3.14. All the acquired 3D models have been processed using Innovmetric Software, PolyWorks.

Figure 3.13: Detailed view of the collected point cloud

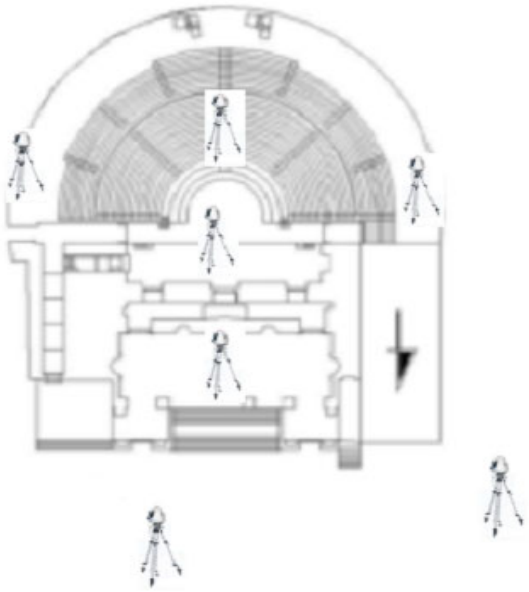


Figure 3.14: Selected scanning positions of the North theatre

The produced model for Al-Khasneh has an average resolution of 2 cm with more than 10 million triangles, as depicted in figure 3.15. In another hand, the north theatre model consists of more than 4.6 million triangles; some snapshot for the final models are depicted in figures 3.16 and 3.17.

As have been discuused in the introduction chapter, GS100 system has a low resolution camera already integrated in the system but the main uses is to support the semi automatic evaluation such as visual detection of the tie points and visual check of completation of the work. Thus these images are not sufficient for high quality texturing which is desired for documentation. The first row in figures 3.18 and 3.19 depicts images collected using the integrated camera for both Al-Khasneh and the North theatre. It could be noticed the recoreded images have different radiometric properties and this problem become clear in the case of the North theatre. The North theatre is more complex structure and its architectures are scatterd over a large area. Thus time differences between the collected scans result in varying light conditions and changing shadows. For this reason, additional images were collected by a Fuji S1 Pro camera, which provides a resolution of 1536x2034 pixel with a focal length of 20 mm. These images depicted in the bottom rows of figures 3.18 and 3.19. Figure 3.20 shows the 3D textured model of the treasury using the images collected by the integrated camera.

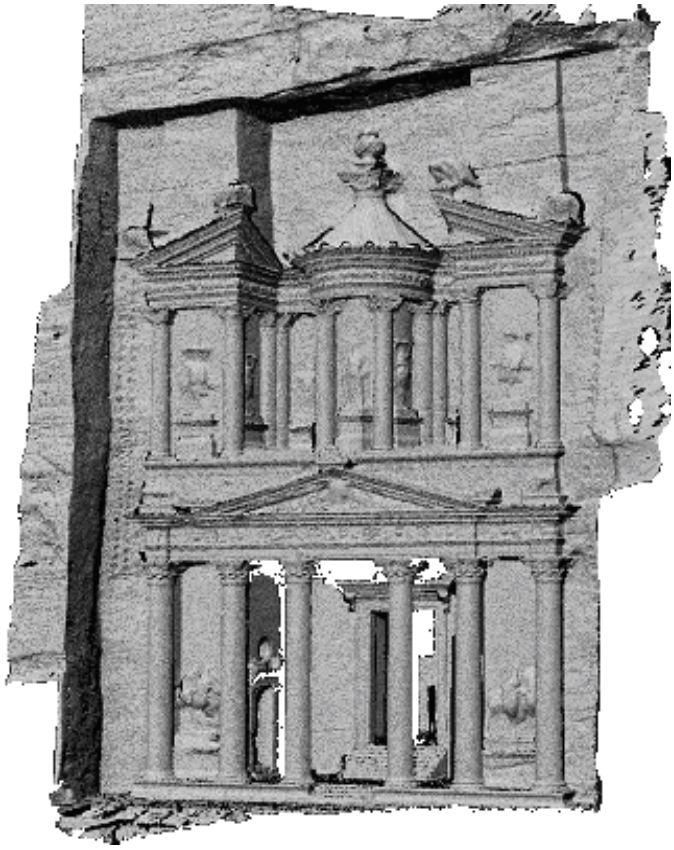


Figure 3.15: 3D triangulated model of Al-Khasneh

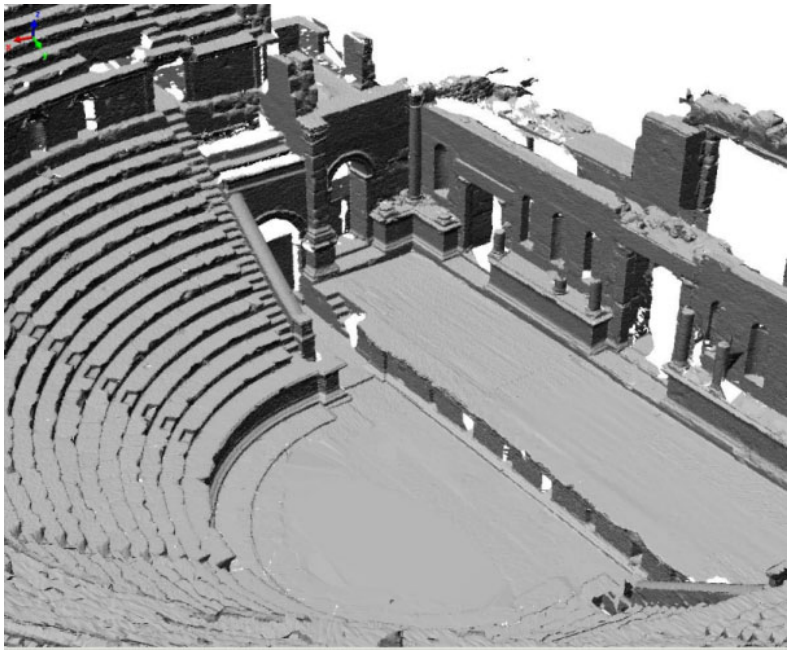


Figure 3.16: 3D triangulated model of the North theatre

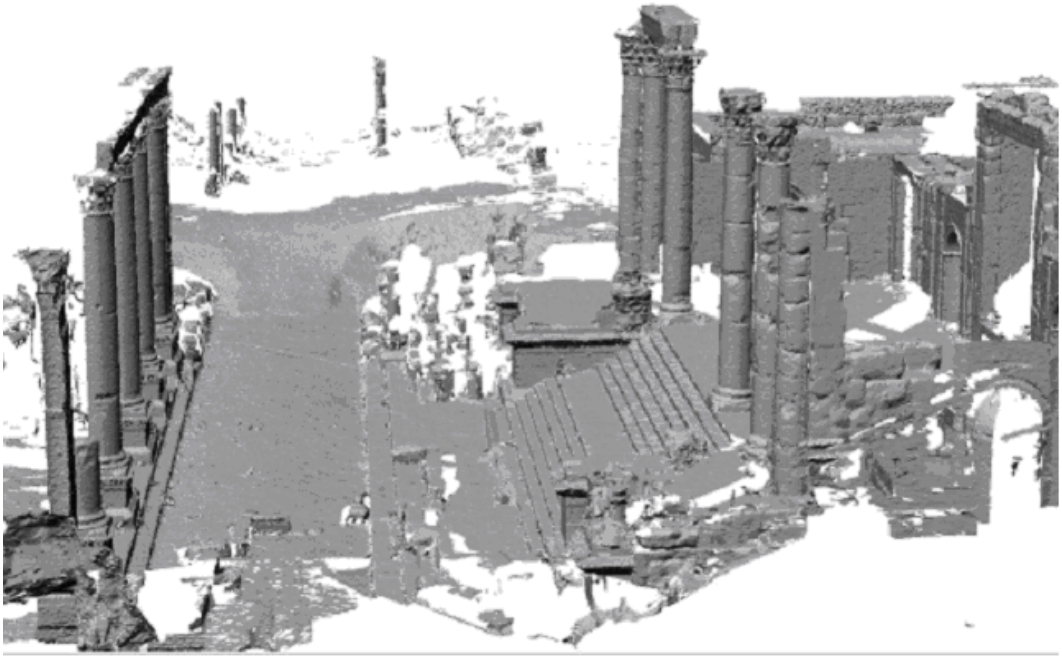


Figure 3.17: 3D triangulated model the North theatre entrance

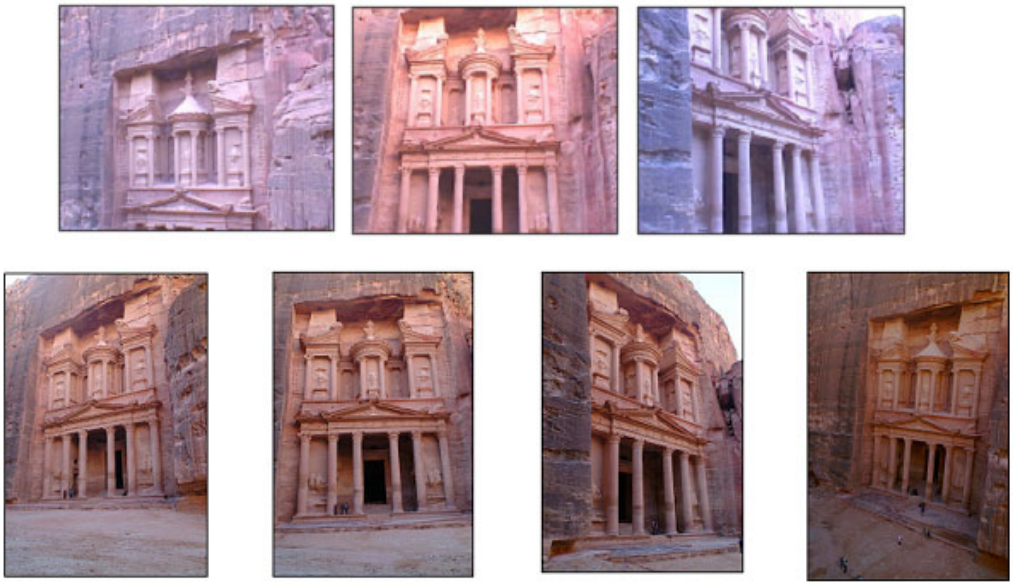


Figure 3.18: In the first row: images collected for Al-khasneh using the camera integrated by Mensi laser scanner system (768x576). In the second row; high-resolution images using Fuji camera (2304x1536)

Figure 3.19: In the first row: images collected for the North theatre using the camera integrated by Mensi laser scanner system (768x576). In the second row; high-resolution images using Fuji camera (2304x1536)

Figure 3.20: 3D textured model of the treasury using the integrated camera (3 scans)

4 Range Image Segmentation

The segmentation algorithms developed for intensity image have been discussed extensively in the literatures [Palmer, 1996]. The algorithms now can apply a real time segmentation of the intensity images; a widely used example is canny operator [Canny, 1986]. In the other hand, ready-made solutions for range image segmentation are not available as it is the case for the intensity images [Gächter, 2005]. Range images are gaining popularity in different fields since they permit the efficient acquisition and representation of 3d information. A long with the intensity images, range images provide the fundamental contribution toward the goal of understanding 3-D shape, which is required for general-purpose object recognition and image understanding. Range image represented by a two dimensional grid of depth values as measured from a single point (perspective) using a range camera. The distances to points imaged are recorded over quantized range. For display purposes, the distances are often coded in grey scale; such as the darker a pixel is, the farther it is to the camera.

Recently, Range image sensors are evolving continuously, allowing nowadays digitizing a full scene with high resolution in a short time. Thus, algorithms and techniques developed to process range images have also to evolve with the same dynamism to enforce the processing of whole information rapidly. The most significant problem in the early stages of image analysis is image segmentation, a process of partitioning of pixels in a 2-D image into meaningful parts in order to extracting important image features. In general, segmentation techniques for range images can be classified into two categories: region and edge based approaches. Region based approach attempts to group pixels into surface regions based on the homogeneity or similarity of surface properties. Although the region-based techniques of range image segmentation bring acceptable results, they have as a common problem the definition of initial seed regions and the determination of the number of clusters [Adams & Bischof, 1994]. On the other hand edge based approach attempts to extract discontinuities in properties that form the closed boundaries of components [Jiang & Bunke 1996]. The main problem of these techniques is that they cannot guarantee closed boundaries due to noise problems. In spite of that edge-based approaches seem to be more attractive option when a large range image has to be processed. Some authors try to solve these problems by means of hybrid approach techniques, edge based segmentation and region-based segmentation, through these hybrid approaches the information provided by the edge detection stage is used to estimate the number of the clustering algorithm. Techniques such those involve high CPU time, and usually the algorithms are difficult to be implemented [Wani & Batchelor, 1994].

Comparatively to the much interest in region segmentation in the literatures, few segmentation works are based on edge based techniques only [Tomita & Kanada, 1984; Hermann, 1985; Fan et al, 1987; Sze et al, 1998]. Most of the researches focused on simple polyhedral objects, and searched for particular structure in the data such as straight lines, circles. In the last decay, long-range laser scanners have become accessible to many users. These scanners can capture accurate 3D range images of real world scene at large scale. These new data set pose much more serious challenges for range image analysis than the traditional polyhedral world. The main difficulties lie in the following aspects: first, natural scene contains many types of objects, which should be represented by various families of surfaces models that have different parameterization dimensions. Second, objects and surfaces in natural scenes appear at multiple scales. Third, although range data is accurate it is rather noisy around object boundaries, depth data are missing for objects at infinite distance like the sky and at metal, glass, and ceramic surfaces where the laser rays never return to the scanner. One of the most significant approaches in real scene edge detection process is the scan line approximation technique [Jiang & Bunke, 1999]. The edge detection method proposed is based on the approximation of a scan lines by set of polynomial functions, then the scan lines are partitioned into curve segments using splitting algorithm. The main drawback is the spurious segments produced by the splitting algorithm causing what is called over segmentation. For this reason [Sappa & Devy, 2001; katsoulas & Werber, 2004] tries to improve the approach by applying a merging steps to deal with over segmentation problem. Recently [Han et al, 2004] present a stochastic jump-diffusion algorithm for the segmentation of range images in the Bayesian framework. Although it is considered as the most acceptable algorithm applied in real scene with satisfactory results there are many drawbacks the author mentioned such as; Complexity and large number of parameters used. In addition the experiments reveal that when the models are not sufficient, the segmentation is not good. Generally, most of based edge work has focused on extracting the edges for matching purposes [Stamos & Allen 2003]. This is because of the fact that images of man-made environment are rich with linear features. The edges of human made objects often can be approximated by straight lines [Lee & Habib, 2002]. Those features can be used to increase the redundancy and improve the geometric strength of photogrammetric adjustments.

In this chapter we present and efficient edge detection algorithm that can detect the line features in both range and intensity images. The purposes are to simplify the dense datasets and provide stable features of interest, which are used to recover the positions of the 2D cameras with respect to the geometric model. In our algorithm the distinguished points, which will comprise the edges, depend on the spatial analysis of the numerical description of the mean curvature values. Although the central task of edge detection is to reliable detect and locate edge points, a rich description of edge points give the ability to reliably detecting and characterizing the edge types as a crease and step edges, and then go further to classify the crease edges as concave or convex types. Compared to known methods in literature, our algorithm exhibits the following features: computational efficiency, high accuracy in the localization of the edge points, easy to implement, and Image noise does not degenerate its performance. The algorithm was initially proposed for range image segmentation and has been extended to segment the intensity images with some improvements. The generality and robustness of the algorithm is illustrated on scene images with different available range sensors.

The outline of the chapter is the following: we begin the chapter by a brief description of surface curvature characteristics. Detailed presentation of the segmentation algorithm is contained in section 2. Section 3 discuses the characteristics and the performance of the algorithm. Section 4 presents a new edge-linking algorithm that developed in order to improve the produced edge map. In section 5 we show experimental results, applying the proposed algorithm to images that come from different range scanners. Section 6 discusses how the proposed algorithm was developed to segment the intensity images.

4.1 Surface Curvature Characteristics

Many complex real scene structures consist of unknown number of objects and surfaces with various sizes and types. The ability to compare these 3-D shapes is necessary in order to identify similarities and quantify the differences between objects, or to search for similar shapes in a database. Comparing surfaces requires metrics techniques, an example is the curvature, which consider as a useful method of describing and classifying surfaces. Principle curvatures and the measures derived from them like mean and Gaussian curvatures are viewpoint invariant. Therefore they play a valuable role in shape representation for tasks including surface segmentation, object recognition and surface reconstruction. Differential geometric concepts for curvature estimation are applicable to the digital surfaces even in the presence of quantization and measurements noise. They provide invariant surface characteristics that are general enough to describe the polyhedral and objects with arbitrary curved surfaces. That is each point on the digital surface can be characterized by the spatial properties of other points on the surface in small neighbourhoods surrounding the given point.

Two questions immediately arise for the defining of the surfaces: how do we define a neighbourhood about an inspection point, and how we do define a robust and fast parameterization method that allows the surface patch to be approximated by polynomial. In fact the researchers concentrates on the issue that involving the parameterization and the curvature derivation of a surface patch using either implicit functions $f(x, y, z)=0$ or explicit functions such as $z= f(x, y)$. Rather, parameterized representations of the surface can express a coordinate of a surface point (x, y, z) as a function of u and v defined on the unit square such that $u, v \varepsilon [0,1]$. Given a parametric representation on a local patch, the values of parameters u and v determine the position of each point on the patch. In the first part of this section we will briefly discuss the geometric characteristics of digital surfaces. The mathematical estimation of the curvature values will be presented the second part. Then in more details we will discuss the properties and the advantages of using of the Mean curvature value for the surface representation.

4.1.1 Geometric Characteristics of Digital Surfaces

In general surfaces can be expressed in an explicit or an implicit form. An explicit form of a surface is the graph of a function of two variables. In context of a computer vision system, the depth maps produced by a range finder can be considered as sample data sets from explicit surface $z= f(x, y)$. Here z is the distance from the camera origin to the surface, and (x, y) are the image plane coordinates. On other hand, the implicit form of a surface in R^3 is expressed in the form

$$f : \Re^3 \rightarrow \Re \mid f(x,y,z) = constant$$

Where (x, y, z) are the Cartesian points on the surface *f* is a function which maps a point from the space of possible surface R^3 to a real number (x, y, z) in the space *R*. In general the explicit form is a special case of the implicit. As an example, the equation of an ellipsoid depicted in figure 4.1 is:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

This can be re-arranged in the explicit form

$$z=c\sqrt{(\,1-\frac{x^2}{a^2}-\frac{y^2}{b^2}\,)}$$

The previous equation describes the position of all points of the surface; there are two solutions for z for a given value of (x, y). However, if we view an ellipsoid from the one side, we cannot see the points on the rear of the surface. Thus, in the definition, graph surface is the single root of equation that corresponds to orthogonal depth data we have obtained from the rangefinder.

We can define a general surface S in explicit parametric form

$$S=\left\{\left(x,y,z\right):x=\ d(u,v),\ y=\ e(u,v)\ ,\ z=\ f(u,v)\mid(u,v)\ \in\ D\subseteq R^2\ \right\}$$

For range and intensity images which have several smooth surfaces separated by points of discontinuity, Monge patch surface can be use to represent the graph surfaces. The monge patch is a patch $X:U\to R^3$ of the form

$$x(u,v)=(u,v,f(u,v))$$

Where U is an open set in R^2 and $f:U\to R$ is a differential function. Thus, the parameterization for a graph surface can take the following simple form:

$$S=\left\{\left(x,y,z\right):x=u\ ,\ y=\ v\ ,\ z=\ f(u,v)\mid(u,v)\ \in\ D\subseteq R^2\ \right\}$$

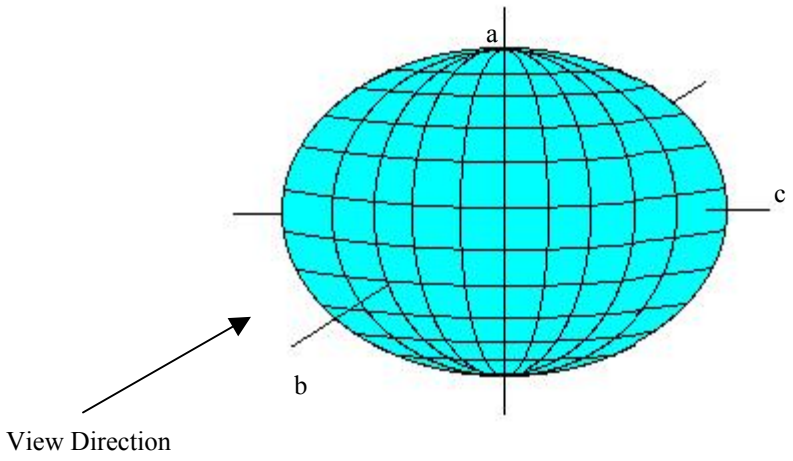


Figure 4.1: Viewing an ellipsoid

In our study, only smooth parametric surfaces are considered in which all three parametric functions possess continuous second partial derivatives. Differential geometry of smooth surfaces depends on two classical mathematics forms. They are known as the first and second fundamental forms of a surface. Complete knowledge of either of these forms at every surface point uniquely characterizes and quantifies general smooth surface shape. The first fundamental form *I* of the surface patch S is given by:

$$I(u,v,du,dv) = dS.dS = Edu^2 + 2Fdudv + Gdv^2$$

$$= \begin{bmatrix} du & dv \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} du \\ dv \end{bmatrix}$$

$$= d\vec{u}^T \; [g] \; d\vec{u}$$

The [g] matrix elements are defined as:

$$g_{11} = E = S_u . S_u \quad , \qquad g_{22} = G = S_v . S_v \quad , \qquad g_{12} = g_{21} = F = S_u . S_v$$

$$S_u(u,v) = \frac{\partial S}{\partial u} \qquad , \quad S_v(u,v) = \frac{\partial S}{\partial v}$$

S_u (u, v) and S_v (u, v) define the tangent vectors of the surface at the point (u, v), and they may or may not be orthogonal to each other. These two tangent vectors are shown in figure 4.2. The matrix [g] is the first fundamental form matrix. The E, F, G notation of Gauss is used along with the matrix element subscript notation because both are useful and often used in differential geometry literature. The first fundamental form *I* (*u, v, du, dv*) is a measure of the small amount of movement $\|ds\|^2$ on the surface at a point (u, v) for a given small vector movement in the parameter plane (du, dv) as shown in figure 4.2. The first fundamental form represents the intrinsic properties of the surface. Which mean it is depend on the surface properties of the surface itself, not how is the surface embedded in the 3D apace. Generally in the differential geometry the functions E, F, G determine all the intrinsic properties of the surface.

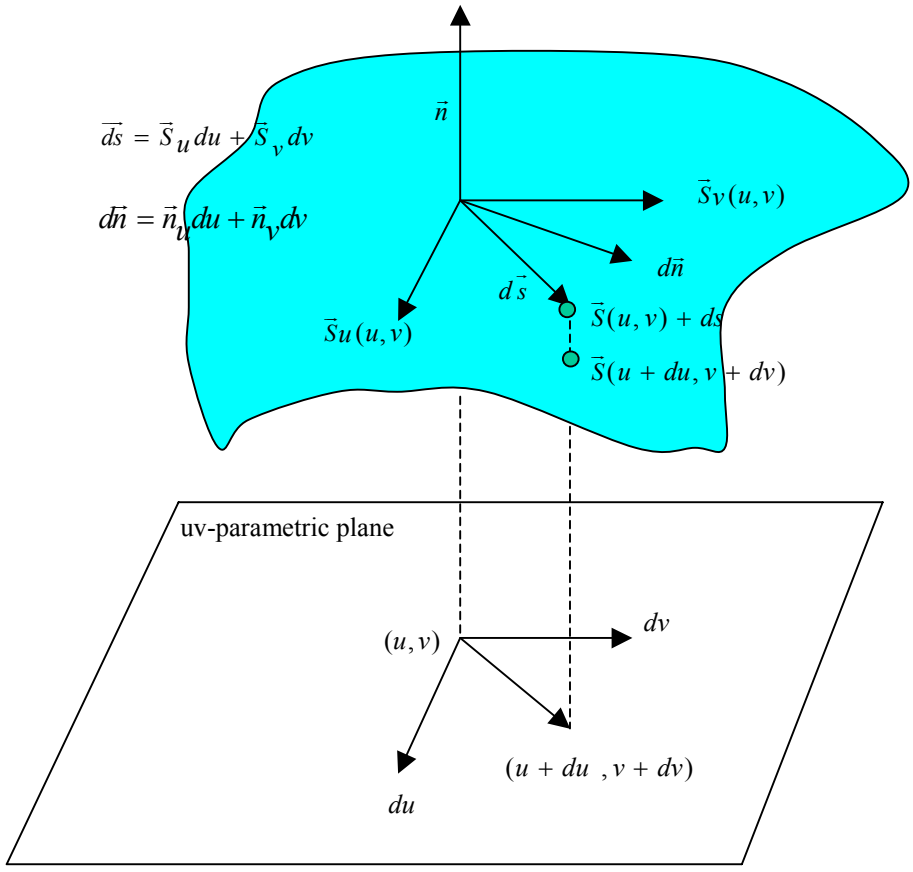


Figure 4.2: The differential geometry of surface patch

In the other hand, the second fundamental form represents the extrinsic property of the surface, and it depends on the position of the surface in the 3D space. The second fundamental form *II* defined as:

$$II(u,v,du,dv) = dS.d\vec{n} = Ldu^2 + 2Mdudv + Ndv^2$$

$$= [du\ dv] \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} [du\quad dv]$$

$$= d\vec{u}^T [b]\, d\vec{u}$$

The matrix [b] elements defined as:

$$b_{11} = L = S_{uu}.\,\vec{n}\ , \qquad b_{22} = N = S_{vv}.\,\vec{n}\ , \qquad b_{12} = b_{21} = M = S_{uv}.\,\vec{n}$$

Where the unit normal vector $\vec{n}(u,v)$ is defined as:

$$\vec{n} = \frac{S_u \times S_v}{\left|S_u \times S_v\right|}$$

The double subscripts denote the second partial derivatives:

$$S_{uu}(u,v) = \frac{\partial^2 S}{\partial u^2} \quad , \quad S_{vv}(u,v) = \frac{\partial^2 S}{\partial v^2} \quad , \quad S_{uv}(u,v) = \frac{\partial^2 S}{\partial u \partial v}$$

The second fundamental form measures the correlation between the change in the normal vector $d\vec{n}$ and the change in the surface position *ds* at surface point *(u, v)* as a function of the direction of the differential vector *(du, dv)* in the parameter space. The ratio of *II (u, v, du, dv) / I (u, v, du, dv)* is known as the normal curvature function *k normal*. This function depends on the direction of the differential vectors (*du, dv*) in the parameter space. If $d\vec{n}$ and *ds* are aligned for a particular direction of (du, dv) that direction is called a principle direction of the surface at that surface point. The extrema of the normal curvature function occur at that point and are known as the principle curvatures, k_1 the maximum and k_2 the minimum.

For example the parametric equations of an ellipsoid can be written as

$$x = \alpha \cos\theta \sin\phi$$

$$y = b \sin\theta \sin\phi$$

$$z = c \cos\phi$$

For azimuthally angle $\theta \in [0, 2\pi]$ and polar angle $\phi \in [0, \pi]$

In this parameterization, the coefficients of the first fundamental form are

$$E = (b^2 \cos^2\theta + a^2 \sin^2\theta) \sin^2\phi$$

$$F = (b^2 - a^2) \cos\theta \sin\theta \cos\phi \sin\phi$$

$$G = (a^2 \cos^2\theta + b^2 \sin^2\theta) \cos^2\phi + c^2 \sin^2\phi$$

And of the second fundamental for are:

$$L = \frac{abc \sin^2\phi}{\sqrt{a^2 b^2 \cos^2\phi + c^2 (b^2 \cos^2\theta + a^2 \sin^2\theta) \sin^2\phi}}$$

$$M = 0$$

$$N = \frac{a^2 b^2 c^2}{\sqrt{a^2 b^2 \cos^2\phi + c^2 (b^2 \cos^2\theta + a^2 \sin^2\theta) \sin^2\phi}}$$

4.1.2 Surface Curvatures

In the previous section we have shown that the arbitrary smooth surface shape is captured by the six scalar functions E, F, G, L, M, N. Where this information can be reduced into two-curvature functions k_1, k_2. It is possible to combine functions to yield more easily interpretable surface shape characteristics. Such as the mean curvature and the Gaussian curvature. They are direction independent quantities and can be computed by combining the two-direction dependent principal curvatures k_1 and k_2 as follow:

$$K = k_1.k_2$$

$$H = \frac{k_1 + k_2}{2}$$

The principal curvatures k_1 and k_2 are the two roots of the quadratic equation:

$$k^2 - 2Hk + K = 0$$

So, if K and H are known at each point in the range image, the two principle curvatures determined as follow:

$$k_{1,2} = H \pm \sqrt{H^2 - K}$$

In fact both k_1 and k_2 contain the same surface curvature information as K and H, but in different form, and the advantages and disadvantages working with either pair depends on the application.in orderTo compute the mean and Gaussian surface curvatures from digital images, we need the five partial derivatives S_u, S_v, S_{uu}, S_{vv}, and S_{uv}:

$$H = \frac{S_{uu} + S_{vv} + S_{uu}S_v^2 + S_{vv}S_u^2 - 2S_uS_vS_{uv}}{2(1 + S_u^2 + S_v^2)^{3/2}}$$

$$K = \frac{SuuSvv - S_{uv}^2}{(1 + S_u^2 + S_v^2)^2}$$

For the ellipsoid shown in figure 3, the Gaussian and Mean curvature will be

$$K = \frac{a^2b^2c^2}{\sqrt{[a^2b^2\cos^2\phi + c^2(b^2\cos^2\theta + a^2\sin^2\theta)\sin^2\phi]^2}}$$

$$H = \frac{abc[3(a^2 + b^2) + 2c^2 + (a^2 + b^2 - 2c^2)\cos(2\phi) - 2(a^2 - b^2)\cos(2\theta)\sin^2\phi]}{8\,[a^2b^2\cos^2\phi + c^2(b^2\cos^2\theta + a^2\sin^2\theta)\sin^2\phi]^{\frac{3}{2}}}$$

4.1.3 Properties of Mean Curvature Values

To generate useful descriptions of the surface, a useful representation is needed. In general, such a description should be suitable for the tasks of segmentation. It should be rich so that similar objects can be identified. Stable so that local changes do not radically alter the descriptions, and have a local support so that the visible objects can be easily identified. These criteria naturally lead us to a successful segmentation process. The mathematical properties of the mean curvature are discussed in details to stress the importance of these quantities to surface characterization. As the name suggests, it is the mean value of the curvature over all possible directions. Mean curvature is directional free quantities, equal to the sum of the principle curvatures k_1 and k_2. It has the dimension of (1/length) and it is closely related to the first variation of surface area. Unlike the Gaussian curvature, the mean curvature depends on the embedding, for instance, a cylinder and a plane are locally isometric but the mean curvature of a plane is zero while that for a cylinder is non-zero. Comparatively with the principle curvatures the mean curvature is invariant to arbitrary rotations and translation of surface. Which mean that the visible surface in range images have the same mean curvature from any viewpoint under orthographic projection. This invariance necessary for surface shape characterization. Additionally, mean curvature is the average of the principal curvatures, thus, it is slightly less sensitive to noise in numerical compu-

tations than the principal curvatures. Depending on the previous characteristics, the numerical values of the mean curvature could be used as a useful measure for detecting surface irregularities in range and intensity image.

4.2 Range Image Segmentation Algorithm

Previous efforts in edge detection reveal that an optimal edge operator depends on how it fits the nature of edges embedded in the image. It is known that there is a great deal of diversity in the requirements of edge extracted from images. For most image analysis processes used edge information, it is required that the edge should be presented correctly. That is, edge points in the image must present actual (true) edges. Additionally, the extracted boundaries should be represented in a certain width, not too thin and not too thick. Finally the edges should have a certain degree of continuity that is; edge points belonging to one edge segment should form a continuous line and one edge segment should have connectivity with some other edge segments. These connections then form closed boundaries of the objects in the images. Typically, the researchers classify the edges in 3D image in two main categories; step (jump) edges: represents discontinuities on the surface. Crease (fold) edges: correspond to the discontinuities of the surface normals. Some researchers consider another type of edges called ridgelines edges, example of this type are the edges a long the extrema of the major axis of the cross section of an elliptical cylinder.

For the intensity image, the devised segmentation algorithms usually aim to detect step edges. This is due to the natural limitation of this type of images. Thus the fact of using intensity image algorithms for range image segmentation will not work. Figure 4.3 depicts the result of using canny operator for segmenting curve block range image. For range image since depth information is available, it is possible in this case to detect both the step and crease edges. But it is worth to out line that crease edges are usually the most difficult to detect since information is usually embedded in noise. The edge detection technique used in this study search to detect all types of edges in 3D range images achieving the criteria of optimal edge detectors mentioned above. The algorithm depends on computing the mean curvature value at every image point, where a threshold related to the mean curvature values is used to classify edge points. Using of multiple scales allows us to increase our confidence in estimating the mean curvature values in noise image and detecting features without loss of localization. Large-scale mask may result a thick edge, so some morphological operations such as image resizing and skeletonized is necessary in order to have the final segmentation with one pixel wide edge. The method has been extensively tested on a large number or synthetic and real range images acquired by different range scanners with quite different characteristics and demonstrated good results. The following section will present the general methodology of the segmentation algorithm. In more details the mathematical aspects used for estimating the mean curvature values and its spatial analysis behavior will be discussed in sections 2 and 3. Section 4 presents an automatic method for selecting threshold values for mean curvatures used to detect the edges in the range images. Then morphological operations for correcting detection of the corners and trihedral vertices will be discussed along with the skeletonized operation in sections 5 and 6.

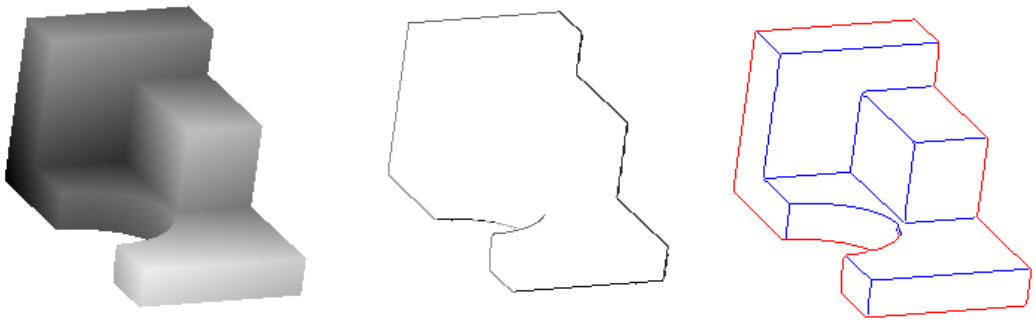


Figure 4.3: Curve block range image (OSU). Mid image shows segmentation result using canny operator. The right image shows the types of edges in the range image; crease edges (blue) and step edges (red).

4.2.1 Methodology

The approach is developed from a theoretical model based on the analysis of classical differential geometry of 3D surfaces. The idea behind the proposed algorithm is to assess the curvature at each pixel in the given image; the surface is locally approximated by an analytic representation (Mango surface patch). The different properties of the patch at the point of interest are then calculated analytically and assigned back to the point. The process is repeated for each pixel in the range image. Then the approach tries to find the local maximal or zero crossing of the mean curvature in a range image. So, if we assume that the scanner delivers a set of coordinate in a cloud of *(x(i), y(i),z(i))* for *i= 1,……,N,* data points sampled from the body surface. Our aim is to find the form of the surface within some small neighborhood $N(s)$ of a point s on the surface to estimate the curvature there. To do so, we choose a local analytic approximation to the surface and fit it to the data points in N(s). The parametric representation: $x = x(u,v), y = (u,v), z = (u,v)$ is preferred in differential geometry. The steps of the algorithm are as follows: For every point P in the set:

a) A neighborhood of points (size N) around P is selected.

b) A mango patch f is fitted onto the neighborhood of points.

c) The mean curvature properties of f are calculated at P and assigned back.

d) An automatic threshold is used to help detecting and classifying the diseconomies in the range image.

e) Further processing using multi scale approach and skeletonized operation used to increase the confidence in detecting and localizing the edges.

For the proposed algorithm to work, the following assumptions must be valid:

a) The input data set is a point cloud interpolated in a sampled range image. It is not necessary for the point cloud to be connected, although the use of connectivity information wherever available would improve the generation of the range images.

b) The point cloud represents a significant coverage of the original surface. Since no curvature is calculated at areas of the object surface not covered by the point cloud.

c) The surface must be sufficiently sampled; the point cloud must be sufficiently dense.

4.2.2 Mathematical Properties

This section presents techniques for efficiently estimating mean curvature values at multiple scales. Since we work with digital data, mean curvature values can be obtained from analytic method by fitting analytical surfaces to a given data. The general strategy of the analytical methods is to fit a surface in a local neighborhood of the point of interest and then compute the partial derivatives needed to determine curvatures. The main differences in the analytic methods are in the method for the local surface fitting. In our algorithm we will use Besl and Jain proposed technique for estimating the mean curvature. The advantage of this approach is the potential speed of the process using optimized convolution operations. A local least squares surface fit is computed within the (N X N) window around each pixel of the digital surface using separable convolutions. This method is based on a local least squares surface model using discrete orthogonal polynomials, which provide local biquadrate surface fitting capability. Consider a local parametric representation of the range image with coordinates u and v the procedure for estimating the mean curvature is as follows; for a given odd N x N window each data point is associated with a position (u, v) from the set U x U where U described as the following;

$$U=\{-(N-1)/2,\ldots\ldots\ldots,-1,0,1,\ldots\ldots\ldots\ldots,(N-1)/2\}.$$

The local biquadratic surface fitting capability is provided using the following discrete orthogonal polynomials:

$$\emptyset_0(u)=1, \qquad \emptyset_1(u)=u \ , \qquad \emptyset_2(u)=(u^2 - M(M+1)/3) \qquad \text{Where } M=(N-1)/2.$$

To estimate the first and second partial derivatives, orthogonal set of di(u) functions using the normalized versions of the orthogonal polynomials $\emptyset_i(u)$ is used.

$$\vec{d}_i(u) = \frac{\emptyset_i(u)}{P_i(M)} \ .$$

Where the Pi (M) are normalizing constants define as:

$$P_0(M) = N.$$

$$P_1(M) = \frac{2}{3}M^3 + M^2 + \frac{1}{3}M.$$

$$P_2(M) = \frac{8}{45}M^5 + \frac{4}{9}M^4 + \frac{2}{9}M^2 - \frac{1}{9}M^2 - \frac{1}{15}M.$$

Since the discrete orthogonal quadratic polynomials over the 2D window are separable in u and v, partial derivative estimates can be computed using separable convolution operators. This is more efficient on general-purpose computer. These derivatives estimates can then be plugged into the equation for mean curvature. The equally weighted least squares derivative estimation window operators are given by:

$$[D_u] = \vec{d}_0\vec{d}_1^T \quad [D_v] = \vec{d}_1\vec{d}_0^T \quad [D_{uu}] = \vec{d}_0\vec{d}_2^T \quad [D_{vv}] = \vec{d}_2\vec{d}_0^T \quad [D_{uv}] = \vec{d}_1\vec{d}_1^T$$

Where the column vectors $\vec{d}_0$, $\vec{d}_1$, $\vec{d}_2$ for 5 x 5 window are given by

$$\vec{d}_0 = \frac{1}{5}[1\ 1\ 1\ 1\ 1]^T \quad , \quad \vec{d}_1 = \frac{1}{10}[-2\ -1\ 0\ 1\ 2]^T \quad , \quad \vec{d}_2 = \frac{1}{14}[2,-1,-2,-1,2]^T$$

If $\tilde{g}(i,j)$ represents noisy, quantized discretely sampled version of piecewise-smooth graph surface. Then the partial derivative estimate images are computed via appropriate 2-D image convolutions.

$$\tilde{g}_u(i,j) = D_u \otimes \tilde{g}(i,j) \quad , \quad \tilde{g}_v(i,j) = D_v \otimes \tilde{g}(i,j) \quad , \quad \tilde{g}_{uu}(i,j) = D_{uu} \otimes \tilde{g}(i,j) \quad ,$$

$$\tilde{g}_{uv}(i,j) = D_{vv} \otimes \tilde{g}(i,j) \quad , \quad \tilde{g}_{uv}(i,j) = D_{uv} \otimes \tilde{g}(i,j)$$

The mean curvature are computing using the partial derivatives estimate images as the following:

$$H(i,j) = \frac{(1+\tilde{g}_v^2(i,j))\ \tilde{g}_{uu}(i,j) + (1+\tilde{g}_u^2(i,j))\tilde{g}_{vv}(i,j) - 2\tilde{g}_u(i,j)\ \tilde{g}_v(i,j)\tilde{g}_{uv}(i,j)}{2(\sqrt{1+\tilde{g}_u^2(i,j)+\tilde{g}_v^2(i,j)})^3}$$

During the operations of mean curvature estimation, the mentioned masks are slid over an area of the input image, change that pixels value then shifts one pixel to the right and continues to the right until it reaches the end of a row. It then starts at the beginning of the next row. The center of the mask is placed over the pixel you are manipulating in the image. It is important to notice that pixels in the border of the image (first and last rows as well as the first and last columns) cannot be manipulated. This is because when placing the center of the mask over a pixel in the first row for example the mask will be outside the image boundaries.

4.2.3 Mean Curvature Spatial Analysis

Having computed mean curvature values at each vertex of the range image, the spatial behaviour of the curvatures along a cross section in the surface can be plotted. Figure 4.4 shows the spatial distribution profile of the mean curvature values for block and bigwye range images at selecting cross sections. The cross section is chosen normal to the significant line on the surface. From the analysis of the figures it was found the followings:

a) For Jump boundaries (J) where surface depths are discontinue, the mean curvature exhibits a zero crossing and two peaks of opposite sign moving a way from each other as shown in figure 4.5a. When we apply differential operators to images, the zeros rarely fall exactly on a pixel. Typically, they fall between pixels.

b) For crease edge (C) the curvature response is a smooth peak and its sign depend on the orientation of the faces; concave (Cv) or convex (Cx), as depicted in figure 4.5b and 4.5c. The position of the convex crease edges in the positive sign where as the position of the concave crease edge is in the negative one.

c) For jump edge, the position of the true edge is in the mid of the two peaks of the mean curvatures. The position represents zero crossing surrounded by non-zero numbers of opposite sign. Whereas for both crease and ridge edges the true edge represents the maximum value for the peaks.

d) Mean curvature values for the crease edges are smaller comparatively with the jump edges, this depend mainly on the magnitude of the depth discontinuity.

e) For the ridge edges (R) the Mean curvature response like the crease edges but with small values.

After computing the mean curvature values for all the pixels in the image. The pixel location is declared an edge location if the value of the gradient exceeds some threshold. The edges will have higher pixel range values than those surrounding it. So once a threshold is set you can compare the mean curvature values to the threshold value and detect an edge whenever the threshold is exceeded.

$$E(x,y) = \begin{cases} 1 \; if \left\| H \right\| > T \; for \; some \; threshold \; T \\ 0 \; otherwise \end{cases}$$

$\{(x,y) : E(x,y) = 1\}$ The set of edge pixels.

The output results can be depicted in figure 4.6 after we have used 5x5 widows (mask size for estimating the mean curvatures values). It can be seen that the algorithm locates the position of the step and crease edges. Since we have used mask size 5, it is expected to locate the true position of the edge along with some positions around it. Although the algorithm localized the edges in a continuous shape, some of junctions and corners (represented in red arrows in figure 4.6) are still missing because of the presence of another close edges in its neighborhood which lying within the mask size. How to get rid from this problem and how to compute the optimal threshold value will be discussed in following sections.

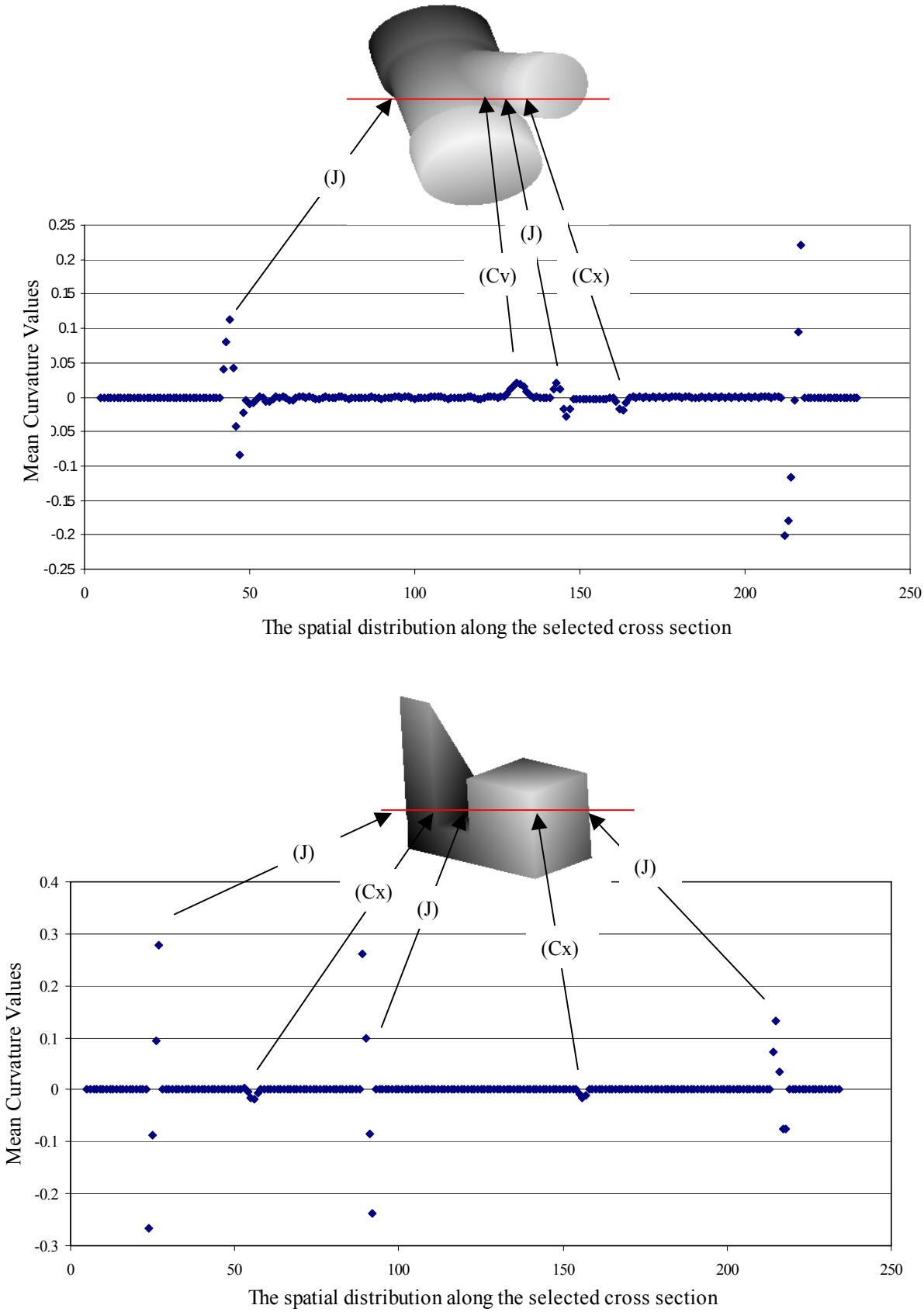


Figure 4.4: Spatial distribution of the Mean curvature values for the showing cross sections in both bigwye and block range images using mask size 5.

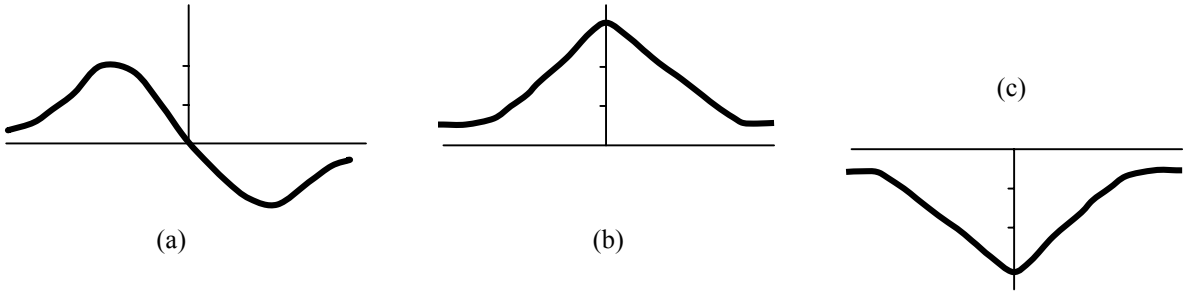


Figure 4.5: (a) zero crossing of mean curvature represents step edge in the range image, the true edge position is the pixel in the mid of the two surrounding peaks (J). (b) Local maxima of mean curvature represents crease edge or ridgeline, true edge position is the pixel in the mid (correspond to the peak value), the positive peak indicates concave edge type (Cv). (c) Local maxima of mean curvature represents crease edge or ridgeline, true position is the pixel in the mid (correspond to the peak value), negative peak indicates the convex edge type (Cx).

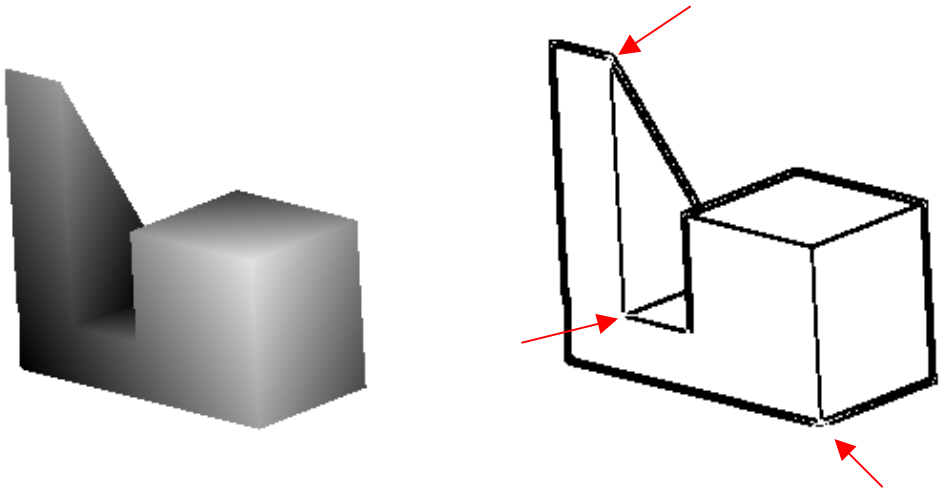


Figure 4.6: Depth and orientation discontinuities detection using mask size 5, the red arrows show the missing parts of the edges in the junctions and the corners of the object

4.2.4 Selecting Thresholding Values

The purpose of a thresholding operation is to exclude the false edges and to enhance true edge response. Once we have calculated our derivatives, the next stage is to apply a threshold to determine where the results suggest an edge. The lower the threshold the more lines will be detected and the results become increasingly susceptible to noise. Conversely a high threshold may miss important information. Thus, the output of the thresholding stage is extremely sensitive and the selection of the best threshold value is usually a disturbing problem. This is due to the insufficient knowledge about the mean curvature values for the features in the range image. Mean curvature values vary and depend on object surface types, object scale and the mask size used for estimating the curvatures. To avoid wasting time in finding a suitable threshold value and towered automatic threshold processing, our approach presents initial threshold values for each processed image to guide the selection of the optimal threshold value. The algorithm estimates the candidate threshold values as follow;

a) Estimating the mean curvature values across the mid of the image or in any selected cross section in the image contains the majority of the surface types.

b) The differences in the mean curvature values are estimated to locate zero crossings at all types of edges; jump, crease and ridgelines, see figure 4.7.

c) Around each of the zero crossing position, the largest mean curvature values around those positions are found and sort in a descending way, P {p0, p1, p2……..pi-1}.

Typically, you can start by one of the candistes threshold values, and depend on the output results you can decide to select the next candidates. In this process some times losing some edges and some times getting many extraneous background pixels. Figure 4.8 depicts examples for the bigwye using different values of thresholds.

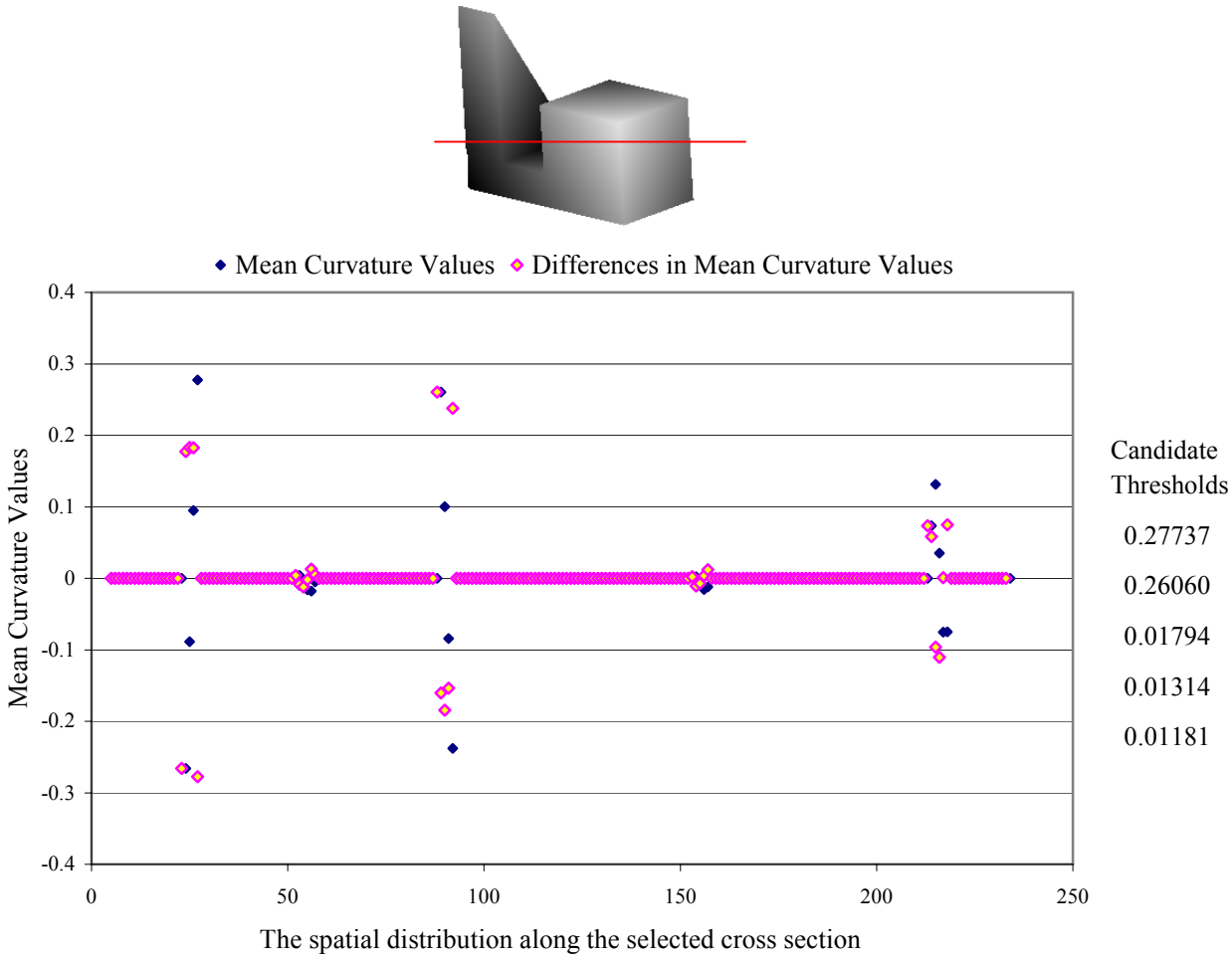


Figure 4.7: Spatial distribution of the differences in the mean curvature values for a cross section in the block range image in order to determine the candidate threshold values.

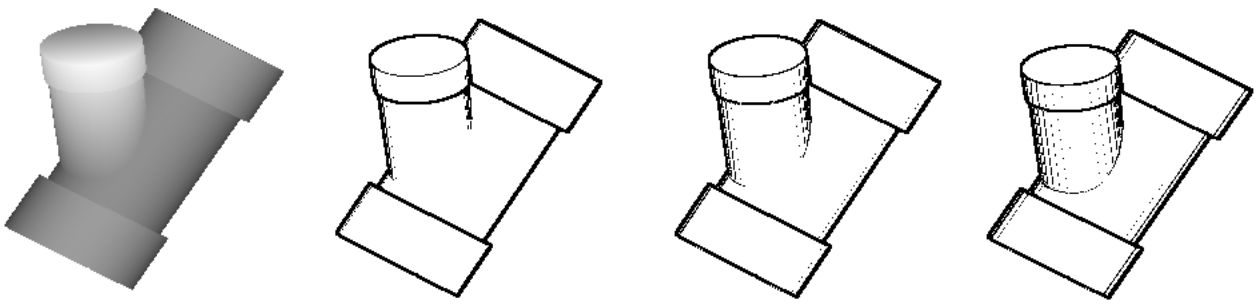


Figure 4.8: Bigwye segmentation using different values of candidate's thresholds.

4.2.5 Multi-scale Approach

No edge operator with a single scale can perform well on any edge image since the effectiveness of an edge operator is also related to the signal-noise ratio of the image. It is often necessary to select a proper scale for the operator to detect the edges reliably. Mean curvature, being a local measurement, is a highly noise sensitive. So larger masks to compute the partial derivatives may be used to have reliable estimation. By applying different sizes of edge operator on an image, we can get a description of the signal changes at different scales. Every setting of the scale parameter yields a different description where new extreme points may appear. For small-scale operator we get fine details depth changes whereas for a large-scale operator we get course depth changes. In case of using large mask sizes, it has to be noticed that in the image there are more than one edge curve. These edges either intersect or close to each other. Thus, if there is only one edge curve on the operator disk, then the operator will detect the edge correctly. But if there is more than one edge curve on the operator disk, the edges curve will affect each other. This will result disappearing of parts of the detecting edge, as can be depicted in the figure 4.6. It can be seen that the junctions and corners are missing because of the presence of another close edges in its neighborhood lying within the mask size. T-junctions and corner detection is an important task in various computer vision and image understanding systems. Much work has been carried out on corner detection [DeMicheli et al, 1989; Rothwell et al, 1995].

In order to compromise between edge localization and noise sensitivity, we have to handle the scale thresholds parameters to a case the size of a finite operator disk contains only one edge curve. Generally, the operator scale is adjusted according to the objective evaluation of the resulting edge image. In our perevious example, we are handling the junction problems using different scale threshold parameters. The result of the segmentation is shown in figure 4.9. It can be seen that the missing parts are recovered. In general, the problem of choosing the suitable scale is strongly knowledge problem depends on analysis of the operator behavior at different types of images (synthetic, real close range images or airborne images).

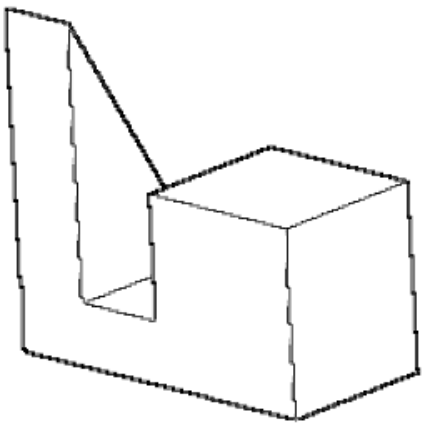


Figure 4.9: Handling the junction problems using different scale threshold parameters

4.2.6 Skeleton

For scale dependant operator every setting of scale parameterizes yield a different description, new extremal point may appear and existing ones may be disappeared. Thus, the thickness of edge curve will expand as the scale increases. Edge should be properly represented, to evaluate edge strength one of criteria is that the extracted edges should be continual and connected thin line in a binary image. Edge width stands how the edge points are concentrated to their true location. This mean that in general only one pixel is seen at the cross section of an edge line. As it is discussed in the spatial analysis section, the position of the true edge is in the mid of the two peaks of the mean curvature for the jump edge and represent the maximum value (mid point) for the peaks in the crease and ridge edges. One way to find the middle of something is through the thinning process, and what is left from the thinning is a structure called a skeleton. Thus we benefit from this fact to yield one pixel wide edge of the results from the multi-scale operator. Figure 4.10 below shows the features detected with one pixel precision using skeletonized process.

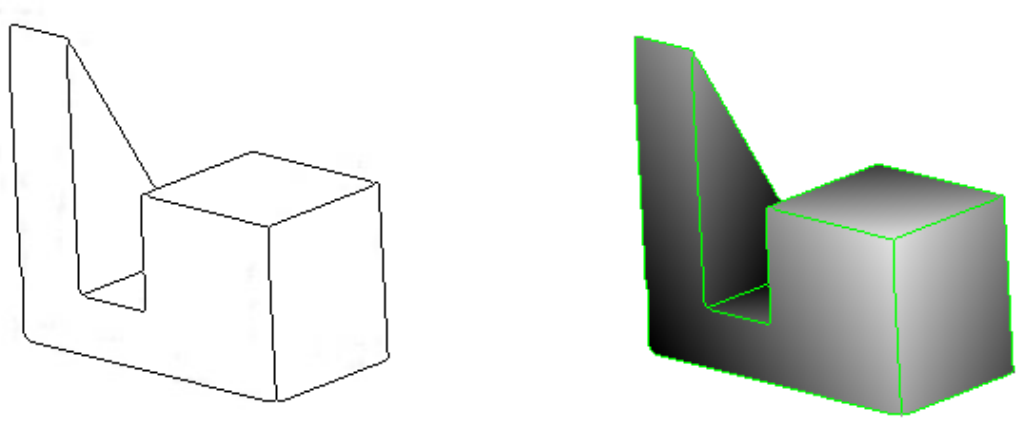


Figure 4.10: The segmentation result after skeltonized process to yield one pixel wide edge

4.3 Algorithm Characteristics

Good edge detection required that the operator be designed should fits the nature of a specific image. Beside this fact it was found that there are many other interesting characteristics of the proposed edge detector performance related to the nature and properties of the mean curvature values. The following sections will describe some of these characteristics; first section discuss the ability and the performance of the developed algorithm to reliably detecting and characterizing the edge types as a crease and step edges, and then go further to classify the crease edges as concave or convex types. The immunity against the noise will be discussed in details in section two. The ability of the proposed algorithm to detect the edges of the object in all possible posses using the same threshold value will be discussed in section 3. The last section shows how the edge detection technique deals successfully with free form objects.

4.3.1 Edge Classification

In spatial analysis section of the mean curvature values, we have mentioned a fact that " *the mean curvature values for the crease edges are smaller comparatively with the jump edges, and this depend mainly on the magnitude of the depth discontinuity*". Depending on this definition, it will be possible to classify the edge types of the object by applying different values of thresholds. Thus, a low threshold value can be selected for detecting the small peaks represented the crease edges and high value can be used for step edge detection. The examples in figures 4.11 below demonstrate the ability of our algorithm to reliably characterized edge types.

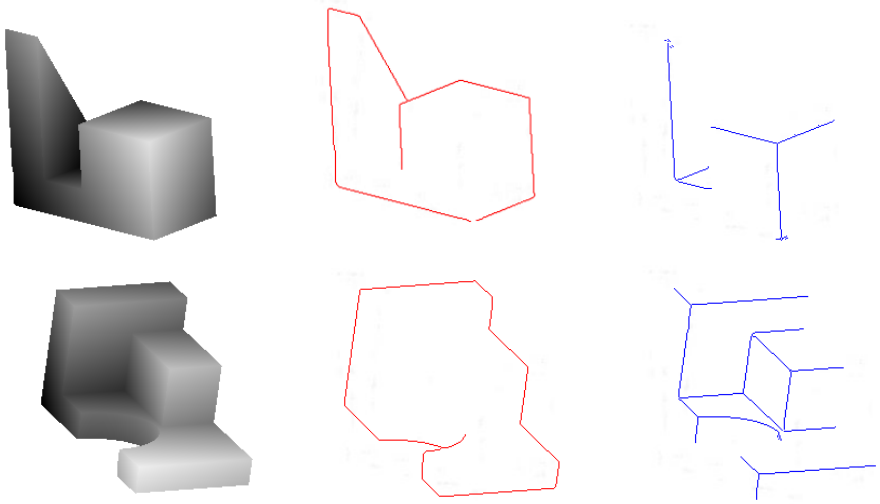


Figure 4.11: Block and curve block range image segmentation using different threshold values to detect step edges (red) and the crease edges (blue).

4.3.2 Noise Immunity

The first challenge in most of segmentation algorithms is the measurement of noise. Although the fidelity of scanners has improved dramatically over the past decade, noise is ever present in any practical system. The sampling quality is dependent not just on sensor accuracy but also surface properties and viewing angles [Vitulano, fast.seg]. Since for a second order derivative the numerical values will be more sensitive to noise. Thus, edge detection algorithms based on numerical differentiation are ill posed algorithms. Most edge detection techniques apply a smoothing process before the extraction of range changes. The raw image is convolved with a smoother operator like Gaussian mask. The smoothing filter used in the first stage affects the results of the edge detector. Smaller filter causes less blurring, and allow detection of small, sharp lines. A larger filter causes more blurring, smearing out the value of given pixel over a large area of the image. One of the costs of the smoothing is the possible cost of accuracy of localization. In addition most of the algorithms presented in the literature are sensitive to the noise and smoothing will crucially affect the results algorithms, some of them totally fail if the smoothing of parameters mask is not suitable [Besl, 1988; Hoover et al, 1996; Böhm & Fritsch, 2003].

During the discussion of the properties of the mean curvature value, we have mentioned a fact that " *Because it is the average of the principal curvatures; mean curvature is slightly less sensitive to noise in numerical computations*". In order to examine the immunity of the detector against the noise, we have tested the behavior of our algorithm in a noise image generated by adding more than 220 randomly positioned noisy points to the original image of the curve block, as can be seen in the figure 4.12. The range value of the noisy point is set to z max if the original value is larger than (zmin+zmax)/2, and set to z max otherwise. The z min and z max is the minimum and the maximum range value of the original image [Jiang & Bunke, 1999]. After applying the segmentation algorithm on the noisy image, it can be seen that image noise does not degenerate the performance of our edge detection algorithm. Addition examples on noise real images are presented in the experimental section using terrestrial and airborne laser range images.

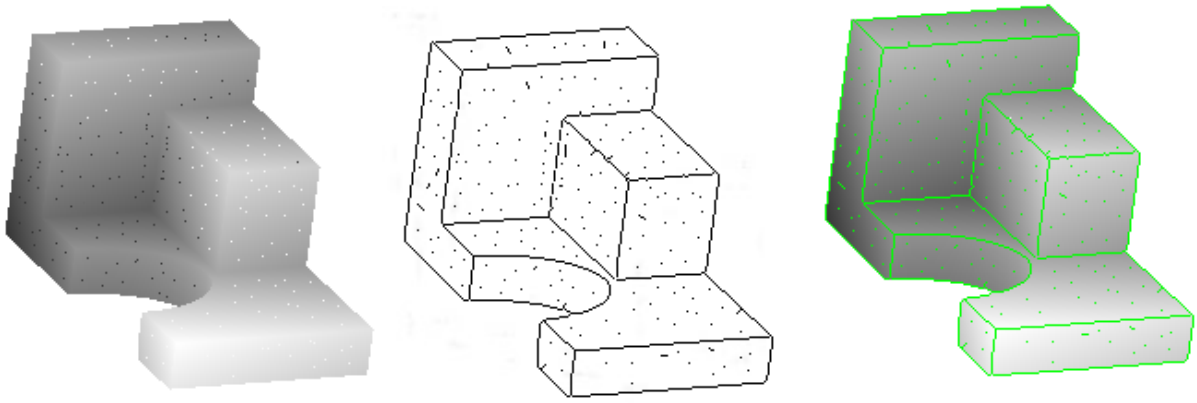


Figure 4.12: A sparse image for the curve block generated using 220 randomly positioned noisy points. From the segmentation results it can be shown that image noise does not degenerate the performance of the proposed edge detection algorithm.

4.3.3 Pose Independent:

Range image object recognition usually decomposes into surface recognition problem, which is in turn relies on surface characterization. Segmentation process depends mainly on those characteristics. Viewpoint invariance is a desirable property for any surface characteristics derived from discrete image data. A quantity is invariant with respect to a group of transformations if those transformations do not change its value. *Mean curvature value is invariant to arbitrary rotations and translations of the surface because of the invariance of the E, F, G, L, M, N functions to rotation and translation*. Thus, we can expect that the algorithm can reliably detect the edges independent of the pose of the object. The performance of our technique for curvature estimation and edge detection is tested using range images with different poses. It was found that the algorithm could find the edges in all object directions using the same threshold value. The result not affected by the arbitrary direction and rotation of the surfaces. Examples are shown in figure 4.13 for the bigwye range images and in figure 4.14 for the block range images.

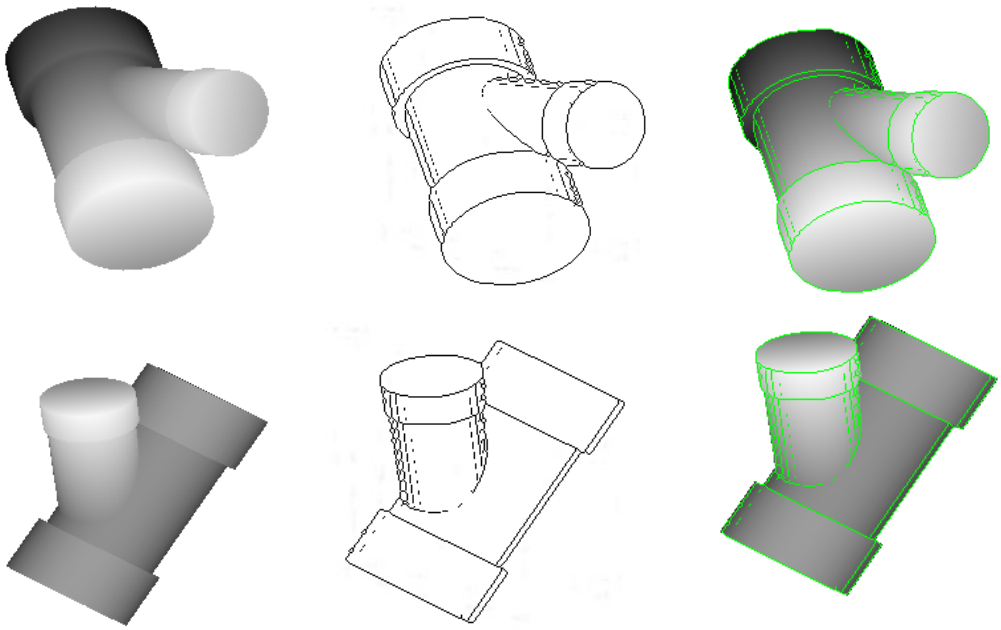


Figure 4.13: Bigwye collection segmentation with different poses

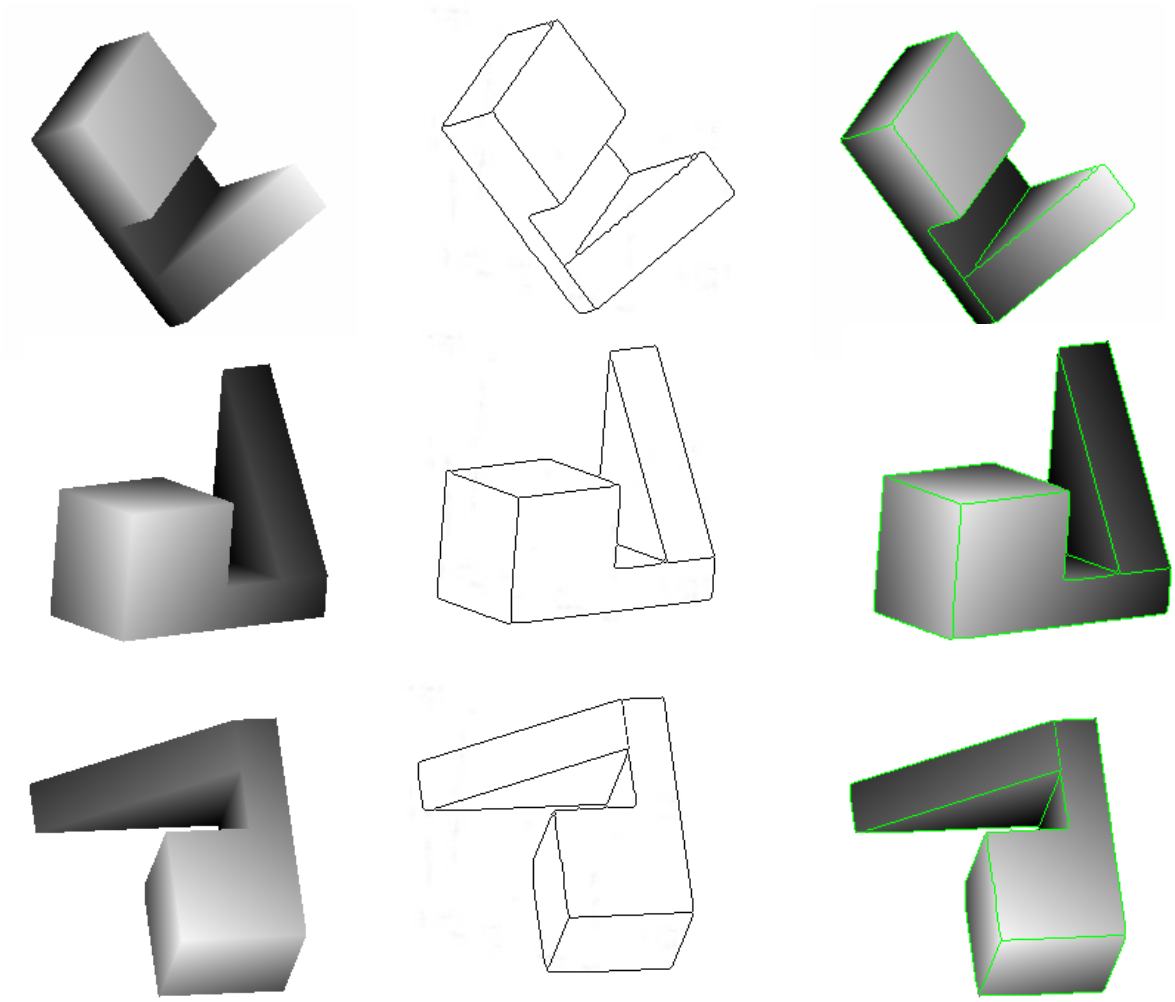


Figure 4.14: Block collection segmentation with different poses

4.3.4 Free Form Objects

A free form surface has a well-defined surface normal that is continues almost everywhere except at the edges and cusps. Discontinuities in surface normal may be presented any where on the free form surface. Similarly, discontinuities in the surface depth may be present any where in a projection of the object. The curves that connect these points of discontinuities may meet or diverge smoothly [Dorai, 1997]. Human faces and sculptures are typical examples of free form objects. It is widely accepted in the researches that the reliable segmentation and recognition of arbitrary viewed complex curved objects is still a challenging task [Campbell, 2001]. This is because of the lack of clearly defined and stable boundaries to the segments. Many techniques for free form object segmentation were proposed [Benlamri, 2003]. Typically, edge based techniques which attempt to extract closed boundaries of components by detecting continuity in both depth and surface orientation, present difficulties when dealing with incomplete broken edges. Region based techniques on the other hand, attempt to cluster surfaces based on their intrinsic differential geometric properties [Ernest & Wu, 1992], but also present difficulties when dealing with noise, occlusion and surface continuity at the boundaries. This is because free form objects are not entirely smooth but piecewise smooth. The proposed edge detection technique has been tested when free form objects considered. Figure 4.15 give examples using sculptured objects obtained from OSU (MSU/WSU) database. The obtained results prove that the presented technique also can deal with this kind of objects.

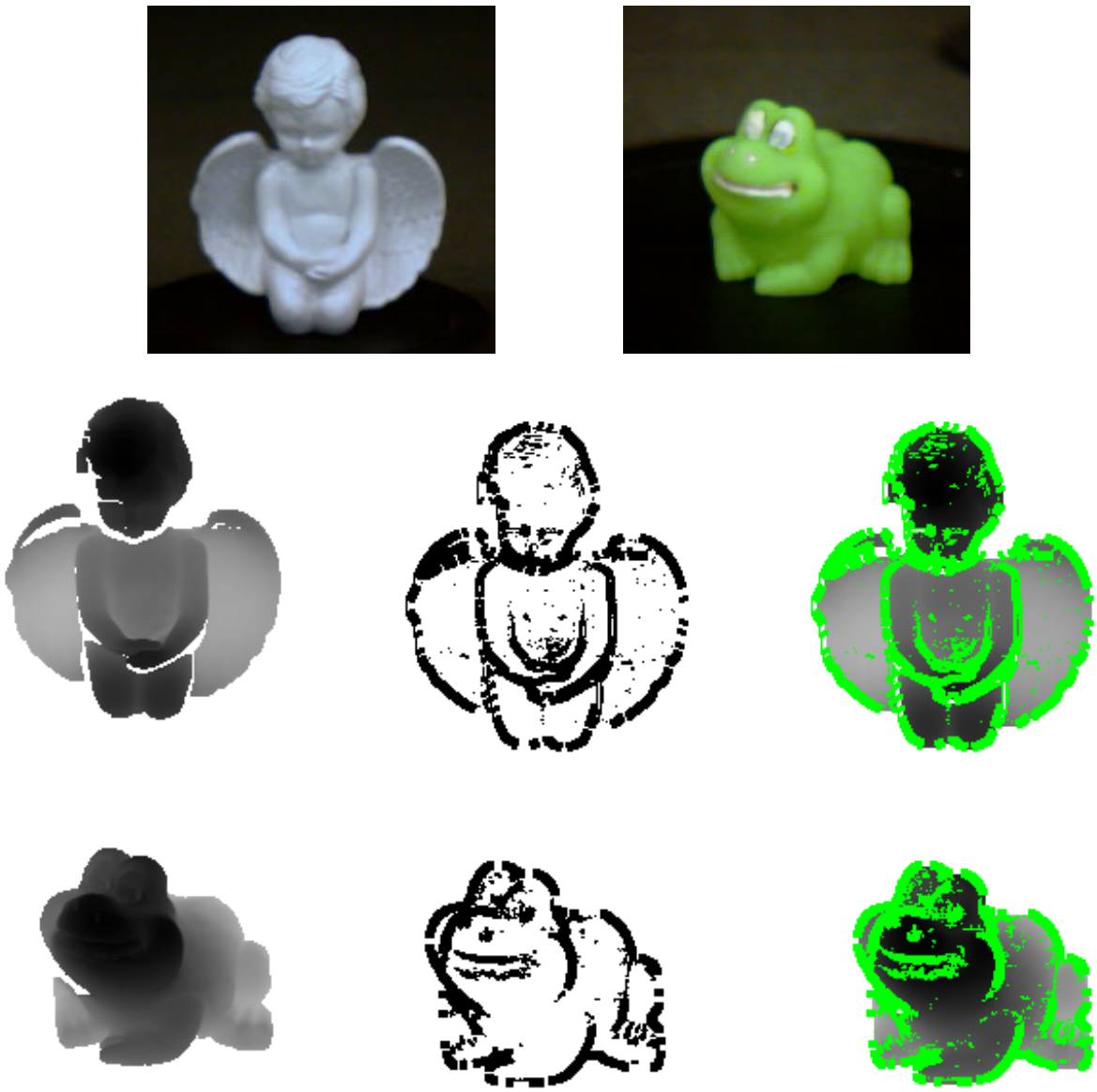


Figure 4.15: Segmentation results of free form objects, angle status and frog range images collected using Minolta scanner (OSU) database.

4.4 Edge Linking Algorithm

In general, edge detectors based on partial derivative do not guarantee producing edge maps with continuous edge contours because of its high sensitivity to noise [William 1989]. The major problem with thresholding process in edge detection is that we consider only the range differences, not any relationships between the pixels. For this reason, there is no guarantee that the pixels identified by the thresholding process are contiguous. So there is probability to miss pixels, which represent an edge, and these effects get worse as the noise get worse. Because many object recognition algorithms prefer closed contours of objects for comparison and matching, a separate edge-linking step is some times needed to generate a more complete edge map.

In this section, we propose an edge-linking algorithm to fill the gaps between the edge segments in order to form a more continues edge. The main step in the edge-linking algorithm is how to find the break points. The break point is defined as a point at which an edge line is terminated. In our algorithm, recognizing the break points in a given edge map done by considering the structures of edge line in a window of 3x3, as shown in figure 4.16. We benefit from natural structure of the break point, so if a given line counts 2 pixels within the 3x3 widow then the center point of that edge line is a break point, see figure 4.16. Whether the center pixel is a break point can be tested using the following criteria:

$$Break(x,y) = \begin{cases} 1 \;\; if \;\; \sum_{i=0}^{i=8} I_i \;\; = \;\; 2 \\ 0 \quad otherwise \end{cases}$$

Ones the break points in an edge image are identified. For each break point the algorithm searches to find the closest break point using distance threshold as follow:

$$\left| P(E1(x,y)) - P(E2(x,y)) \right| \le T$$

$P(E1(x,y))$: the position of the break point 1.

$P(E2(x,y))$: the position of the break point 2.

Then a straight line will be constructed between those break points to close the gap. Generally the choice of the distance threshold value depends on the requirements for edge linking accuracy and the strong of the original edge map. Figure 4.17 shows an example how the edge-linking algorithm works to close the gaps in the edge image of a column range image.

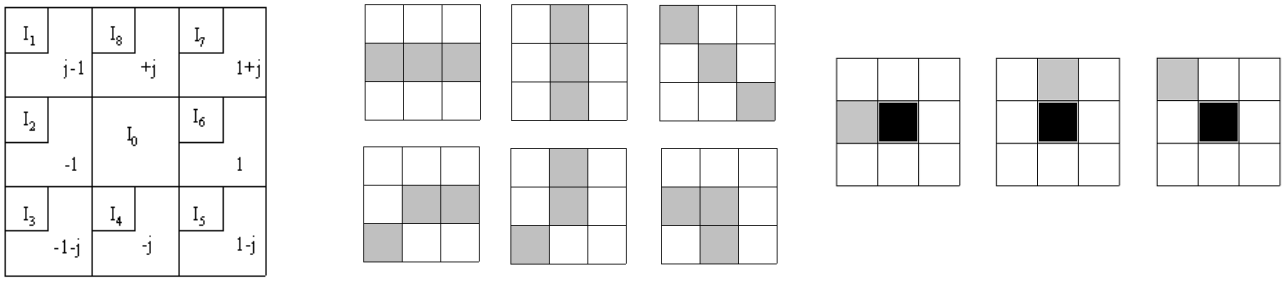


Figure 4.16: To the left, window size (3x3) used to convolve the edge image. In the mid, examples of continues edges. The right figure shows examples of how the break points are defined (black).

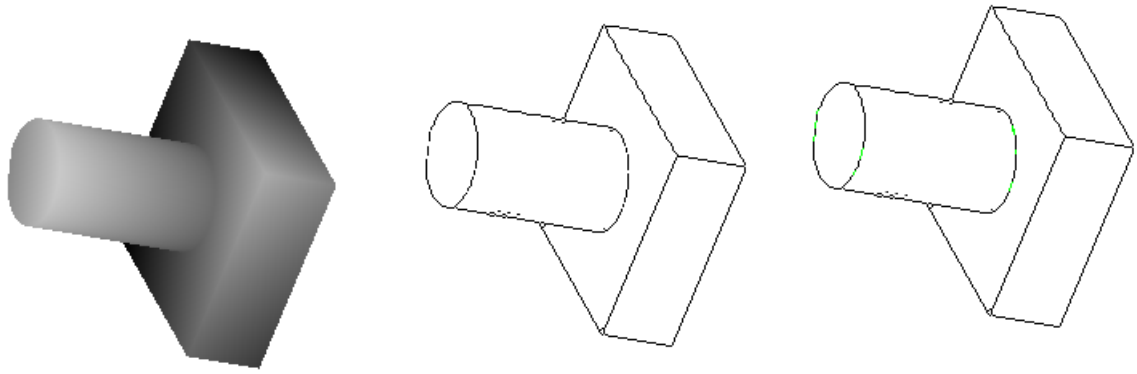
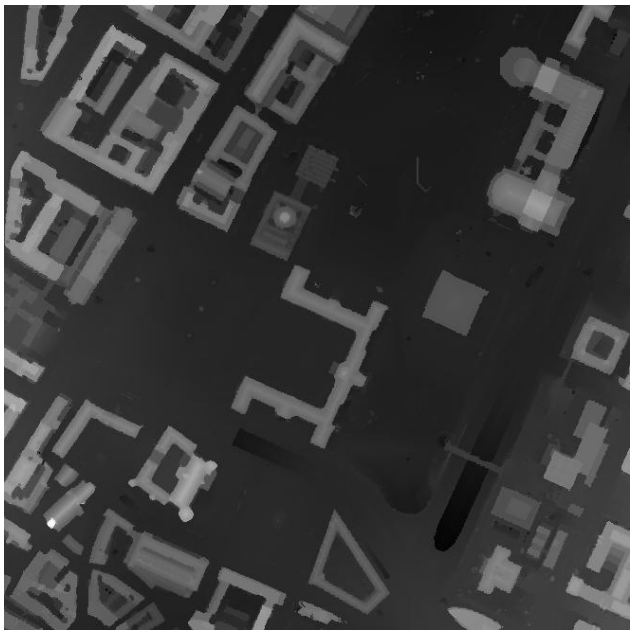


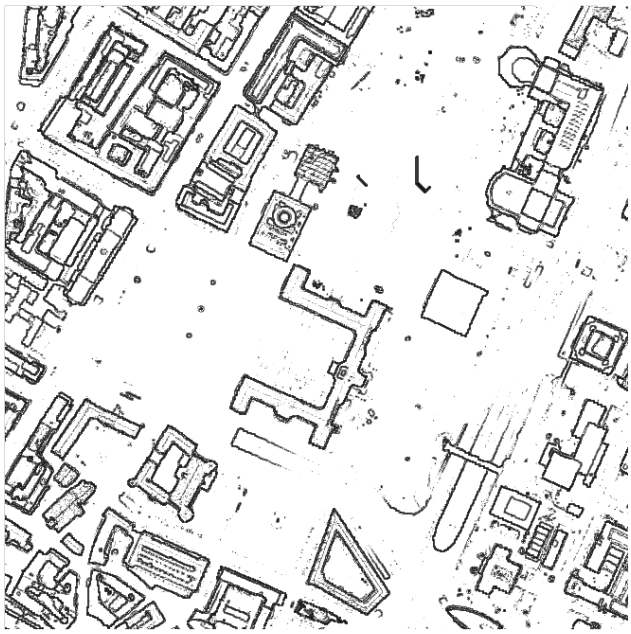
Figure 4.17: Column range image. In the mid; edge map produced using the proposed segmentation algorithm. The right figure shows closing the available gaps using edge-linking algorithm (green).

4.5 Experimental Results

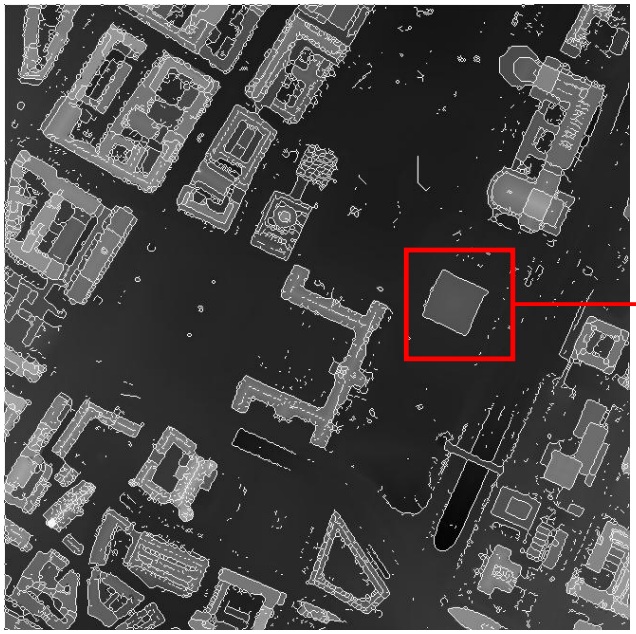
The performance analysis of has been conducted on a standard PC with an Intel 4, 3GHz Processor, and 1GB RAM. The proposed edge detection algorithm and edge-linking algorithm have been implemented in C++. In order to demonstrate the efficiency and robustness of the method. The proposed algorithm has been tested on a large number of real range images acquired by different range sensors; terrestrial and airborne laser scanners. Figure 4.18- 4.20 shows range images acquired by toposys laser scanner system for some parts of Stuttgart city and Alps Mountains. The system provide terrain points measured at approximately one point each 1x1 m^2 with an accuracy of 0.3 m in planimetry, and 0.1 m in height. The measurements are provided along strips, which are usually processed and resampled to obtain a regular raster. Figure 4.18a shows range image for the Schloss Platz area of Stuttgart city. The produced edge map is shown in figure 4.18b. It can be noticed that the edges are thick since we have used mask size 11, which is necessary for robust estimation of the mean curvature values in such high noise images. Figure 4.18c depicts the binary edge map projected on the range image after skeletonized process. Figure 4.19 and 4.20 show other results of the edge detection algorithm for airborne range images of a highway intersection area in Stuttgart city, and Alps Mountains chain range images. For terrestrial application, the algorithm is tested on range images collected from Petra treasury (Al-kahsneh) as can be depicted in figure 4.21. The terristrail data has been collected using Mensi GS100 laser scanner.



(a)



(b)



(c)

(d)

Figure 4.18: a) range image for Schloss platz area, Stuttgart city b) segmentation results using mask size 11 c) edge map projected in the range image after skeletonized d) extracting more features using different threshold values

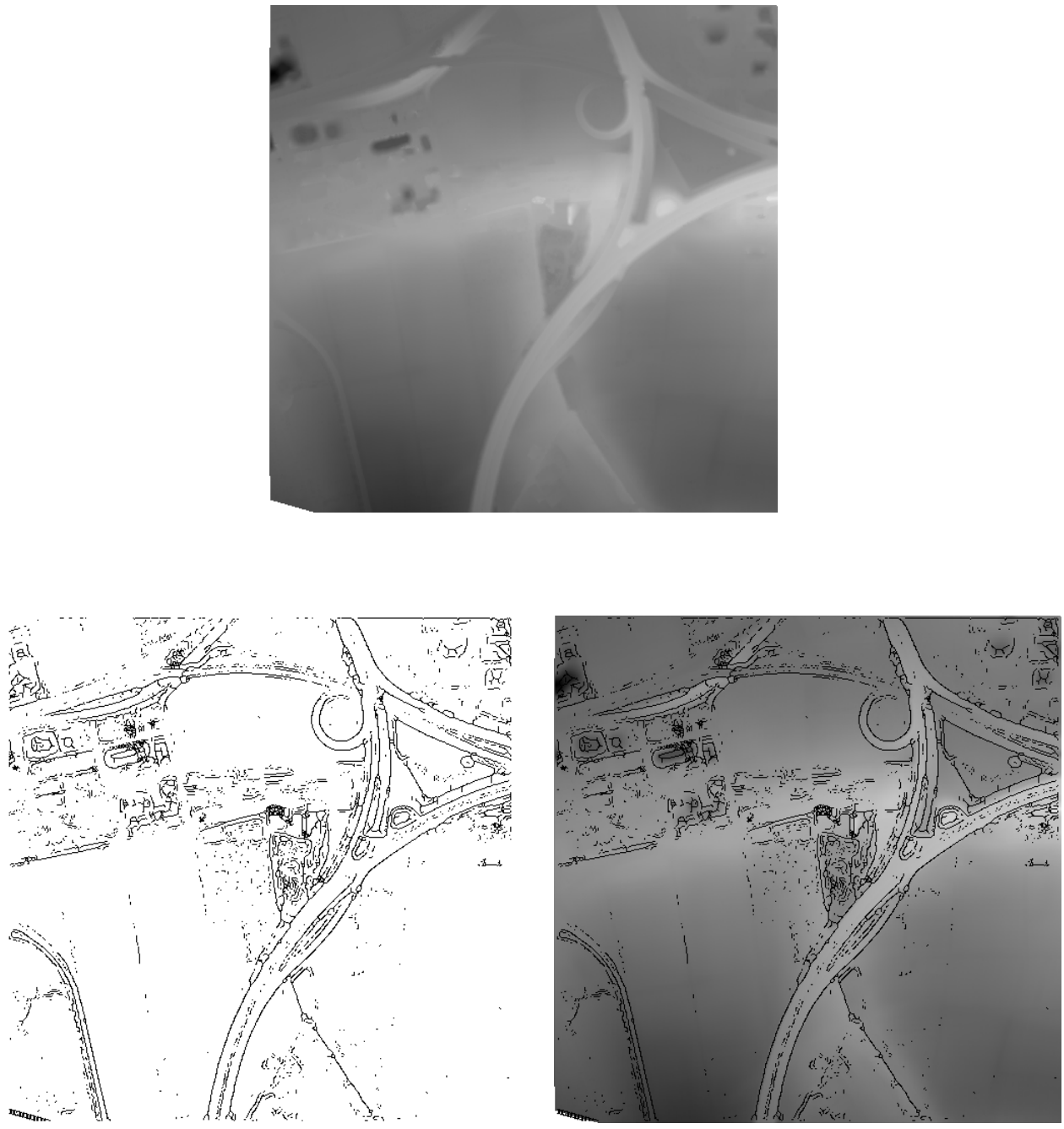


Figure 4.19: Segmentation results for a range image of a highway intersection in Stuttgart city

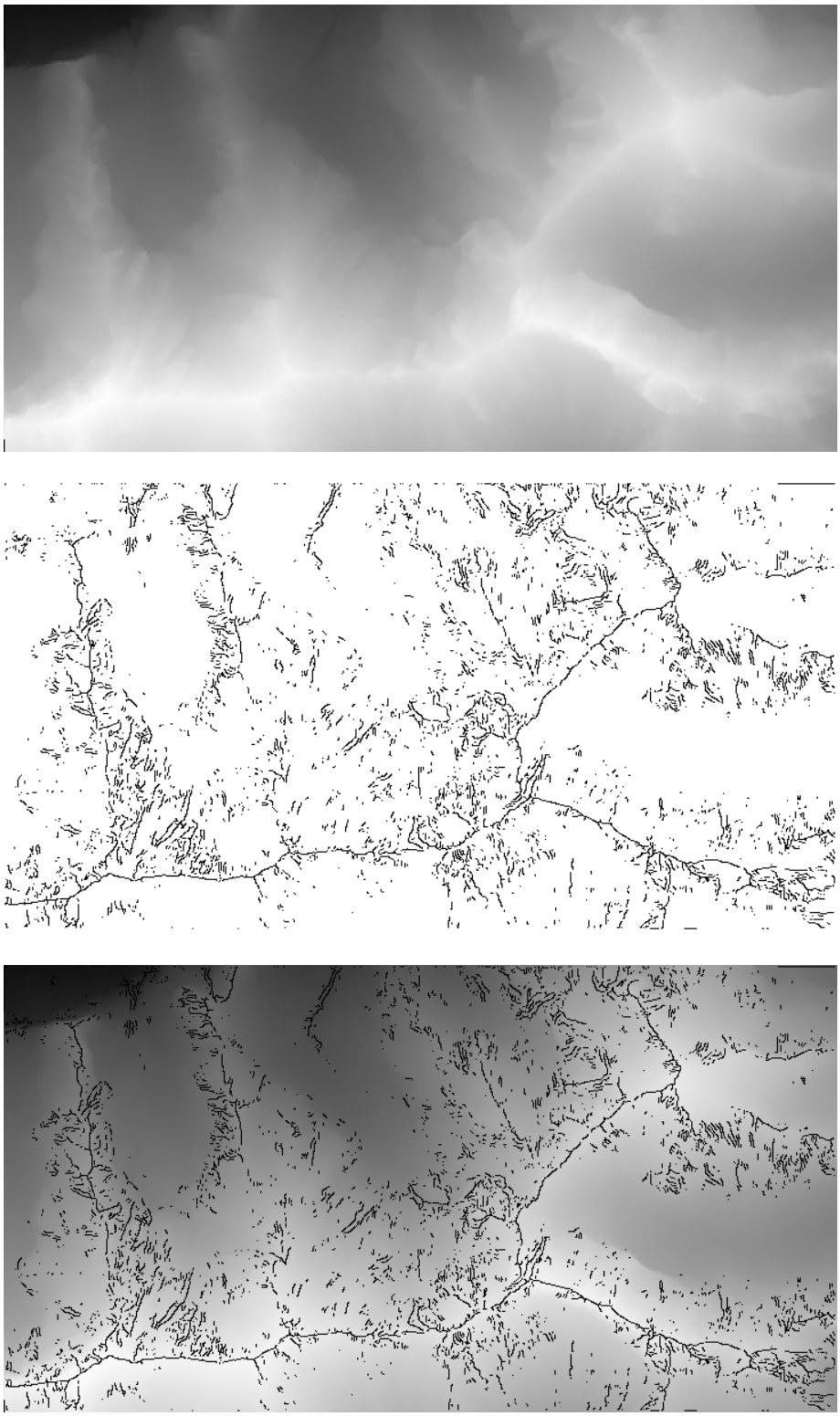


Figure 4.20: Segmentation results for a range image of Alps Mountains

Figure 4.21: Segmentation results for range images collected from Petra treasury (Al-khasneh)

4.6 Intensity Image Segmentation

The technique described above for range image segmentation can be used to segment the intensity images. In order to implement the algorithm, a series steps must be followed. The first step is to filter out any noise in the original image before trying to locate and detect any edges. Because the Gaussian filter can be computed using a simple mask, it is used exclusively in our algorithm. Once a suitable mask has been used, the Gaussian smoothing can be performed using standard convolution method. The larger the width of the Gaussian mask, the lower is the detectors sensitivity to noise. After smoothing the images, the next step is to find the edge strength by estimating the partial derivative images via the following 2-D image convolutions.

$$\tilde{g}_u(i,j) = D_u \otimes S \otimes \tilde{g}(i,j) \quad , \quad \tilde{g}_v(i,j) = D_v \otimes S \otimes \tilde{g}(i,j) \quad , \quad \tilde{g}_{uu}(i,j) = D_{uu} \otimes S \otimes \tilde{g}(i,j) \quad ,$$

$$\tilde{g}_{uv}(i,j) = D_{vv} \otimes S \otimes \tilde{g}(i,j) \quad , \quad \tilde{g}_{uv}(i,j) = D_{uv} \otimes S \otimes \tilde{g}(i,j)$$

The spatial analysis of the mean curvature values along a cross section in the surface of Al-kasneh intensity image has been plotted in figure 4.22. As mentioned before, edge will have higher pixel intensity values than those surrounding it. Smooth surfaces have a close intensity values creating small values of the derivatives and as a result zero mean curvature. Whereas the irregular surfaces creating large values of derivatives with mean curvature value larger than zero. The value of mean curvature depends on the strength of intensity change. Since we deal with byte image the computed partial derivatives and the mean curvature images will have integer values. Examples of such values shown in figures 4.23 and 4.24 for both smooth surface and for area contains an edge. It can be noticed that the mean curvature value estimation depends mainly in the second partial derivative in the direction of the convolution mask $\tilde{g}_{uu}(i,j)$. This value is larger than zero in the presence of edges and equal zero in smooth areas:

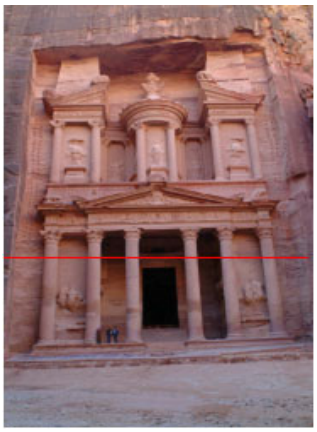
$$H(i,j) = \frac{(1+\tilde{g}_v 2\,(i,j))\;\tilde{g}_{uu}(i,j) + (1+\tilde{g}_u 2\,(i,j))\tilde{g}_{vv}(i,j) - 2\tilde{g}_u(i,j)\;\tilde{g}_v(i,j)\tilde{g}_{uv}(i,j)}{2(\sqrt{1+\tilde{g}_u{}^2(i,j)+\tilde{g}_v{}^2(i,j)})^3}$$

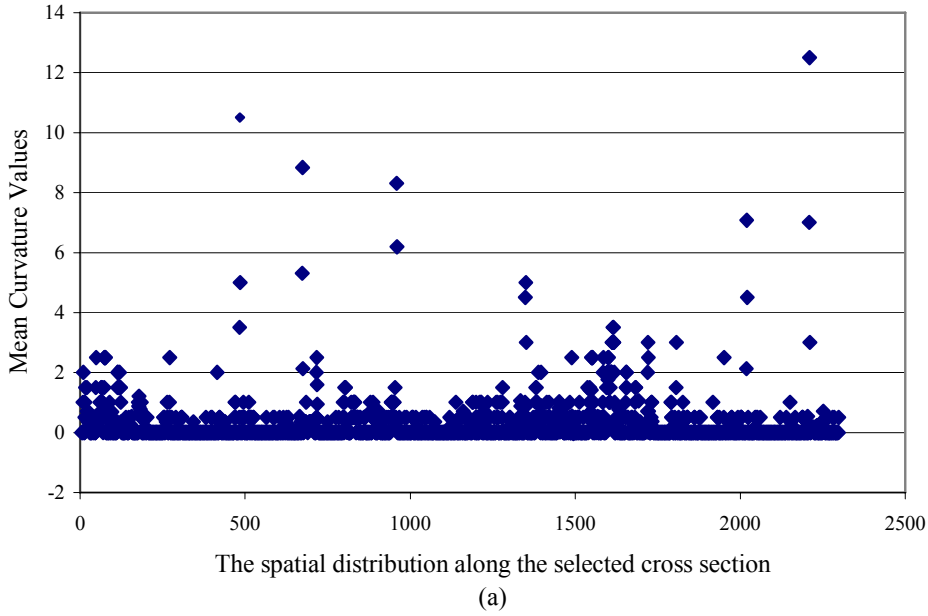
After computing the mean curvature values for all the pixels in the image. The pixel is declared an edge location if the value of the mean curvature is greater than zero. So the technique is based upon a simple threshold concept:

$$E(x,y) = \begin{cases} 1 \; if \; \| H \| > zero \\ 0 \quad otherwise \end{cases}$$

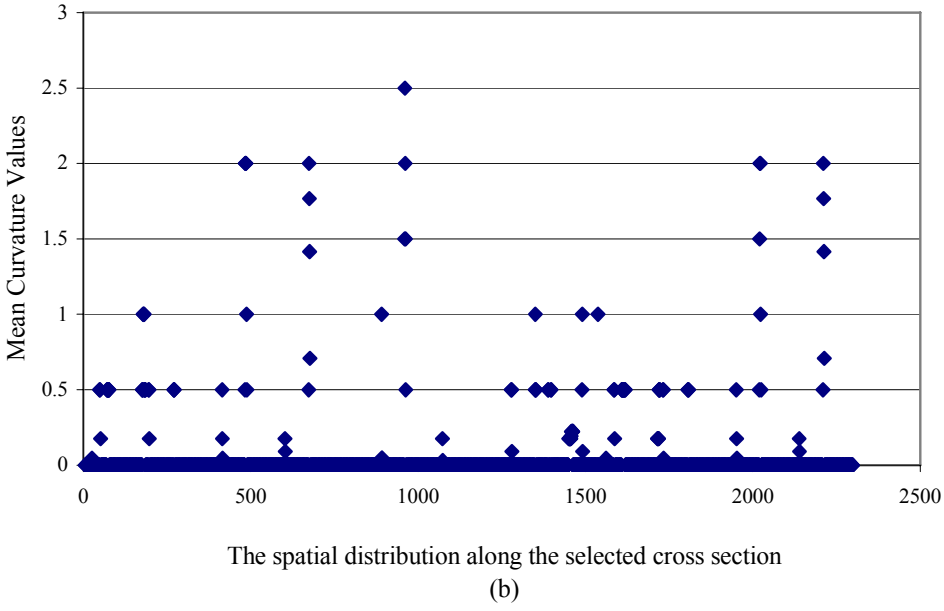
$\{(x,y) : E(x,y) = 1\}$ The set of edge pixels.

We have tested the proposed algorithm in large number of intensity images, Figures 4.25 and 4.26 show different examples for segmenting Petra treasury, Mandrill face, Ziess plate images (Halcon images collection) and Avenches image.



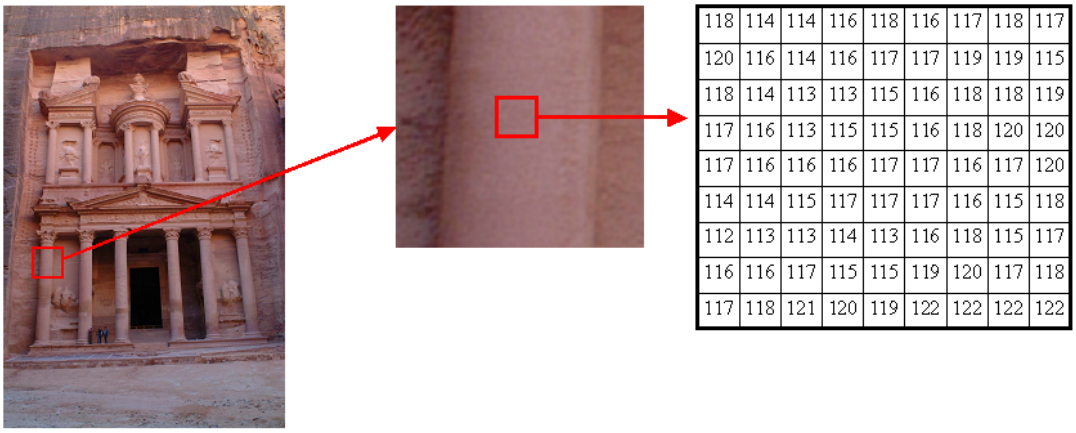


(a)



(b)

Figure 4.22: Spatial distribution of Mean curvature values using mask size 7 along the selected cross section of Al-khasneh/ Petra (a) before image smoothing (b) after image smoothing using Gaussian operator.



118	114	114	116	118	116	117	118	117
120	116	114	116	117	117	119	119	115
118	114	113	113	115	116	118	118	119
117	116	113	115	115	116	118	120	120
117	116	116	116	117	117	116	117	120
114	114	115	117	117	117	116	115	118
112	113	113	114	113	116	118	115	117
116	116	117	115	115	119	120	117	118
117	118	121	120	119	122	122	122	122

P.D.E.I (Mask Size 9) Real Values

$\tilde{g}_u(i,j) = 0.3796$
$\tilde{g}_v(i,j) = 0.2018$
$\tilde{g}_{uu}(i,j) = 0.1130$
$\tilde{g}_{vv}(i,j) = 0.2876$
$\tilde{g}_{uv}(i,j) = 0.0458$

P.D.E.I Integer Values

$\tilde{g}_u(i,j) = 0$
$\tilde{g}_v(i,j) = 0$
$\tilde{g}_{uu}(i,j) = 0$
$\tilde{g}_{vv}(i,j) = 0$
$\tilde{g}_{uv}(i,j) = 0$

⇨ H= 0

P.D.E.I (Mask Size 7) Real Values

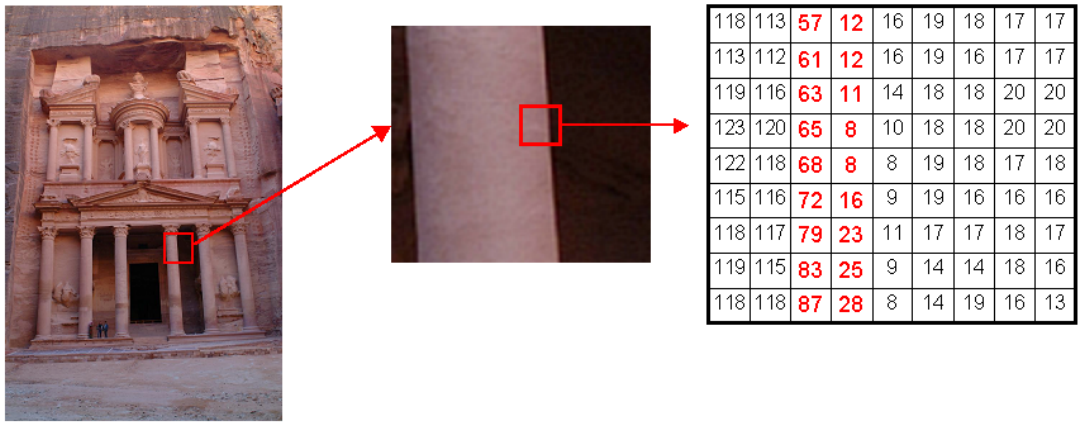
$\tilde{g}_u(i,j) = 0.56632$
$\tilde{g}_v(i,j) = -0.05612$
$\tilde{g}_{uu}(i,j) = 0.091836$
$\tilde{g}_{vv}(i,j) = 0.180272$
$\tilde{g}_{uv}(i,j) = -0.59$

P.D.E.I Integer Values

$\tilde{g}_u(i,j) = 0$
$\tilde{g}_v(i,j) = 0$
$\tilde{g}_{uu}(i,j) = 0$
$\tilde{g}_{vv}(i,j) = 0$
$\tilde{g}_{uv}(i,j) = 0$

⇨ H= 0

Figure 4.23: Partial derivative estimation images (P.D.E.I) computed for a smooth area (no edge available)



118	113	57	12	16	19	18	17	17
113	112	61	12	16	19	16	17	17
119	116	63	11	14	18	18	20	20
123	120	65	8	10	18	18	20	20
122	118	68	8	8	19	18	17	18
115	116	72	16	9	19	16	16	16
118	117	79	23	11	17	17	18	17
119	115	83	25	9	14	14	18	16
118	118	87	28	8	14	19	16	13

P.D.E.I (Mask Size 9) Real Values

$\tilde{g}_u(i,j) = -13.4259$
$\tilde{g}_v(i,j) = 0.47222$
$\tilde{g}_{uu}(i,j) = 7.0052$
$\tilde{g}_{vv}(i,j) = 0.0477$
$\tilde{g}_{uv}(i,j) = -0.2469$

P.D.E.I Integer Values

$\tilde{g}_u(i,j) = 0$
$\tilde{g}_v(i,j) = 0$
$\tilde{g}_{uu}(i,j) = 7$
$\tilde{g}_{vv}(i,j) = 0$
$\tilde{g}_{uv}(i,j) = 0$

⇨ H= 3.5

P.D.E.I (Mask Size 7) Real Values

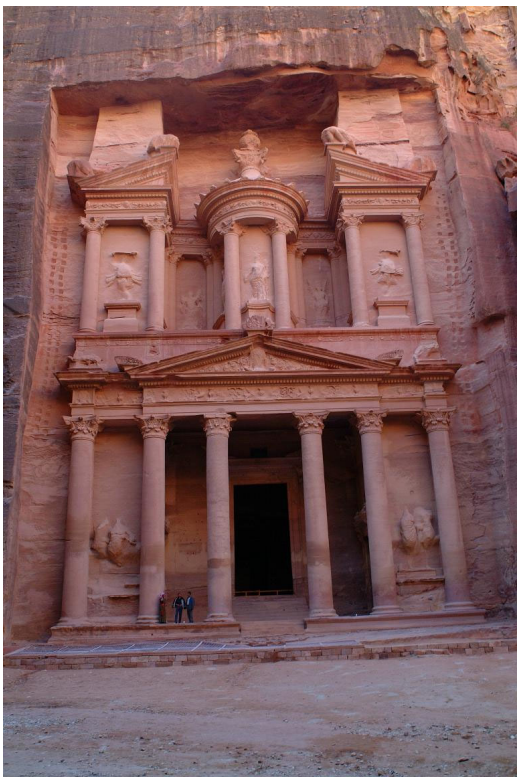
$\tilde{g}_u(i,j) = -14.2398$
$\tilde{g}_v(i,j) = 0.6326$
$\tilde{g}_{uu}(i,j) = 12.622$
$\tilde{g}_{vv}(i,j) = 0.21088$
$\tilde{g}_{uv}(i,j) = -3.530$

P.D.E.I Integer Values

$\tilde{g}_u(i,j) = 0$
$\tilde{g}_v(i,j) = 0$
$\tilde{g}_{uu}(i,j) = 12$
$\tilde{g}_{vv}(i,j) = 0$
$\tilde{g}_{uv}(i,j) = 0$

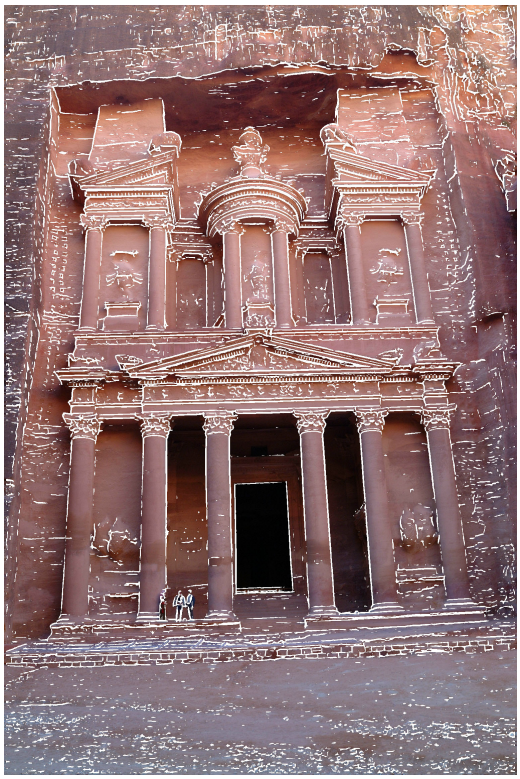
⇨ H= 6

Figure 4.24: Partial derivatives estimation images (P.D.E.I) computed for irregular area



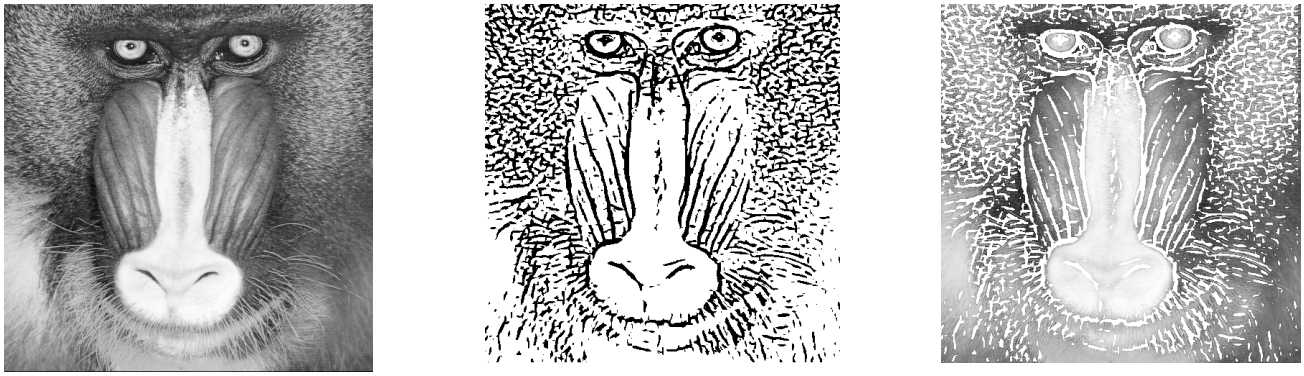
(a)

(b)

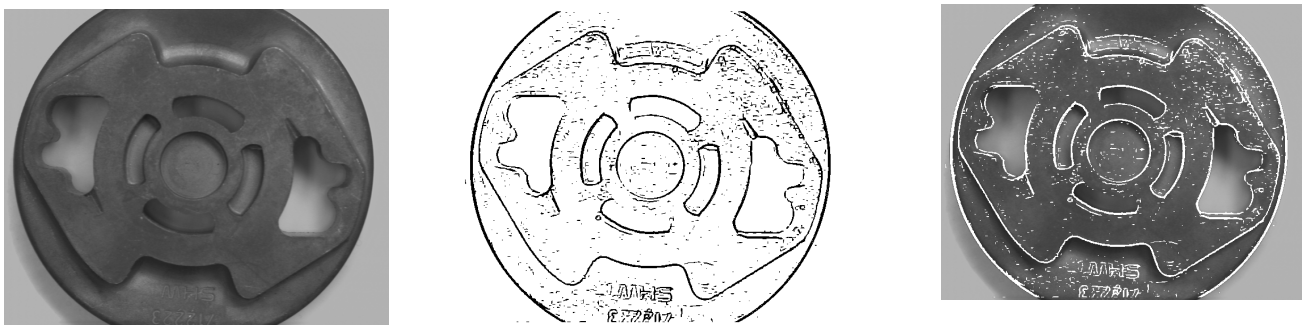


(c)

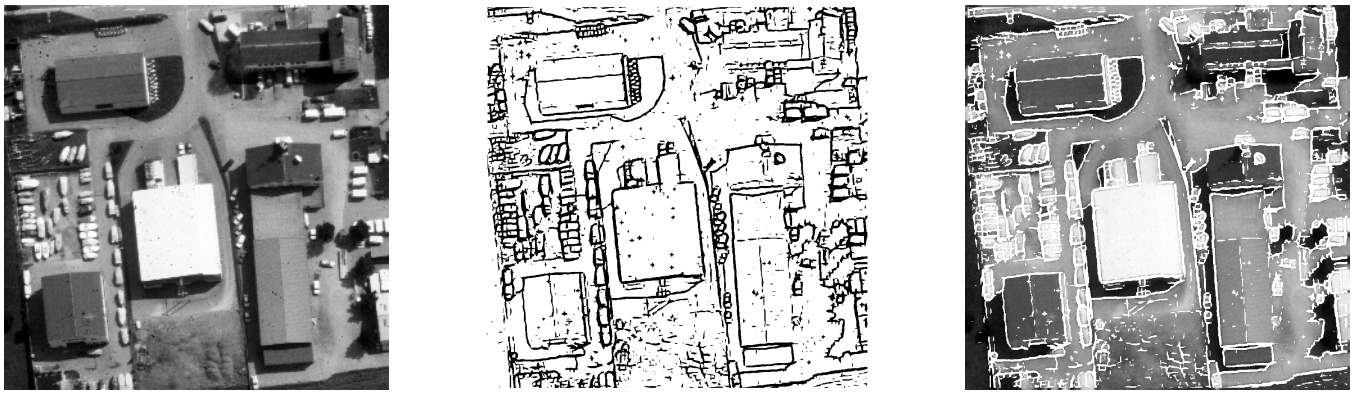
Figure 4.25: a) The segmentation results of Al-khasneh using mask size 5 and Gaussian smoothing b) the segmentation results using the proposed algorithm c) segmentation results projected on the colored image.



(a)



(b)



(c)

Figure 4.26: Examples of segmentation results using the proposed algorithm for intensity images
(a) Mandrill image (b) Ziess image (c) Avenches image

5 Automatic Multi-Image Photo-Texturing

5.1 Introduction

The issue of acquiring and constructing high-quality texture maps has received less attention than the issue of capturing high-quality geometry [Bernardini et al, 2001]. Traditional texture mapping techniques, which are simply draping static imagery over geometry, are not sufficient to represent reality. For high quality photorealism texture mapping using multiple images different issues have to be taken into consideration, these issues involve occlusion problem, accurate data registration, image radiometric and geometric distortion problems. Those issues have been already discussed in the introduction chapter and will be discussed in details in the next section of this chapter.

Alignment of the local coordinate systems of the range and colour data sets is the first step in the fusion problem. Many solutions have been studied in the pose estimation framework from 2D-3D correspondences. Most of the methods coming from the photogrammetry community use point matches. But in the application of range images, the perception of objects structure from point clouds is limited and not very appropriate for registration [Liu & Stamos, 2005], an example is shown in figure 5.1. Moreover, those method correspond to laser footprints rather than a distinct points that could be identified in the imagery. With the recent trend towards automatic extraction and recognition of features from digital imagery, it is becoming advantageous to utilize linear features in photogrammetric applications. From a practical point of view, line detection is more accurate and more reliable than points, and line matching is more robust than point matching with respect to partial occlusions [Christy & Horaud, 1999]. Additionally, those features can be used to increase the redundancy and improve the geometric strength of photogrammetric adjustments. Beside the fact that images of man made environment are rich with linear features, it is easier to automatically extract features than distinct points from imagery, these edges can be often approximated by straight lines [Horaud et al, 1997, Lee & Habib, 2002].

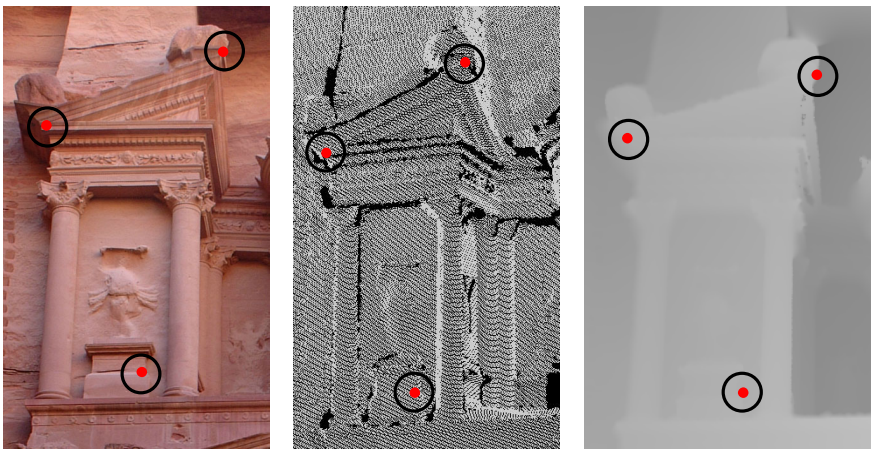


Figure 5.1: Limited perception in finding the corresponding points in point clouds and range image.

Beside the previous issues, if multiple images are used for texture mapping, calibration of the camera and image radiometric correction is required to guarantee the quality of the mapping. Radiometric distortions occur from different sensed brightness between different camera positions. If several images are mapped on a single triangular model, the discontinuities and artefacts are present along the edges of adjacent triangles textured with different images, as depicted in figure 5.2. These differences create a visually very disturbing effect. To create seamless texture maps blending methods have been developed [Baumberg, 2002]. Those methods use the weighted average of pixel values of triangles in the overlapping sections between two adjacent textures. In another hand, geometric distortion is resulting from incorrect camera calibration and orientation. Geometric distortion, like line discontinuity will be visible at common edges of adjacent triangles mapped from different images, can be minimized by lens distortion correction and accurate image registration [El-Hakim et al, 2003].

In this chapter we will discuss an effective automatic approach to the problem of high-resolution photo-realistic texture mapping onto 3D complex model with minimum effort. Our approach allows taking the images at different time

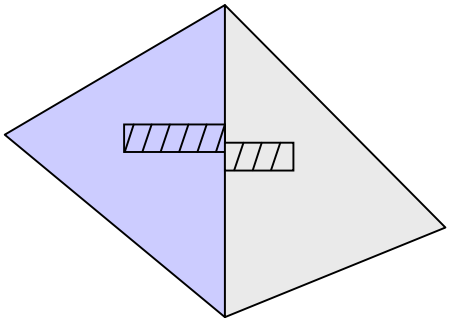


Figure 5.2: Distortion of texture at the boundaries of triangles, which are textured from different images

from laser scanning and at whatever locations that will be the best for texturing. The reconstruction of texture maps from multiple views needs the information concerning the positions of a pixel in texture space in all images. This information is present in the transformation from texture to model and further on to image space. In order to guarantee accurate fusion linear feature 2D-3D matching are used to recover the positions of the 2D cameras with respect to the geometric model. In this framework, the edge detection algorithm which is discuused in chapter 4 will be used to detect the line features in both range and intensity images .The purposes are to simplify the dense datasets and provide stable features of interest, which are used to recover the positions of the 2D cameras with respect to the geometric model. Before performing projective texture mapping, we need to compute visibility information to map the image only onto the portions of the scene that are visible from its camera viewpoint. For this reasons an effective automatic visible surface algorithm was developed to address the occlusion and images fusion. For high quality texture mapping, pre-processing steps for generation geometric and radiometric distortion free images have been done. The presented approaches are demonstrated in the framework of a project aiming at the generation of a 3D virtual model of the Al-Khasneh/ Petra, and the Romanian Theatre in ancient Jerash city. In this chapter, section 2 discusses the basics of texture mapping. Section 3 provides overview of the main issues associated with the modeling of high textured real environments and the techniques currently used to address these challenges. Imagery segmentation and feature matching are provided in section 4. An efficient proposed visible surface algorithm is presented in section 5. In section 6 the over all procedure of the proposed texture mapping approach is discussed with the pr-processing process.

5.2 Basics of Texture Mapping

The purpose of texture processing is to integrate the 3D measurements from the laser scanner with 2D information taken with an external or internal camera. The external parameters are the translation and rotation of the camera relative to the global reference frame; the internal parameters vary according to the mathematical model that is used to describe the camera. The simplest model uses 5 parameters (focal length, centre of project, pixel aspect ratio and 1st degree radial distortion [Tsai, 87]. The source image (*texture*) is mapped onto a surface in 3-D *object space*, which is then mapped to the destination image (*screen*) by the viewing projection. Texture space is labeled (u, v), object space is labeled (x_o, y_o, z_o), and screen space is labeled (x, y), as can be depicted in figure 5.3.

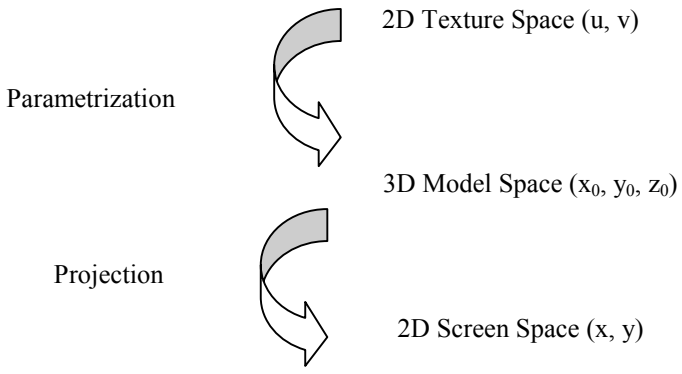


Figure 5.3: Mapping from texture space to screen space using parameterization and viewing projection

The relation between object space and texture space is simply a coordinate system transformation. Texture coordinates are transformed into object coordinates by inverse texture mapping function, which assigned texture points *(u, v)* to 3D object points (x_o, y_o, z_o). For example VRML standard uses this kind of transformation for indexed face sets. In another hand the object-screen space transformation is usually either affine projection or perspective projection (central). Affine mapping imaging with parallel rays and exhibits uniform scale. For texture mapping the general 2Da affine mapping may be written algebraically as:

$$(x \;\; y \;\; 1) = (u \;\; v \;\; 1) \begin{bmatrix} a & d & 0 \\ b & e & 0 \\ c & f & 1 \end{bmatrix}$$

In the other hand the perspective transformation involves imaging with rays that converge to a focal point and exhibits scale that varies from foreground to background. The general form of a projective mapping is a rational linear mapping:

$$x = \frac{au + bv + c}{gu + hv + i}, \quad y = \frac{du + ev + f}{gu + hv + i}$$

The homogenous representation for points provides a consistent notation for projective mapping. Thus, mnipulation of projective mappings is much easier in the homogeneous matrix1 notation:

$$\begin{pmatrix} x' & y' & w \end{pmatrix} = \begin{pmatrix} u' & v' & q \end{pmatrix} \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$$

where $(x, y) = (x'/w, y'/w)$ for $w \neq 0$, and $(u, v) = (u'/q, v/q)$ for $q \neq 0$.

Although there are 9 coefficients in the matrix above, these mappings are homogeneous, so any nonzero scalar multiple of these matrices gives an equivalent mapping. Hence there are only 8 degrees of freedom in a 2-D projective mapping. The projective mapping is affine when g=h=0. Heckbert presents the most basic 2-D mappings such as affine mappings and projective mappings and discusses them in detail.

Typically most of complex scenes converted to polygons for 3D modeling purposes, thus, the projected vertices determine the boundary edge of the projected polygon. In our proposed approach the perspective projective is selected to ensure the proper perspective foreshortening to have more realistic projection of the triangles. This is done by associating the scaling texture coordinates *(u, v)* for each triangle vertex as depicted in figure 5.4.The texture is then cast onto a geometric model using the camera position as the center of projection.

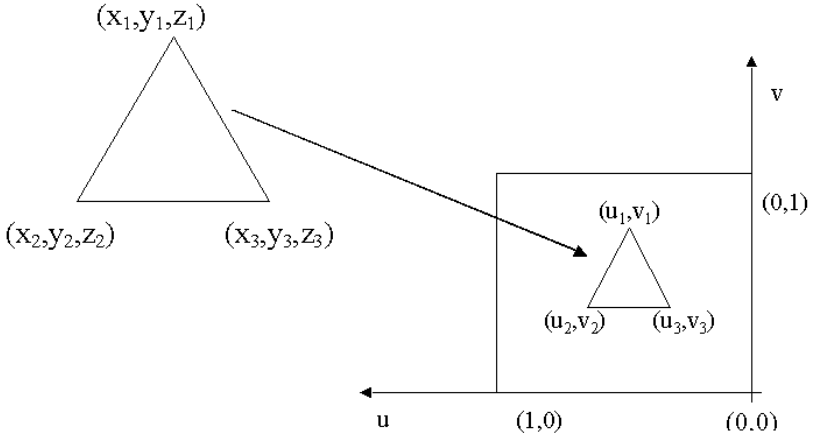


Figure 5.4: Applying a texture to a 3D triangle (right), the scaling texture coordinates (u,v) of the vertices will be copied to the 3D triangle.

5.3 Issues in Texture Mapping From Multiple Images

For high quality texture mapping using multiple images different issues have to be taken into consideration during the projective process, these issues involve (a) geometric distortions; camera calibration (b) 2D-3D image registration. (c) Radiometric distortions; color blending and correction. (d) Occlusion problem. In the following sections, we will discuss those factors which affecting photorealism and possible solutions to reduce their effect.

5.3.1 Camera Calibration

The central perspective projection is an abstract entity. When it is used in photogrammetry as the mathematical model for image formation in a camera, significant discrepancies leading to errors will occur. The most obvious discrepancy between the abstract and the real is the lens. For this reason, camera calibration used to identify how much the geometry of image formation in a real camera differs from the geometry of central perspective projection. Through calibration, the camera intrinsic parameters are obtained. The intrinsic parameters comprise of:

a) Effective focal length of the pinhole camera.

b) Two radial and two tangential distortion coefficients.

c) Coordinates of the center of distortion in the image space.

d) Scale factor that accounts for the uncertainty in the frame grabber's re-sampling of the horizontal scan line.

There are different ways to obtain the required parameters; the manufacturer usually specifies the internal parameters on the data sheet. However, for a non-metric camera, these values are usually not precise enough accurate texture mapping. The internal parameters can also be computed for a specific camera using standard computer vision algorithms, which process images taken of a calibrated targets either in the laboratory (e.g. white dots on a black ground) or in the filed during the data collection; in this framework, the test field needs numerous control points that have to be precisely surveyed prior to the calibration process. Establishing and maintaining such a test field is an expensive procedure that has to be conducted by professionals. For some heritage sites and monuments it is not allowable for pasting these targets on the surface of the objects. Moreover, these artificial objects may disturb the information needed from the intensity images for purposes such as texture mapping [Lensch et al, 2000]. Another alternative which does no need any calibrated targets is the self-calibration method where the internal (and external if necessary) parameters for each image can be calibrated using the 3D point cloud or the corresponding model. The calibration algorithm needs a set of corresponding 2D coordinates from the image and 3D coordinates from the point cloud to compute the required parameters. This procedure mainly depends on the accurate extraction of the corresponding points.

Thus, in general, when cameras with standard lenses are used, lens distortion corrections must be applied else geometric distortions, like line discontinuity, will be visible at common edges of adjacent triangles mapped from different images. It is therefore important that full camera calibration is applied and that it remains valid for all images by keeping camera settings constant or applying self-calibration.

5.3.2 2D-3D Data Co registration

After creating zero distortion images using the calibration parameters, the camera position and rotation relative to the 3D model must be determined for each view. Only if geometry and texture are acquired at the same time with the same sensor, the images are already aligned to the model and no further 3D to 2D registration is needed. This could be found in some commercial 3D systems**.** The system already provides model-registered color texture by capturing the RGB values of each LIDAR point using a camera already integrated in the system. However, these images frequently are not sufficient for high quality texturing, which is desired for documentation, since the ideal conditions for taking the images may not coincide with those for laser scanning. So it is therefore more useful to

acquire geometry and texture by two independent processes and allow for an image collection at optimal position and time for texturing. This is especially true for the high requirements of realistic documentation of heritage sites. [El-Hakim et al, 2002; Alshawabkeh & Haala, 2005]. To calculate the pose for arbitrary rotation and translation, we need to know the correspondences between features in the image and the geometry. Different methods exist each of them has advantages and disadvantages according to the given scene and the required accuracy. This section provides a brief overview of the different methods with the respective pros and cons:

- Registration using targets: Targets (planar or spherical) are the standard way to register two scans. Each pair of scans needs to contain at least three targets, which can be recognised by the processing software.but these marks destroy the texture and have to be removed afterwards. In addition for this method more planning must go into scene capture. To get targets' center fine position, high-resolution scan is needed, which is time-consuming.
- Feature-based registration: With the recent trend towards automatic extraction and recognition of features from digital imagery (e.g. lines, corners). It is becoming advantageous to utilize linear features in photogram metric applications. This is due several facts; linear features are easier to extract than distinct points, especially in a digital environment [Habib et al, 2004]. Those features can be used to increase the redundancy and improve the geometric strength of photogrammetric adjustments. the fact that images of man made environment are rich with linear features, the edges of human made objects often can be approximated by straight lines. The accuracy of the transformation depends on the accuracy with which the features have been extracted from the scans. For providing stable features of interest efficient segmentation algorithms have to be used to extract the features automatically for the co registration of data sets.

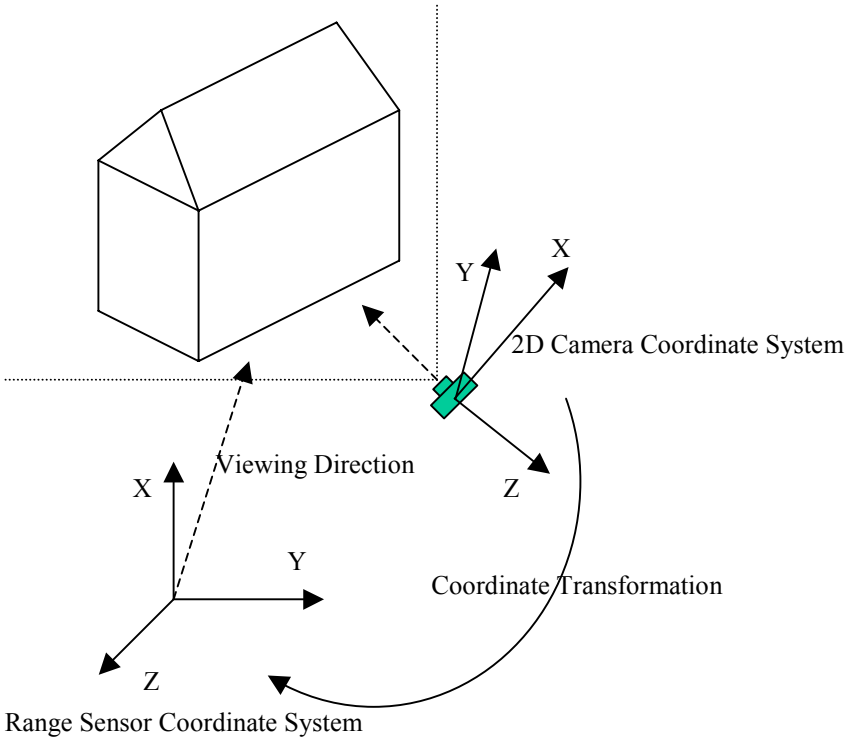


Figure 5.5: The pose estimation problem

5.3.3 Color Correction

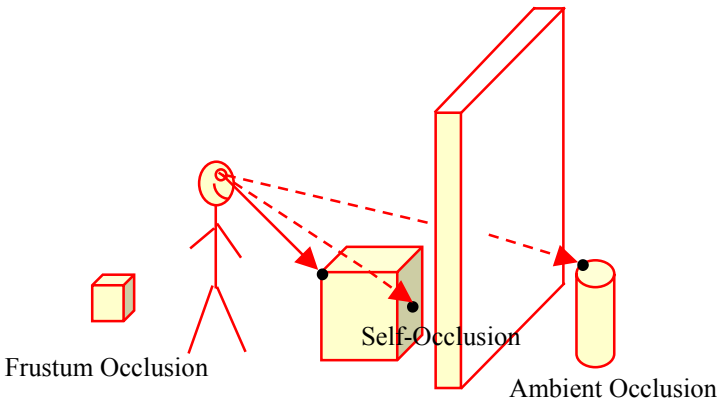
Image radiometric distortion comes from the use of different images acquired in different positions, with different cameras or in different daily moments (e.g. varying light conditions). If several images are mapped on a single triangular model, the discontinuities and artefacts are present along the edges of adjacent triangles textured with

different images. These differences create a visually very disturbing effect. To reduce these effects, the images need to be balanced on a global per image basis and finally blended locally along the transition between images. Blending methods to create seamless texture maps have been developed [Pulli et al, 1998]. Those techniques use the weighted average of pixel values of triangles in the overlapping sections between two adjacent textures.

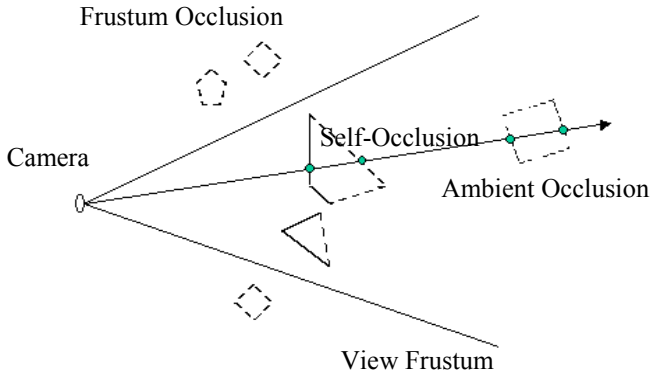
5.3.4 Occlusion Problems

Once the pose parameters have been found. The image is wrapped over the geometry of a triangle mesh constructed from the range data. In order to texture map, we re-project the color image to the range scanner's frame of reference using the registration results. Ideally, warping the image over the geometry associated each 3D point to a pixel in the color image, the texture pierce through the geometry and gets mapped onto all occluded polygons on the path of the projected ray. So the geometry that is occluded in the image will receive incorrect texture coordinates instead of remaining in shadow (double projection problem). Figure 5.6 and 5.7 show this problem and defines the three types of occlusion problems as they are known in computer graphics applications; ambient, self and view frustum occlusions.

- Ambient occlusion: the polygons are occluded by object nearer the viewpoint.
- Self-occlusion: back facing polygons.
- Frustum occlusion: polygons are out side the field of view.

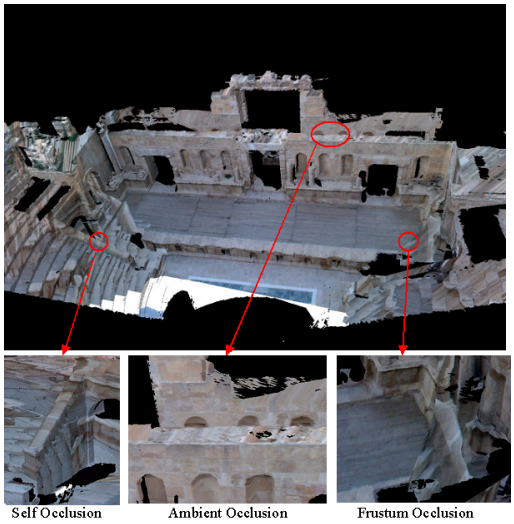


(a)



(b)

Figure 5.6: (a) Types of occlusaion in computer graphics. (b) Visibility from a point and double projection problem.



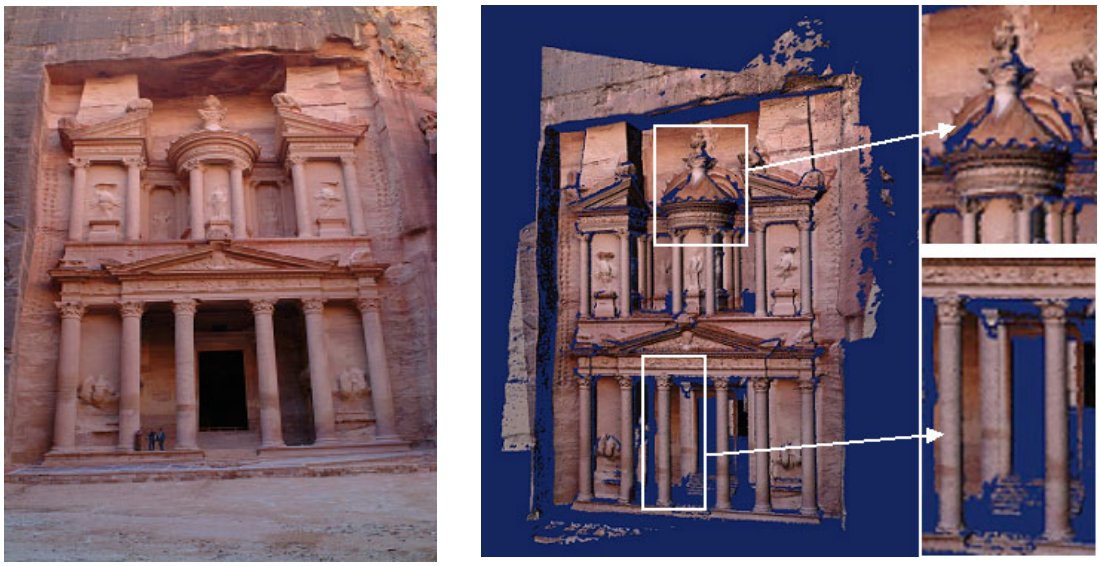


Figure 5.7: Occlusion problems resulted from direct texture mapping of the selected colored images

The occlusion problem can be avoided by manual texture extraction and mapping, however, such tedious task can take up to several days for good results. Fully and semi-automatic texture extraction and placement of the real scene have been presented in different works [Grammatikopulos et al, 2004; Neugebauer & Klein, 1999] but with more complex processing. Those techniques use computer graphics occlusion culling algorithms. In general these algorithms classify under two main approaches; the first approach working in image space takes the advantage of the discredited nature of the output image, computing visibility only to the precision required to decide what is visible at a particular pixel. Examples are the Z-buffer value, ray castings, and Watkins algorithms. Although it is easy to be implemented, Image space algorithms are need computation of degree *O* (n, N) where n is the number of objects (polygons) in the scene and n is the number of pixels. This requires large memory and long computation time, additionally these algorithm show bad performance in a complex scenes. In another hand object space approach use geometric tests on object description to determine the visible surfaces, examples are Painter`s, depth sorting and BSP algorithms. Object space algorithms are O (n^2) where n is the number of object in the scene. Since they compare the objects in the object space, the algorithms show high precision and good results in high complex scenes. But still the complexity of the algorithms is high, and computationally expensive.

5.4 Imagery Segmentation and Matching

The quality of the registration process, which aligns the laser scanner data with the imagery, is a crucial factor for the aspired combined processing. This registration can realized if correspondence coordinates are available in both systems. Since the accurate detection and measurement of point correspondences can be difficult especially for the point clouds from laser scanning, straight lines are measured between the image and the laser data as corresponding elements. The accuracy of the transformation depends on the accuracy with which the features have been extracted from the scans. For providing stable features of interest efficient segmentation algorithms have to be used to extract the features automatically for the co registration of data sets. In our approach we have used the methodology for extracting the 2D straight-line features from range and intensity image using the proposed segmentation algorithm discussed in chapter 4. We are handling the scale thresholds parameters to compromise between localization and noise sensitivity. Figure 5.8 shows an example of some selected edges (lines) used to registre both colored and range images for Petra treasury (Al-Khasneh).

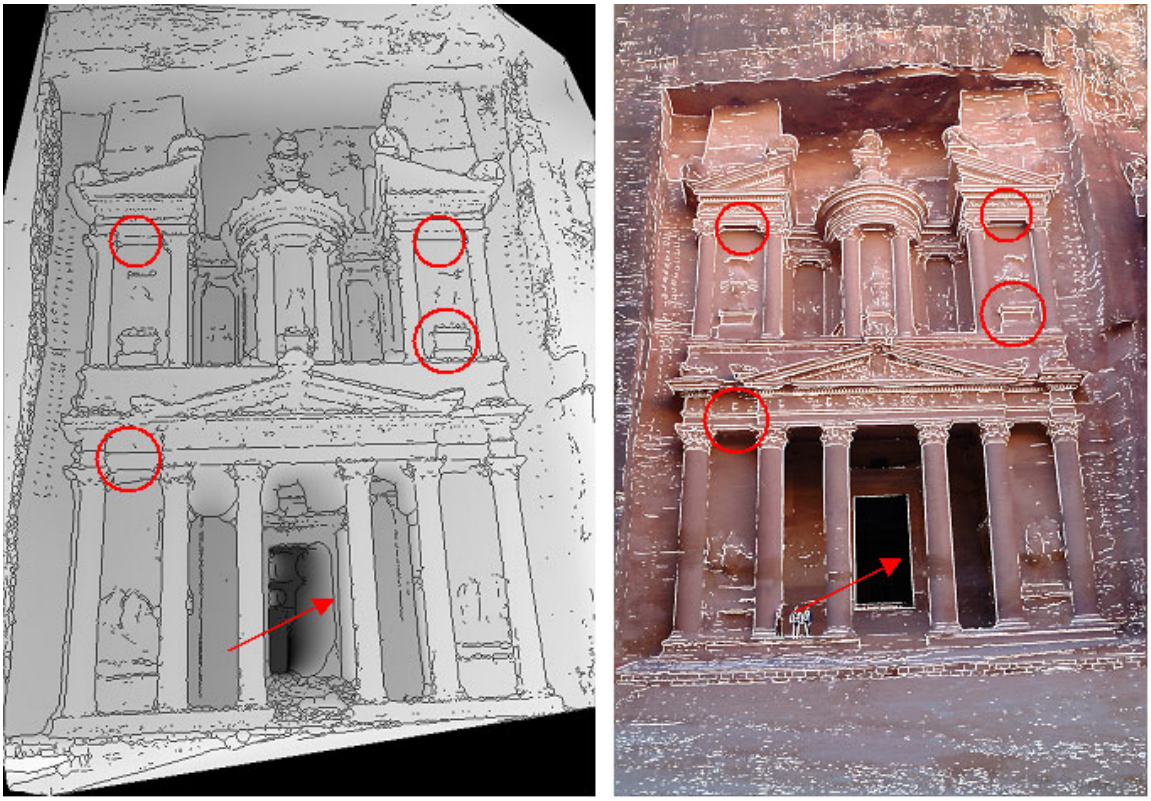


Figure 5.8: Corresponding registration features

These lines are then used by a shape matching followed by a modified spatial resection [Klinec & Fritsch, 2003]. The algorithm transforms the 3D straight lines extracted from the laser data and the corresponding 2D image lines into a parameterized representation:

$$y = \alpha_x \times x + b_x \quad , \quad Y = \alpha_x \times X + \gamma_x \quad , \quad Z = \beta_x \times X + \delta_x \tag{5.1}$$

The principle of the feature based spatial resection process is illustrated in Figure 5.9. For the feature based spatial resection the well-known co linearity equations are utilized. As we are using straight lines as tie-information, the commonly used co linearity equations are modified by the parameterizations (5.1). The resulting equations are shown in (5.2) and (5.3). The camera parameters (xk, yk and ck) in equation (5.3) are fixed values, estimated by a separate camera calibration process.

$$a_x = \frac{r_{31}\left(Y_0 - \alpha X_0 - \gamma\right) + r_{21}\left(\delta - Z_0 + \beta X_0\right) - r_{11}\left(\alpha\left(\delta - Z_0\right) - \beta\left(\gamma - Y_0\right)\right)}{r_{32}\left(-Y_0 + \alpha X_0 + \gamma\right) + r_{22}\left(Z_0 - \delta - \beta X_0\right) + r_{12}\left(\alpha\left(\delta - Z_0\right) - \beta\left(\gamma - Y_0\right)\right)} \tag{5.2}$$

$$b_x = x_h \frac{r_{31}\left(-Y_0 + \alpha X_0 + \gamma\right) + r_{21}\left(Z_0 - \delta - X_0\beta\right) + r_{11}\left(\alpha\left(\delta - Z_0\right) - \beta\left(\gamma - Y_0\right)\right)}{r_{32}\left(-Y_0 + \alpha X_0 + \gamma\right) + r_{22}\left(Z_0 - \delta - \beta X_0\right) + r_{12}\left(\alpha\left(\delta - Z_0\right) - \beta\left(\gamma - Y_0\right)\right)} + y_h \tag{5.3}$$

$$+ c_k \frac{r_{33}\left(-Y_0 + X_0\alpha + \gamma\right) + r_{23}\left(Z_0 - \delta - X_0\beta\right) + r_{13}\left(\alpha\left(\delta - Z_0\right) - \beta\left(\gamma - Y_0\right)\right)}{r_{32}\left(-Y_0 + \alpha X_0 + \gamma\right) + r_{22}\left(Z_0 - \delta - \beta X_0\right) + r_{12}\left(\alpha\left(\delta - Z_0\right) - \beta\left(\gamma - Y_0\right)\right)}$$

To solve the spatial resection problem, a least squares algorithm using the Gauss-Markov-Model was implemented. Here the unknown parameters (X_0, Y_0, Z_0, ω, φ, κ) are estimated and an exact co-registration of object and image space can be provided. But in a least squares approach the non linear equations cannot be used without initial values. To overcome this dilemma and for estimation of rough (initial) exterior orientation parameters the photomodelr software is used. Once the initial values for the spatial resection approach are available. We are able to compute the exterior orientation by the feature based spatial resection. At least three corresponding straight lines are required to obtain a unique solution for the spatial resection.

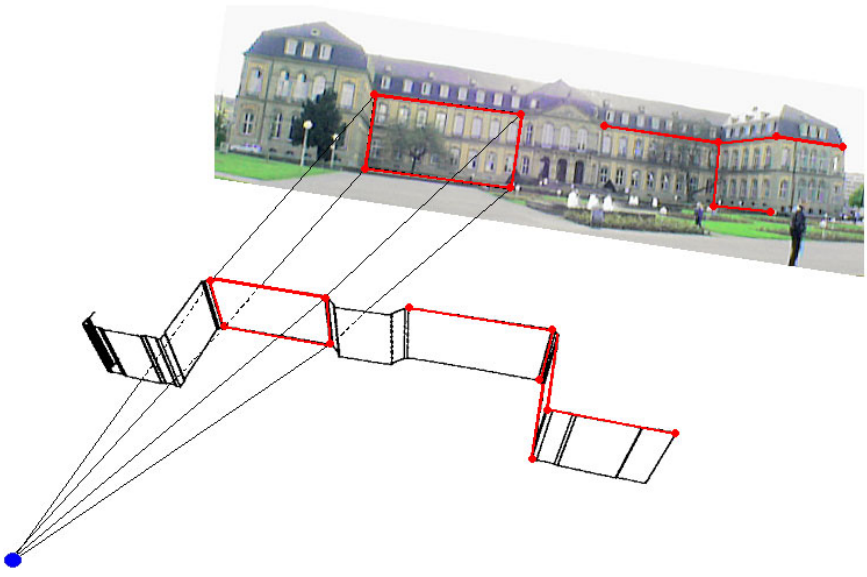


Figure 5.9: Principle of the feature based spatial resection (Klinec & Fritsch, 2003)

5.5 The Proposed Visible Surface Algorithm

In order to handle the resulting problem of occlusions in our texture mapping approach, we have developed an efficient visible surface algorithm. In order to optimize its performance, the algorithm works in both image and object space. The main characteristics of this algorithm are the ability of detecting all types of occlusion problems; ambient, self and view frustum occlusions. The algorithm works as following; after given the exterior orientation parameters and the size of the corresponding image, the algorithm checks the frustum occlusion to exclude the geometry parts that are not within the range of the image. The visible polygons will be stored in a matrix with the z values of their vertices. For self and ambient occlusion two thresholds values have to be defined: image space threshold (T1) represents the maximum number of pixels of triangle edge projected onto the image and object space threshold (T2) represents maximum length of the triangle edge in the scene space (sampling interval). These thresholds values are defined in figure 5.10. The threshold (T1) will define the searching area in the image space so if the threshold is defined by 3 pixels, the searching area will be 9 pixels. The polygons will be sorted with the specified area. The algorithm will search and select the nearest polygon as the visible one depending on the Z value and the threshold (T2) as depicted in figure 5.11. Each polygon will have an ID to classify it as visible or non-visible polygon. This step is important for the further texture processing.

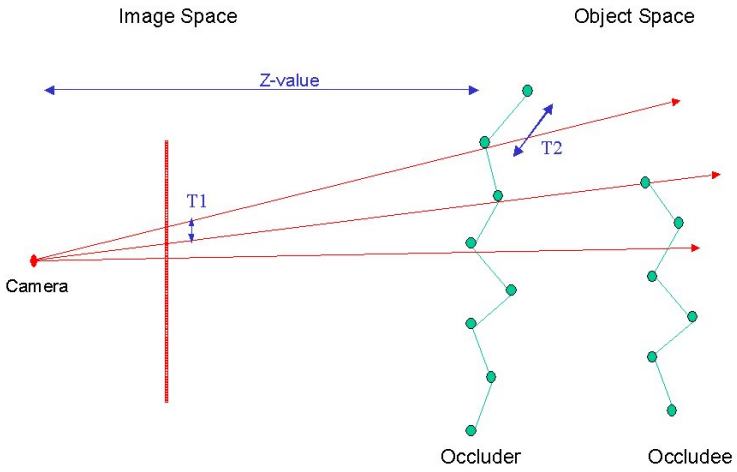


Figure 5.10: Principle of the visibility algorithm

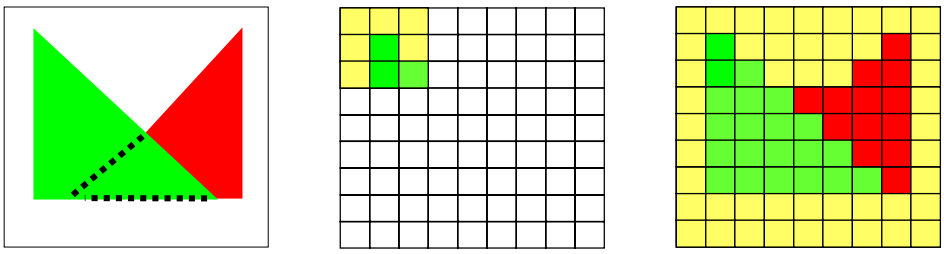


Figure 5.11: The visible algorithm detects the green triangle as the visible one using threshold T1= 3 and searching area = 9 pixels.

The algorithm tries to take the advantages of both image and object space approaches. Comparatively to other algorithms, it has efficient computation time since it does not search pixel by pixel as in the image space algo-

rithms. In addition to that, it is flexible to allow the visibility detection of different resolution of geometric and 2D images. The search area will depend on the resolution of the 3D geometric model, which is usually smaller than the resolution of the 2D images. The algorithm needs computations of degree *O* (S x n, N) where n is the number of objects (polygons) in the scene and n is the number of pixels and *S* is a factor equal (n / searching area) ≤ 1. This flexibility of using the suitable thresholds will speed up detecting the visible polygons. Additionally the algorithm shows high precision and good results in high complex scenes since it detects the visible polygons in object space precision. Although the algorithm shows low sensitivity, choosing smaller threshold values may wrongly classify some visible triangles as occluded. Whereas choosing larger thresholds may classify the occluded triangle as a visible one. The output results of the algorithm will be shown during discussion the over all procedure of the texture approach.

5.6 Summary of the Overall Procedure

Our method of texture mapping is based on the selection of a combination of optimal image patches for each triangle of a 3D model. According to the best possible geometric and radiometric conditions, the locations of the image triangles are computed from object faces via co linearity equations. The procedure consists of three main steps: pre-processing, which includes camera calibration, color corrections and data co registration, visible surface computing and texture warping.

For our investigations the interior orientation parameters were computed using laboratory pre-calibrated targets and Australis software. These parameters were used to creating zero distortion images. In order to achieve nearly the same color balance in all images, a histogram matching is performed. This function makes the brightness distribution of the two images as close as possible. The resulting histogram of the display where the function was started is changed to match the current (source) histogram of the selected image display. The histogram source is selected as the input histogram. In general for the procedure, the following input data are required:

- Triangulated 3D mesh in the form of successive XYZ, describing the object surface.
- Calibrated colored images along with their exterior orientation parameters.
- Visible surface algorithm.

The different steps, which summarized in figure 5.12, will be presented exemplarily for the North theatre and Petra treasury.

a) Select most appropriate image that depicts most parts of the object (master image). The aim of choosing master image is to texturing large numbers of visible primitives. This will speed up the further processing steps for the north theatre the master image depicted in figure 5.13.

b) Given the exterior orientation parameters, the texture coordinates associated to each polygon vertex can be located using the co-linearity equations.

c) Check for frustum occlusion to exclude texturing the coordinates that are not within the range of the image. Store the other texture coordinates in a matrix with the Z value of the corresponding vertex.

d) In image space define searching area using threshold (T1), search and sort all the triangles within the specified area, then select the nearest triangle dependent on the Z value and the (T2), give an ID for such polygons as an occluder one. The occluded polygons vertices will take a zero textured values. Figure 5.14 shows the results of mapping the master image onto the 3D model of the North theatre after detecting the three types of occlusion.

e) The occlude polygons will be masked out from the original model with their connectivity, as depicted in figures 5.15 and 5.16.

f) The separated parts will be used as input model for the second selected textured images. The process for visibility and separating of occluding parts will be repeated automatically until the model is textured from all the available images as it can be seen in figures 5.17 - 5.24. Then separated parts merge with the master image part to have the final model, see figure 5.25.

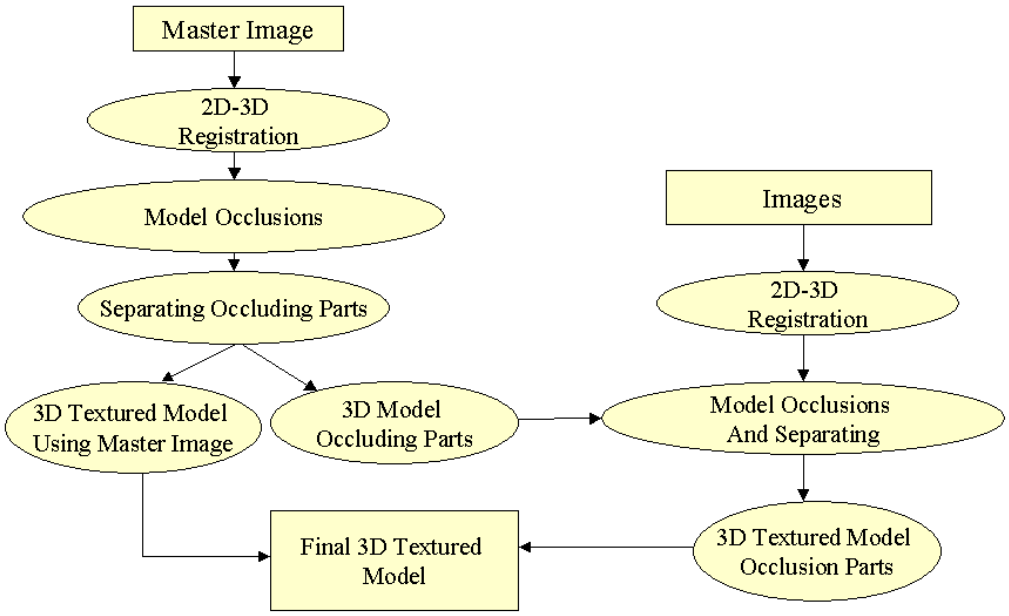


Figure 5.12: Texture mapping procedure

One crucial point during the separating process is to maintain connectivity of the polygons; the coordIndex of the model. This is realised by preserving the ID of each polygon during all processing steps especially the polygons in the border between the splitting areas as it is depicted in figures 5.27 and 5.28. Although we have tried to optimize the collection of the images there are still some observable discontinuity in the color and the brightness between the images as depicted in figure 5.29. For this reason we have used ENVI 4.1 software to perform the histogram equalization for all three channels in all textured images in order to smoothen the colour transitions between the different images as depicted in figure 5.30. The performance analysis has been conducted on a standard PC with an Intel 4, 3GHz Processor, and 1GB RAM. In our approach VRML modelling language is used. The approach has been also demonstrated for texture mapping the Petra treasury as it can be seen in figures 5.31- 5.38.

Figure 5.13: Master image for North theatre –Jerash

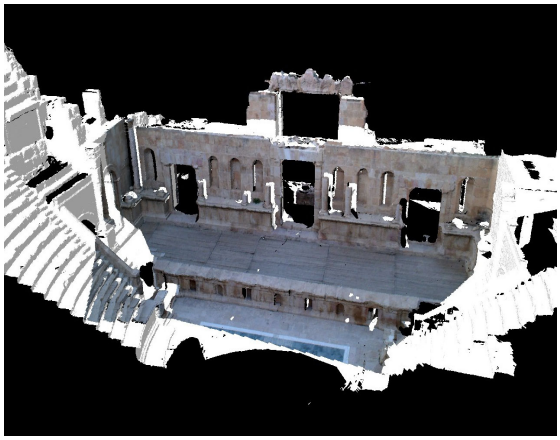


Figure 5.14: 3D Textured model for North theatre using master image after detecting all types of occlusions

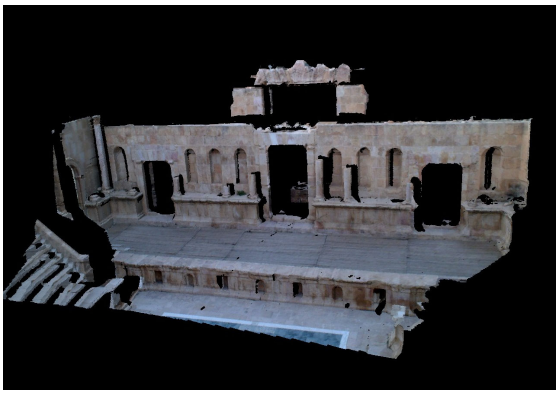


Figure 5.15: 3D Textured model using master image after separating of the occluding parts

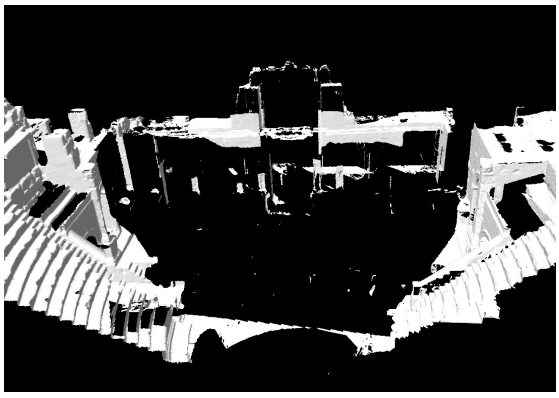


Figure 5.16: 3D Separated occluding parts model

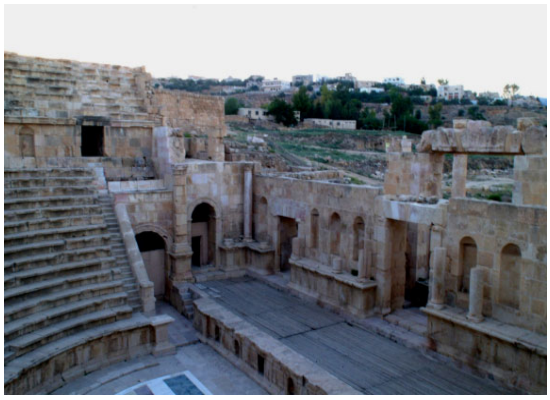


Figure 5.17: The second selected image for texture mapping of the North theatre

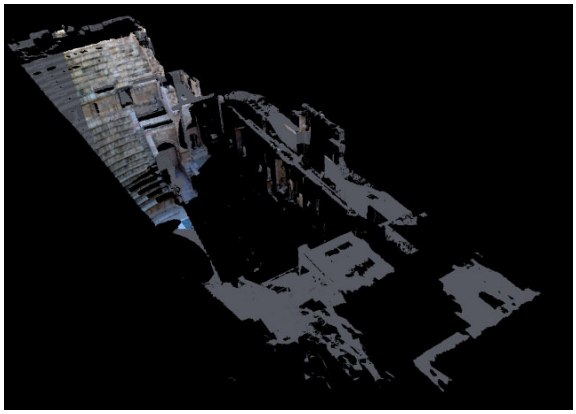


Figure 5.18: 3D Textured model using the second selected image after detecting all types of occlusions

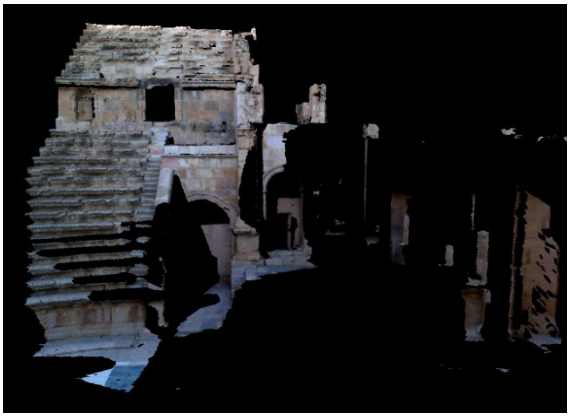


Figure 5.19: 3D Textured model for the north theatre of Jerash using the second selected image after separating of the occluding parts

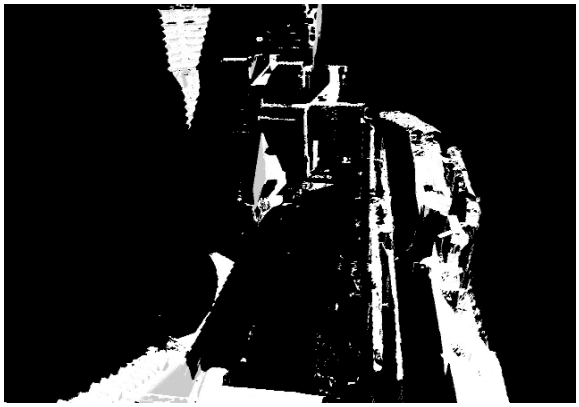


Figure 5.20: 3D Separated occluding parts model of the North Theatre of Jerash separating the textured model from the master and second images

Figure 5.21: The third selected image for texture mapping the North Theatre

Figure 5.22: 3D Textured model for the north theatre of Jerash using the third selected image after separating of the occluding parts

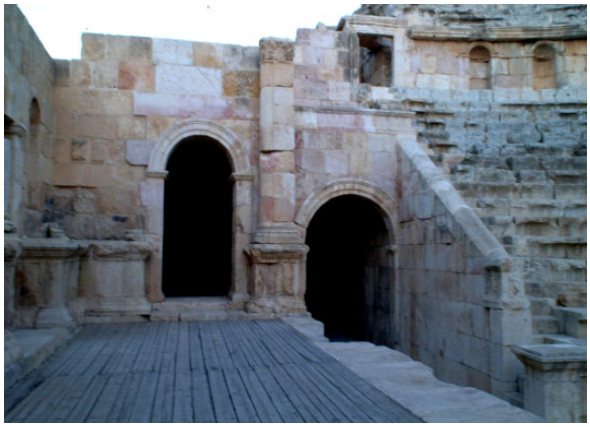


Figure 5.23: The forth-selected image for texture mapping the North Theatre

Figure 5.24: 3D Textured model for the north theatre of Jerash using the forth-selected image after separating of the occluding parts

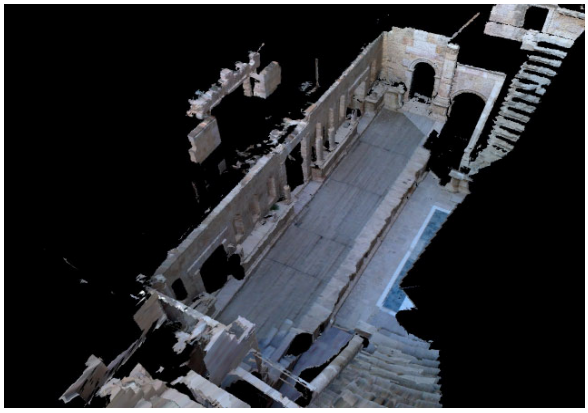


Figure 5.25: 3D Textured model for the north theatre of Jerash using 4 images

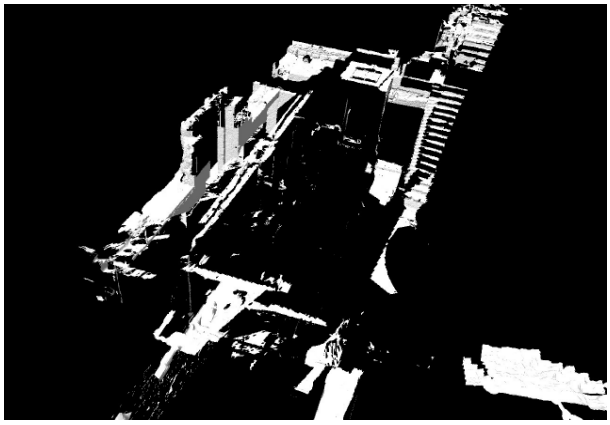


Figure 5.26: Separating of the occluding parts

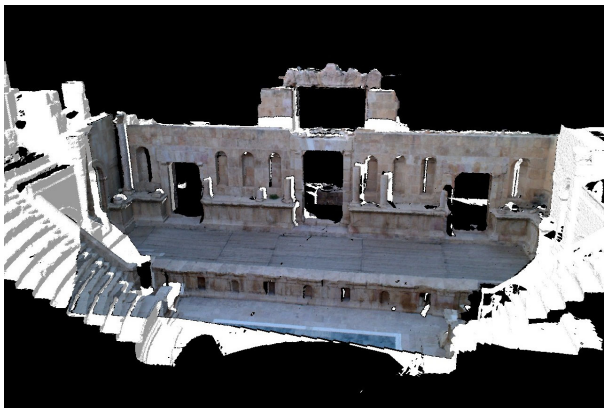


Figure 5.27: Separating the 3D model shows borders gaps between the splitting parts

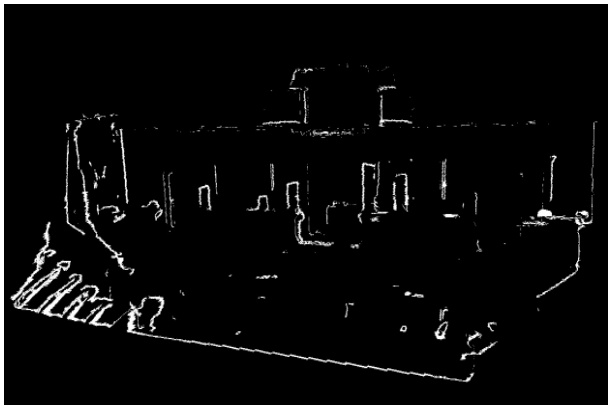


Figure 5.28: Preserving the connectivity of the borders between the splitting parts

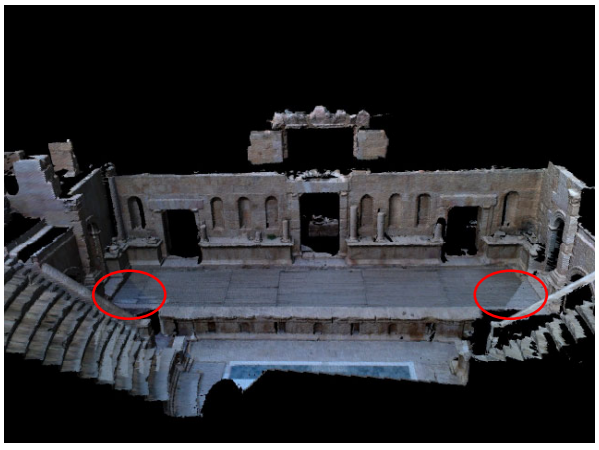


Figure 5.29: 3D Textured model for the north theatre of Jerash using 4 images before colour correction.

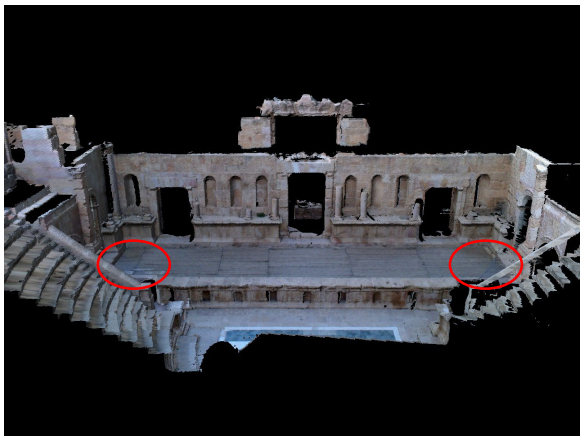


Figure 5.30: 3D Textured model after colour correction

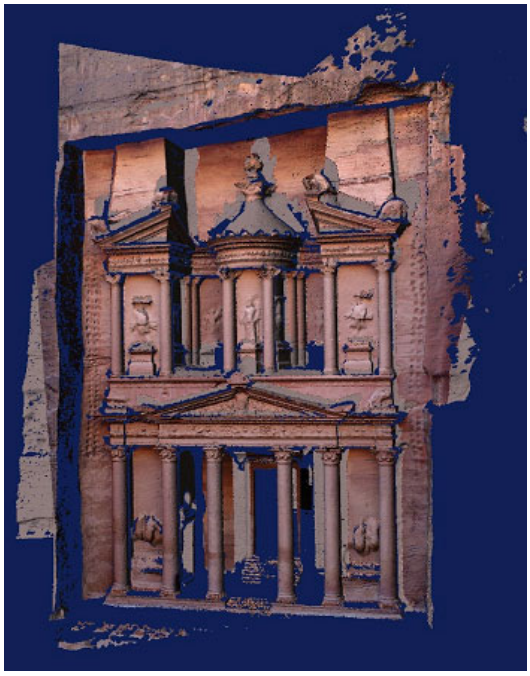


Figure 5.31: 3D Textured model for Petra treasury using master image after detecting all types of occlusions

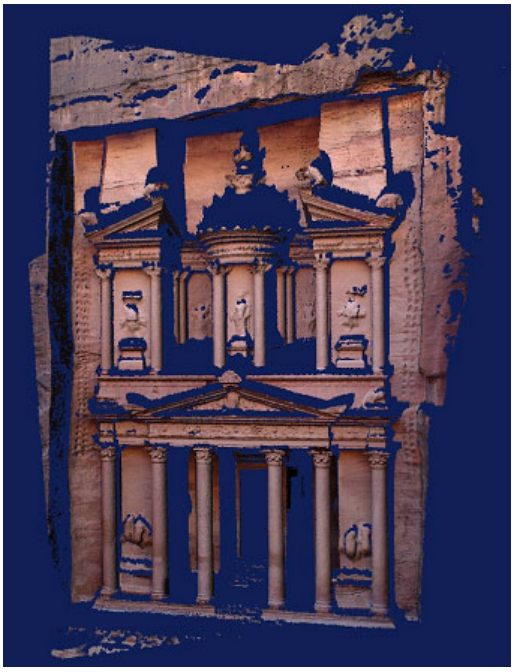


Figure 5.32: 3D Textured model using master image after separating of the occluding parts

Figure 5.33: 3D Separated occluding parts model of the Petra treasury of Jerash.

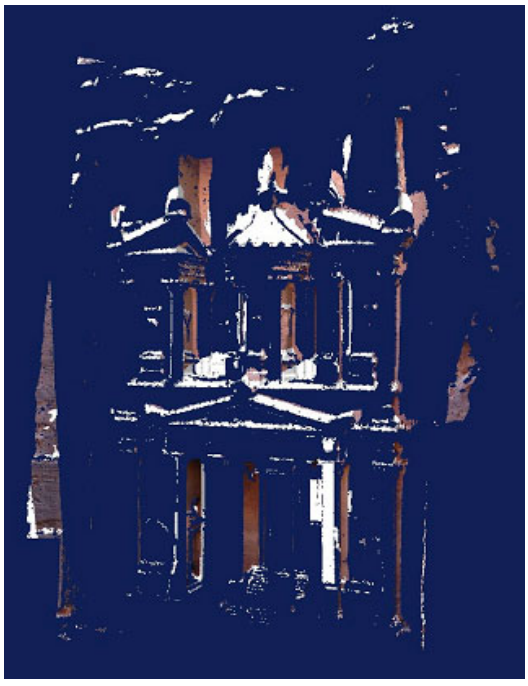


Figure 5.34: 3D Textured model using the second selected image after detecting all types of occlusions

Figure 5.35: Border gaps between the splitting parts

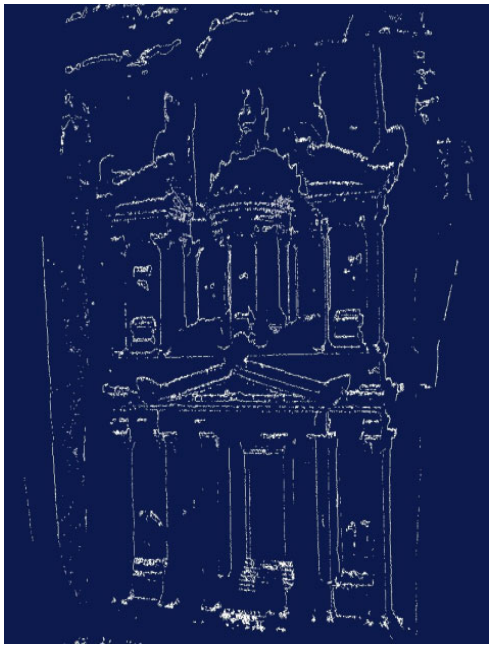


Figure 5.36: Preserving the connectivity of the borders between the splitting parts

Figure 5.37: Texturing 3D model of occluding parts using 3 images

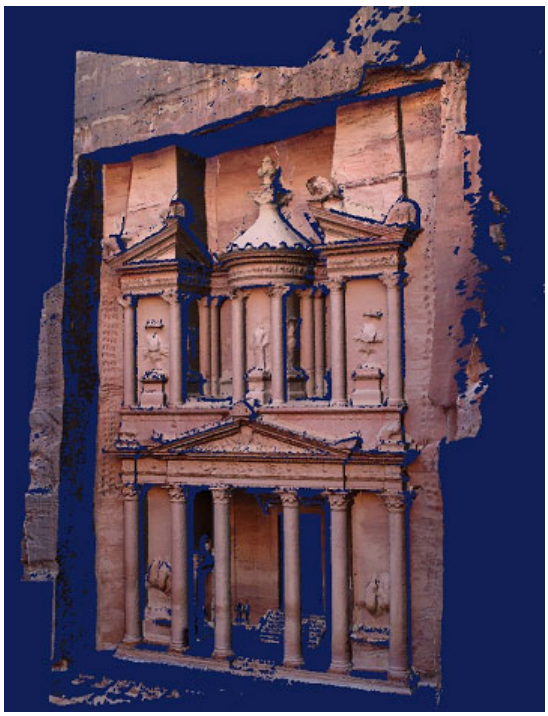


Figure 5.38: Final textured model of Alkasneh using 4 images

6 Automatic Surface Features Extraction

6.1 Introduction

Within the chapter the potential of combining terrestrial laser scanning and close range photogrammetry for the documentation of heritage sites is discussed. Besides improving the geometry of the model, the integration aims on supporting the visual quality of the linear features like edges and cracks in the historical scenes. Although the laser scanner gives very rich surface details, it does not provide sufficient data to construct outlines for all surface features of the scanned object, even though they are clearly defined in the reality. In the contrary, the digital photogrammetry is more accurate in outline rendition and in contrast to terrestrial laser scanning, which requires a relatively high effort for hardware installation; they can be captured easily from different viewpoints. For this reason, it can be advantageous to complete a geometric model, which has been generated from the laser measurement based on intensity images, which are captured independently from the range data. By these means object geometry, which is not available in the range data due to occlusions can be provided based on photogrammetric measurements.

Thus, the highest possible degree in efficiency and flexibility of data collection will be possible, if both techniques are combined during data processing. In our approach this integration helps to improve the geometry and visual quality of the collected 3D model. During data collection the information on edges and linear surface features like cracks, reliefs and deteriorated edges is based on the analysis of the images, whereas information on object geometry is provided from the laser data. Additionally, areas, which are not accessible in the laser scanner data due to occlusions are added based on semiautomatic evaluation of the imagery. By these means, a complete 3D features for the scene can be generated with sufficient and clear details. Within this chapter, problem statement is presented in section 2. Section 3 exemplarily presents our feature extraction approach using the hybrid system. The pre-processing of the relevant image and LIDAR data is discussed. This pre-processing is mainly required in order to have co register laser and image data. To optimize the transformation between the 2D images and 3D range scans, the transformation is based on a match of 3D with 2D straight-line features that maximize the overlap criterion. Section 4 presents image based semi-automated techniques in order to bridge gaps in the laser scanner data in order to add new details, which is required to build more realistic perception of the scene volume.

6.2 Problem Statement

Although the 3D model produced by laser scanner contains a large number of triangles, which are representing the surfaces, it can still difficult to recognize and localize the outlines of the surface features. An example for this type of features, which are clearly visible in an image such as cracks, deteriorated edges, and reliefs, is depicted in figure 6.1. This data was collected from the left door of Al-Khasneh using calibrated Fuji S1 Pro camera, which provides a resolution of 1536X2034 with focal length of 20mm. As it can be seen from the corresponding 3D meshed model shown in figure 6.2, the cracks and the edges outlines are lost beyond the resolution in the available laser data despite the fact that the data contains 1.1 million points. The fact of using range image segmentation algorithms alone for detecting and extracting these features are not worth; for algorithms depend on calculating normal vectors, these features have irregularities and small variation of normal vectors, which is considered the main criteria for edge detection algorithims, usually such algorithm deal with clear palners surfaces. In another hand, for the segmentation algorithms depend on depth discontinuity and partial derivatives, these features also have very small variation depth values. Thus, using large threshold values will not access them as can be depicted in figure 6.3. Additonally, using small threshold values may extract some of those features, but with over segmentation problem, as can be depicted in figure 6.4.

Figure 6.2 also showed that some 3D features of the scene are missing, example is the inner edge of the right column of the door. This is due to the fact that limitations of 3D range sensor positioning (standoff distance, maximum distance) will cause constrains to have complete data collection.

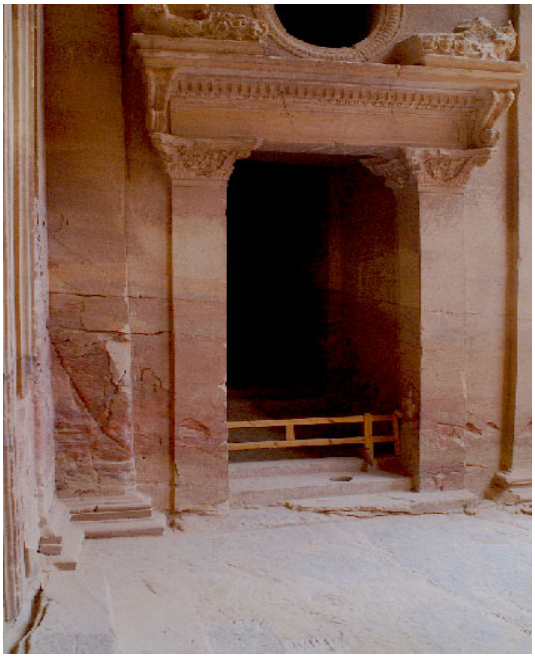


Figure 6.1: The left door of Al-Khasneh 1536X2034 pixel

Figure 6.2: Meshed model for the left door

Figure 6.3: Range image segmentation using large threshold value

Figure 6.4: Range image segmentation using small threshold value (over segmentation problem)

6.3 Hybrid Approach for Automatic Surface Feature Extraction

In order to support the visual quality of such details, a hybrid approach combining data from the laser scanner and the digital imagery was developed. For integration all data sets have to be co registered in the first processing step. This is realized by aligning the extracted straight edges from both data sources. After position and orientation parameters are computed for the sensor stations, distance images are generated from the point cloud in order to

provide the missing third dimension in the available images. Finally, an integrated segmentation process based on the image data is used in order to support the extraction of the details and the surface features outlines from distance images. The left door of the Al-Khasneh is exemplarily processed.

6.3.1 Data Co-registration

In our approach, the input consists of the pair of range image in addition to 2D colored images collected freely. We assume that both the camera and range sensors view the same part of the real scene; so that the 3D and 2D views have significant overlap. In addition, we assume that we know the internal parameters of the camera, such as focal length and radial distortion, with which we can convert the real camera into an ideal pinhole camera. The methodology for extracting the 2D straight-line features from range and intensity image is done using the proposed segmentation algorithm presented in chapter 4. We are handling the scale thresholds parameters to compromise between localization and noise sensitivity; the result of the segmentations are shown in figures 6.5 and 6.6.

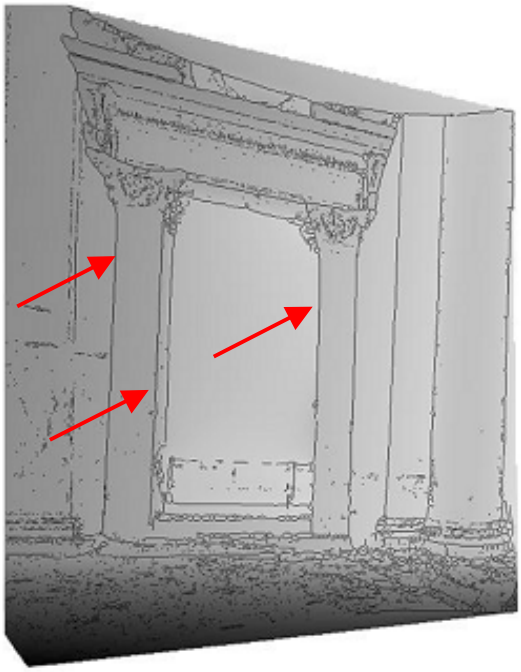


Figure 6.5: Range image segmentation, the features used for registration are indicated by red arrows.

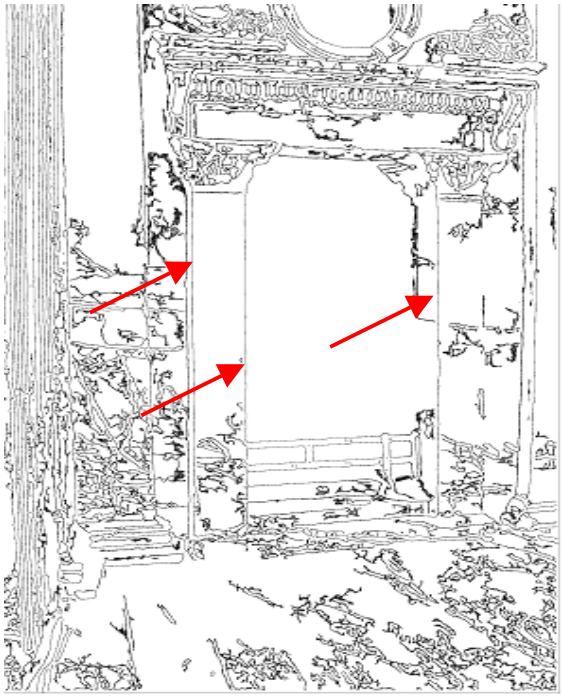


Figure 6.6: Intensity image segmentation, the corresponding features used for registratioin are indicated by red arrows.

These lines are then used by a shape matching followed by a modified spatial resection [Klinec & Fritsch, 2003]. The algorithm transforms the 3D straight lines extracted from the laser data and the corresponding 2D image lines, which are given by two points, into a parameterized representation. Then the unknown exterior parameters of the image are determined by spatial resection. In order to solve the spatial resection problem, the least squares algorithm-using Gauss Markov Model with initial values of the exterior orientation parameters is implemented. For more details see section 5.4.

6.3.2 Distance Images Generation

After the camera coordinates and orientation parameters are registered in the laser scanner coordinate system, the collinearity equations are applied to generate a distance image based on the available point clouds. For our exemplary scene, the point cloud was collected from a close viewpoint resulting in 1 cm average resolution. The distance images were generated with the same number of pixels of the corresponding images (1536x2304 pixels). For the left door of Al-Khasneh, three distance images were generated for the three images available from different camera stations. Figure 6.7 depicts two of these distance images projected on the corresponding colored images.

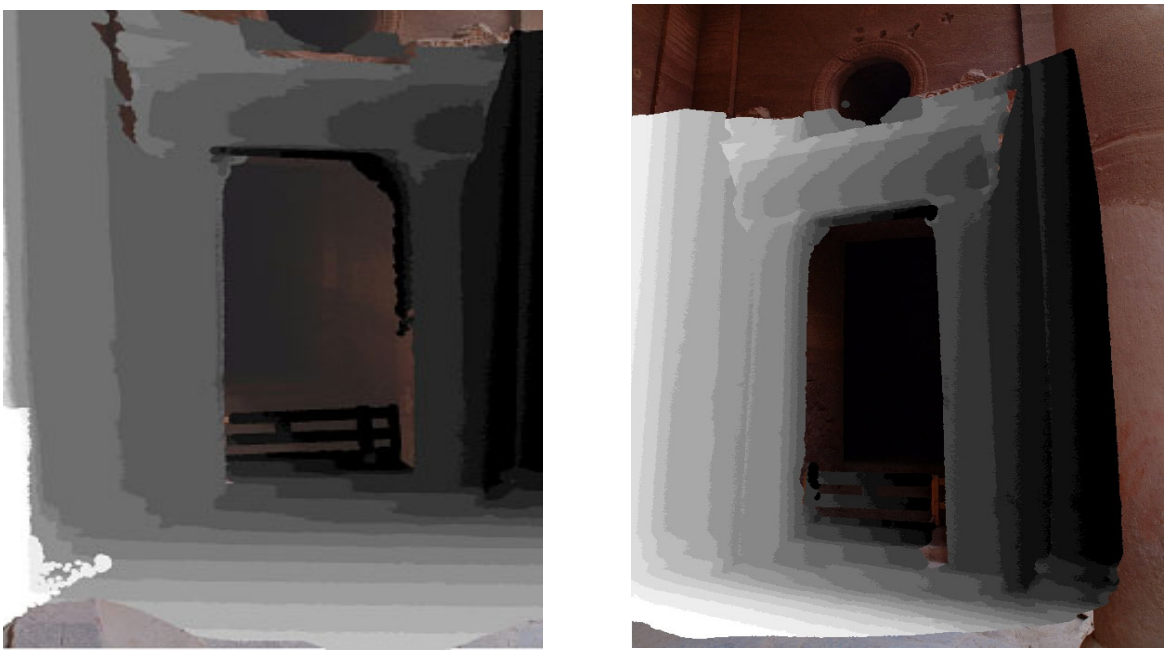


Figure 6.7: Distance images projected on the corresponding images.

6.3.3 Integrated Segmentation Process

The segmentation process is used to automatically extract the 2D coordinates of the surface linear features from three different imagery images. Then the third dimension of the segmented outlines is provided from the distance images. Figure 6.9, depicts the final 3D features extracted for the left door of Al-Khasneh. It can be seen that the data contains all of the edges and linear surface features in a clearly outlines. In total the feature-based representation contains 143.6 thousand points, whereas the original point cloud of the same portion has 1.1 million points.

6.4 Occluded Features from Image Based Measurements

It can be seen from the 3D features presented in figure 6.2 that due to the position of the laser scanner, the inner edge of the right column of the door has no data. The occluded edge can be added based on semiautomatic evaluation of digital imagery to have a complete data set for the scene. In our approach, the 3D coordinates of initial points were extracted manually. Then an automatic stereo matching has been applied for closely spaced images to add more points on the edge. For matching within the segmented parts of the edge epipolar constraint is used. The occluded edge is added to the 3D features as it can be shown in right part of figures 6.8 and 6.9.

Figure 6.8: Image Based Measurements approach in order to add the inner edge to the 3D features

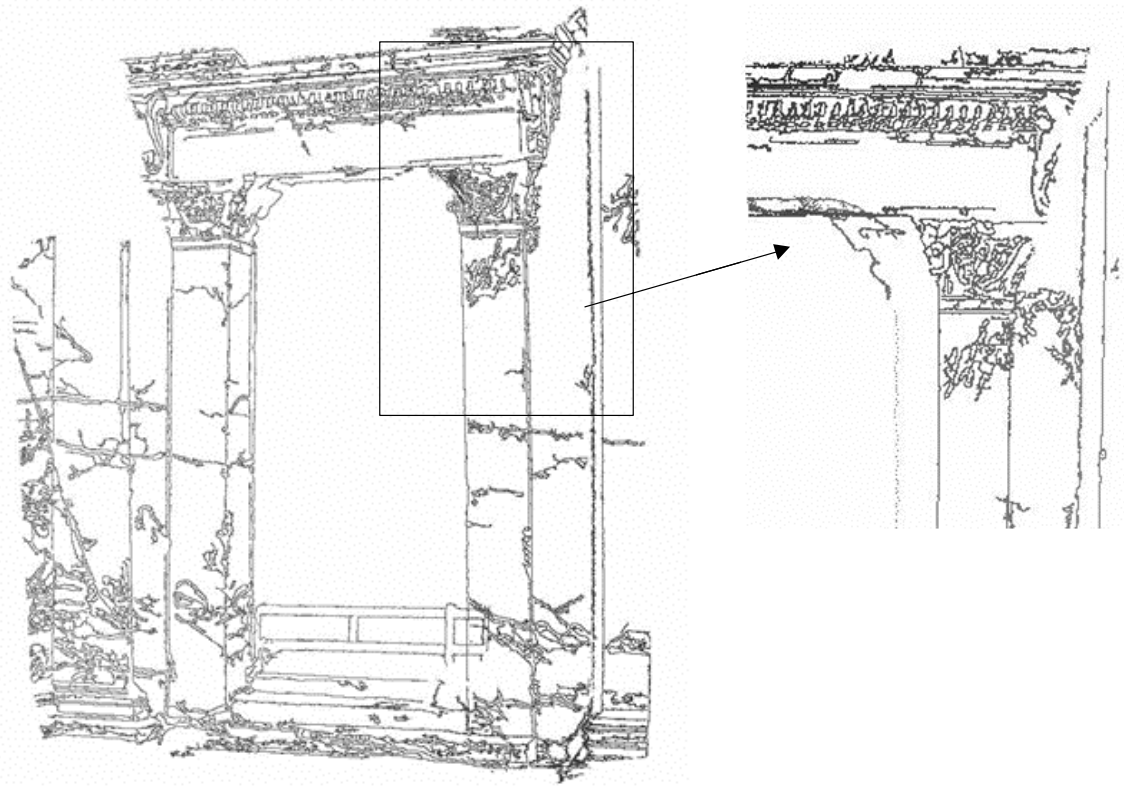


Figure 6.9: 3D features extracted for the left door of Al-Khasneh, the inner edge of the column is added from image based measurements.

7 Conclusion and Future Directions

Within the thesis the potential of combining terrestrial laser scanning and close range photogrammetry for the documentation of heritage sites is discussed. The integration aims to optimize the advantages of the both technique in order to obtain high quality textured 3D models for documentation and visualization purposes. In addition the hybrid approach could be used for supporting the visual quality of the linear features like edges and cracks in the historical scenes. Additionally, we present an effective approach for projective texture mapping from photographs onto triangulated surfaces from 3D laser scanning. By these means, the effort to generate photorealistic models of complex shaped objects can be reduced considerably. In order to handle the resulting problem of occlusions, the visibility of the model areas in the respective images has to be established. For this purpose an efficient algorithm addressing the image fusion and visibility is used.

In order to simplify the dense datasets and to provide stable features of interest, which are used to recover the positions of the 2D cameras with respect to the geometric model for texture mapping task, the research presents an efficient edge detection algorithm for the extraction of linear features in both range and intensity image data. In our algorithm the required features of interest are extracted by an analysis of the mean curvature values. This additionally allows the discrimination of different edge types like crease or step edges. As it has been demonstrated, the algorithm features computational efficiency, high accuracy in the localization of the edge points, easy implementation, and robustness against noise. The algorithm was initially developed for range image segmentation and has been extended to segment intensity images with some modifications. The generality and robustness of the algorithm is illustrated during processing of complex cultural heritage scenes.

However in the broader concept of creation high realistic 3D modeling of the archaeological sites and historical buildings, there are still issues that need to be investigated and studied in order to have further improvements and developing. These issues can be summarized as the follows:

- Further testing and improving could be applied for an efficient and accurate localizing of the segmented edges using hierarchical data structure, which could be also proposed for visible surface algorithm technique.
- Developing a new method for the generating of an accurate and realistic 3D CAD modeling of the historical buildings using the available terrestrial laser data. This step is needed in order to reduce the data size for real-time visualization purposes. By these means, the possibility for modeling and real visualizing of the whole architectures of wide heritage sites could be possible, an example is the ancient Petra or ancient Jerash city which has tenths or hundreds of buildings, tombs, temples and theaters. In this framework hybrid approach of edge based and region growing segmentation could be used for extraction and fitting the scene surfaces. In the approach our edge detection technique may carry out to guide and improve the clustering technique applied to the region-growing algorithm.
- One of the techniques that could be farther integrated is the aerial photogrammetry. The technique could be used for the 3D visualization of virtual landscapes including additional objects like upstanding remains or finds in the context of heritage documentation. In addition it could be used for supplying color images for texture mapping process.
- Integration of the archaeological information system with suggested approach for realistic 3D virtual models of the respective heritage sites will considerably simplify the access to archaeological information especially for non-expert users.

Thus, it can be concluded that this type of applications can be spread further, if different tools for the efficient collection of the required data are available

8 Bibliography

Abdelhafiz, A., Riedel, B. & Niemeier, W. [2005], 3D Image Approach as a Result From The Combination Between The Laser-Scanner Point Cloud And The Digital Photogrammetry.7th Conf. Optical 3-D Measurement Techniques, Vienna.

Adams, R. & Bischof, L. [1994], Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6): 641--647.

Akca, D., [2003], Full automatic registration of laser scanner point clouds. Optical 3-D Measurement Techniques VI, 22-25 September, Zurich, pp. 330-337.

Allen, P., Stamos, I., Troccoli, A., Smith, B., Leordeanu, M. & Hsu, Y. [2003], 3d modeling of historic sites using range and image data. In Proceedings of International Conference on Robotics and Animation (ICRA).

Alshawabkeh, Y. & Haala N. [2004], Integration Of Digital Photogrammetry And Laser Scanning For Heritage Documentation. In: Proceedings of the XXth ISPRS Congress, Istanbul, Turkey, pp. 424-429.

Alshawabkeh, Y. & Haala N. [2004], Laser Scanning and Photogrammetry: A Hybrid Approach for Heritage Documentation. 3rd International Conference on Science & Technology in Archeology & Conservation, Amman, Jordan.

Alshawabkeh, Y. [2005], Using Terrestrial Laser Scanning for the 3D Reconstruction of Petra / Jordan. Photogrammetric Week '05, Wichmann, Heidelberg, pp. 39-47.

Alshawabkeh, Y. & Haala N. [2005], Automatic Multi-Image Photo Texturing of Complex 3D Scenes. CIPA IAPRS Vol. 34-5/C34, pp. 68-73.

Amenta, N., Bern, M. & Eppstein, D. [1999], "Optimal Point Placement for Mesh Smoothing", in *"Journal of Algorithms"*, Volume 30, pp 302-322.

Ansar. A. & Danilidis, K. [2002], Linear pose estimation from points or lines. In A. Heyden et al., European Conference on Computer Vision, volume 4, pages 282–296, Copenhagen, Denmark.Springer.

Aoki, M. & Levine, M. [1978], ‘‘Computer Generation of Realistic Pictures’’, *Computers and Graphics*, Vol. 3, pp. 149-161.

Appel, A. [1967], "The Notion of Quantitative Invisibility and the Machine Rendering of Solids," Proc. *ACM Narwhal Conference,* 387-393.

Atlas Tours. Net (http://www.atlastours.net/jordan/.).

Baccar, M. [1994], “Surface characterization using a Gaussian weighted least squares technique towards segmentation of range images,” Master's thesis, The University of Tennessee, Knoxville.

Baumberg, A. [2002], Blending images for texturing 3D models. Proc. Conf. on British Machine Vision Association, Cardiff, pp 404-413.

Belloni, B. [2000], Jerahs, the Heritage of Past Cultures. Plurigraf, Terni. Italy.

Benlamri, R. & Al-Marzooqi,Y. [2003], 3-D Surface Segmentation of Free-Form Objects using Implicit Algebraic Surfaces, Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), Sydney.

Bernardini, F. & Bajaj, C. [1997], Sampling and reconstructing manifolds using alpha-shapes. Tech. Rep. CSD-TR-97-013, Department of Computer Sciences, Purdue University. Revised version submitted for publication.

Bernardini, F., Martin, I. & Rushmeier, H. [2001], High-quality texture reconstruction from multiple scans. IEEE Trans. On *Visualization & Computer Graphics*, Vol. 7(4), pp. 318-332.

Beraldin, J., Picard, M., El-Hakim, S., Godin, G., Latouche, C., Valzano,V. & Bandiera, A.[2002], “Exploring a byzantine crypt through a highresolution texture mapped 3-D model: combining range data and photogrammetry,” in Proc. Int.Workshop on Scanning for Cultural HeritageRecording, Corfu, Greece.

Bergevin, R., Soucy, M., Gagnon, H. & Laurendeau, D. [1996], Towards a General Multi-View Registration Technique. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(5), pp. 540-547.

Besl, P. J. [1988], *Surfaces in Range Image Understanding*, Springer-Verlag, New York.

Besl, P. & Jain, R. [1988], Segmentation Through Variable-Order Surface Fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(2): 167–192.

Besl, P. & McKay, N. [1992], A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 14, n. 2, pp. 239-256.

Blinn, J. & Newell, M. [1976], ‘‘Texture and Reflection in Computer Generated Images’’, *CACM*, Vol. 19, No. 10, Oct. pp. 542-547.

Blinn, J. [1978] ‘‘Simulation of Wrinkled Surfaces’’, *Computer Graphics*, (SIGGRAPH ’78 Proceedings),Vol. 12, No. 3, Aug. pp. 286-292.

Boehler, W., Heinz, G. & Marbs, A. [2001], The potential of non-contact close range laser scanners for Cultural heritage recording. CIPA International Symposium, Potsdam.

Boehler, W., & Marbs, A. [2002], 3D scanning instruments. ISPRS/CIPA International Workshop on Scanning for Cultural Heritage Recoding , Corfu, Greece, pp.9-12.

Boehler W. [2005], Comparison of 3D laser scanning and other 3D measurement techniques. In: Proceedings of International Workshop on Recording, Modeling and Visualization of Cultural Heritage, Ascona, Switzerland, http://www.ascona2005.ethz.ch.

Böhm, J., Fritsch, D. [2003], Automated Interpretation of Dense Range Data in: 'Electronic Imaging, Videometrics VIII', Vol. 5019, pages 77-86, SPIE, Santa Clara, USA.

Boehm, J. [2004], Multi-image Fusion for Occlusion-free façade texturing. IAPRS Vol. XXXV, part 5, pp.867-872.

Böhm, J., Haala, N. & Alshawabkeh, Y. [2005], Automation in laser scanning for cultural heritage applications in: International Workshop on Recording, Modeling and Visualization of Cultural Heritage, Ascona, Switzerland, May 22-27.

Bolles, R. & Fischler, M. [1981], “A RANSAC-based approach to model fitting and its application to finding cylinders in range data,” in Proc. 7th Int. Joint Conf. Artificial Intelligence, Vancouver, B.C., Canada, Aug. 24-28, pp. 637-643.

Bolles, R., Horaud,P. & Hannah, M. [1983], “3DPO: A three-dimensional part orientation system,” Proc. of 8th International Joint Conf. on Artificial Intelligence (Karlsruhe, West Germany), 1116-1120.

Borg, C., Cannataci & J. [2002], Thealasermetry: a hybrid approach to documentation of sites and artefacts. CIPA-ISPRS Workshop on Scanning for Cultural Heritage Recording, Sept., Corfu, Greece, pp. 93-104.

Bornaz, L., Lingua, A. & Rinaudo, F. [2002], A new software for the automatic registration of 3D digital models acquired using laser scanner devices. *Proc. CIPA WG 6 Int. WS Scanning for Cultural Heritage Recording*, Corfu, September 1-2, pp. 52-57.

Bornik, A., Karner, K., Bauer , J., Leberl, F. & Mayer, H. [2001], High-quality texture reconstruction from multiple views. *IEEE Transactions on Visualization and Computer Graphics* 12,263–276.

Boughorbal, F., Page, D., Dumont, C. & Abidib, M. [1999], Registration and integration of multi-sensor data for photorealistic scene reconstruction. In Proceedings of SPIE Conference on Applied Imagery Pattern Recognition.

Brenner, K. & Haala, N. [1998], Rapid acquisition of Virtual reality city models from multiple data sources. IAPRS, Vol., XXXII, part 5, pp. 323-330

Bruni, V., Vitulano, D. & Maniscalco, U. [2004], Fast Segmentation and Modeling of Range Data Via Steerable Pyramid and Superquadrics. The 12-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2004, WSCG 2004, University of West Bohemia, Campus Bory, Plzen-Bory, Czech Republic, pp 73-80.

Bunker, M., Economy, R. & Harvey, J. [1984], Cell texture—Its impact on computer image generation. In *Proceedings of the Sixth Interservice/ Industry Training Equipment Conference* (National Security Industrial Association, Washington, DC, Oct.), 149–155.

Burt, P. [1981]‘‘Fast Filter Transforms for Image Processing’’, *Computer Graphics and ImageProcessing*, Vol. 16, pp. 20-51.

Canny, J. [1986], A Computational Approach to Edge Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, Nov.

Carbonnell, M. [1989], Architectural Photogrammetry. *In Non-topographic Photogrammetry*, Ch. 19, pp. 321-347, Second Edition, American Society For Photogrammetry and Remote Sensing, (USA: H. M. Karara, Science and Engineering Series).

Carey, R. & Greenberg, D. [1985], ‘‘Textures for Realistic Image Synthesis’’, *Computers and Graphics*, Vol. 9, No. 2, pp. 125-138.

Catmull, E. [1974], A Subdivision Algorithm for Computer Display of Curved Surfaces, PhD thesis, Dept. of CS, U. of Utah.

Catmull, E. [1975], Computer display of curved surfaces. In Proceedings of the IEEE Conference on Computer Graphics, Pattern Recognition, and Data Structure, 11–17.

Catmull, E. & Smith, A. [1980], ‘‘3-D Transformations of Images in Scanline Order’’, Computer Graphics, (SIGGRAPH ’80 Proceedings), Vol. 14, No. 3, pp. 279-285.

Chen. J. & Medioni, G. [1989], "Detection, Localization,and Estimation of Edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol PAMI-11, no 2, pp.191-198.

Chen, H. [1991], Pose determination from line-to-plane correspondences: existence solutions and closed-form solutions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6), 530–541.

Chen, Y. & Medioni, G. [1992], Object Modeling by Registration of Multiple Range Images. Image and Vision Computing, 10(3), pp. 145-155.

Christy, S. & Horaud, R. [1996], Euclidean shape and motion from multiple perspective views by affine iterations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1098–1104.

Chua, C. & Jarvis, R. [1996], 3D free form surface registration and object recognition. *International Journal of Computer Vision*, 17 (1), pp. 77-99.

Cohen, J., Tchou, C., Hawkins, T. & Debevec, P. [2001], Realtime High Dynamic Range Texture Mapping. Eurographics Rendering Workshop.

Coorg, S. & Teller, S. [1999], Extracting textured vertical facades from controlled close-range imagery. IEEE Proc. Of CVPR, pp. 625-632.

Crippa, B., Malinverni, E. & Tucci, G. [1998], Complex Surface Representation. International Archives of Photogrammetry and Remote Sensing, vol. XXXII, part. 6W4, pp. 125-133.

Dallas, R. [1996], Architectural and archeological Photogrammetry. In Atkinson (Ed), 'Close Range Photogrammetry and Machine Vision', Whittles Publishing.

David, P., DeMenthon, D, Duraiswami, R. & Samet, H. [2003], Simultaneous pose and correspondence determination using line features. In Conference on Computer Vision and Pattern Recognition, volume 2, pages 424{431, Madison, WI, June 2003.

Debevec, E. [1996], Modelling and Rendering Architecture from Photographs. Ph.D. Thesis, University of California at Berkeley.

Debevec, P., Yu, Y. & Borshukov, G. [1998], "Efficient view-dependent image-based rendering with projective texture-mapping," in ACMEUROGRAPHICS Workshop, pp. 105-116.

DeMenthon, D. & Davis, L. [1992], Model-based object pose in 25 lines of code. In G. Sandini (Ed.), *Computer Vision—ECCV 92,Proceedings Second European Conference on Computer Vision*, Santa Margherita Ligure, Springer Verlag, pp. 335–343.

DeMenthon, D., and Davis, L. [1995], Model-based object pose in 25 lines of code, *Int. J. Comput. Vision* 15(1/2), 123–141.

De Micheli, E., Caprile, B., Ottanelo, P. & Torre, V. [1989], Localization and noise in edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(10).

Deveau, M., Pierrot-Deseilligny, M., Paparoditis, N. & Chen, X. [2004], Relative Laser Scanner And Image Pose Estimation From Points And Segments, ISPRS.

Dick, A.R., Torr, P., Ruffle, S. & Cipolla, R. [2001], Combining single view recognition and multiple view stereo for architectural scenes. Proc. 8th IEEE International Conference on Computer Vision (ICCV'01), pp. 268-274, July.

Djebali, M., Melkemi, M. & Sapidis, N. [2002], "Range-Image Segmentation and Model Reconstruction Based on a Fit-and-Merge Strategy," ACM Symposium on Solid Modeling and Applications, pp.127-138.

Dorai, C. [1997], COSMOS—A Representation Scheme for 3D Free-Form Objects *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol PAMI-19, no 10.

El-Hakim, S., Beraldin, A. & Blais, F. [1995], A comparative evaluation of the performance of passive and active 3-Dvision systems, in Proc.: St. Petersburg Great Lakes Conference on Digital Photogrammetry and Remote Sensing, St. Petersburg, Russia (in press, SPIE), June 25-30.

El-Hakim S. [2000], A Practical Approach to Creating Preciseand Detailed 3D Models from Single and Multiple Views. XIX Congress of the International Society for Photogrammety and Remote Sensing ISPRS, Proceedings on CD.

El-Hakim, S., Beraldin, A. & Picard, M. [2002], Detailed 3D reconstruction of monuments using multiple techniques. ISPRS/CIPA International Workshop on Scanning for cultural Heritage Recording, Corfu, Greece, pp.58-64.

El-Hakim, S., Gonzo, L., Picard, M., Girardi, S. & Simoni, A. [2003], Visualization of Frescoed Surfaces Buonconsiglio Castle-Aquila Tower, "Cycle Of The Months"Proceedings of International Workshop on Visualization and Animation of Reality-based 3D Models, Tarasp-Vulpera, Switzerland. February 24-28, NRC 44995.

El-Hakim, S., Fryer, J. & Picard, M. [2004], Modelling and visualization of aboriginal rock art in the Baiame cave. ISPRS International Archives of Photogrammetry and Remote Sensing, 35(5): 990-995.

Ernest, M. & Wu, S. [1992], Surface Parameterization and Curvature Measurement of Arbitrary 3D Objects: Five Practical Methods. IEEE Trans. PAMI. Vol.14. 833-840.

Fan, T., Medioni, G. & Nevatia, R. [1987], "Segmented description of 3-D surfaces," *IEEE Transactions on Robotics and Automation*, vol. 3, no.6, pp. 527-538.

Fangi, G., Fiori, F., Gagliardini, G. & Malinverni, S. [2001], Fast and accurate close range 3D modeling by laser scanning system - CIPA International Symposium, Proceedings – Potsdam.

Feibush, E. & Greenberg, D. [1980], Texture rendering system for architectural design. Comput. Aided Des. 12, 2 (March), 67–71.

Feldmar, J. & Ayache, N. [1996], Rigid, affine and locally affine registration of free-from surfaces. *International Journal of Computer Vision*, 18 (2), pp. 99-119. 1128-1135.

Ferencz, A. [2001], Texture Mapping Range Images. Masters Thesis, Computer Science Division, EECS, UC Berkeley.

Ferrie, F. & Levine, M. [1985], "Piecing together 3D shape of moving objects: An overview," in Proc. Computer Vision and Pattern Recognition Conf., IEEE Comput. Soc., San Francisco, CA, pp. 574-584.

Fischler, M. & Bolles, R. [1981] Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. Communications of the ACM, 24(6): 381–395.

Foley, J. [1990], Computer Graphics: Principles and Practice. Reading, Mass.: Addison-Wesley

Forsyth, D. & Ponce, J. [2003], Computer Vision: A Modern Approach. Prentice Hall.

Gardner, G. [1985], ''Visual Simulation of Clouds'', *Computer Graphics*, (SIGGRAPH '85 Proceedings), Vol. 19, No. 3, pp. 297-303.

Gächter, S. [2005], Results on range image segmentation for service robots. Technical Report EPFL-LSA-2005-01, Laboratoire de Syst`emes Autonomes, Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland.

Georgopoulos A., Ioannidis C., Daskalopoulos A. & Demiridi E. [2004], 3D reconstruction and rendering for a virtual visit. In: *XX ISPRS Congress,* Commission V, Istanbul, Turkey, IAPRSSIS Vol. XXXV, Part A5, pp. 632-636.

Gil, B., Mitiche, A. & Aggarwal, J. [1983], "Experiments in combining intensity and range edge maps," Computer Vision, Graphics, Image Processing, vol. 21, pp. 395-411.

Godin, G., Boulanger, P., 1995. Range Image Registration through Viewpoint Invariant Computation of Curvature. IAPRS, XXX (5W1), pp. 170-175.

Godin, G., Laurendeau, D. & Bergevin, R. [2001], A Method for the Registration of Attributed Range Images. In: Proc. of Third Int. Conference on 3D Digital Imaging and Modeling, Quebec, pp.179-186.

Gordon, S., Lichti, D. & Stewart, M. [2001], Application of a High-Resolution, Ground-Based Laser Scanner for Deformation Measurements. Proceedings of 10th International FIG Symposium on Deformation Measurements, Orange, California, 19-22 March, pp. 23-32.

Grammatikopoulos, L., Kalisperakis, I., Karras, G., Kokkinos, T. & Petsa, E. [2004], Automatic Multi-Image Photo-texturing of 3D Surface Models Obtained With laser Scanning. CIPA International Workshop on "Vision Techniques Applied to the Rehabilitation of City Centres", Lisbon, 25-27 October.

Greene, N. & Heckbert, P. [1986], ''Creating Raster Omnimax Images from Multiple Perspective Views Using The Elliptical Weighted Average Filter'', *IEEE Computer Graphics and Applications*, Vol. 6, No. 6, pp. 21-27.

Greene, N. & Kass, M. [1993]. Hierarchical Z-buffer visibility. In *Computer Graphics Proceedings, Annual Conference Series*, pages 231–240.

Gruen, A., Remondino, F. & Zhang, L. [2002], Reconstruction of great Buddha of bamiyan, Afghanistan, *International Archives of Photogrammetry and Remote Sensing* , Vol.XXXXIV, part 5, Corfu, Greece, pp.363-368.

Gruen, A., Remondino, F. & Zhang, L. [2004], 3D modeling and visualization of large cultural heritage sites at very high resolution: the Bamiyan valley and its standing Buddhas. IAPRS, Vol. XXXV-B5.

Gruen, A. & Akca, D. [2004], Least Squares 3D Surface Matching "1.PanoramicPhotogrammetry Workshop", Dresden, Germany

Haala, N. & Kada, M. [2005], Panoramic Scenes for Texture Mapping of 3D City Models. Proceedings of the ISPRS working group V/5 'Panoramic Photogrammetry Workshop', IAPRS Vol. XXXVI-5/W8, on CD.

Habib, A., Ghanma, M., Morgan, M. & Al-Ruzouq, R. [2005], Photogrammetric and LIDAR data registration using linear features. *Journal of Photogrammetric Engineering and Remote Sensing*, June issue .

Han, F., Tu, Z. & Zhu, S. [2004], Range image segmentation by an efficient jump-diffusion method. *In IEEE Trans. On Pattern Analysis and Machine Intelligence.*Vol.26, No.9. pp.1138-1153.

Häusler, G. & Karbacher, S. [1997] Reconstruction of smoothed polyhedral surfaces from multiple range images. In: Girod B, Niemann H, Seidel H-P (eds) Proceedings of 3D Image Analysis and Synthesis '97, Infix, Sankt Augustin, Germany, pp 191–198.

Haralick, R., Lee, C., Ottenberg, K. & Nolle, M. [1994], Review and analysis of solutions of the 3-point perspective pose estimation problem. IJCV 13(3), pp. 331–356.

Hartley R. & Zisserman, A. [2000], Multiple View Geometry, Cambridge University Press.

Hatger, C. & Brenner, C. [2003], Extraction of road geometry parameters form laser scanning and existing databases, *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXIV, Part 3/W13.

Hebert, M. & Ponce, J. [1982], A new method for segmenting 3-D scenes into primitives. In Proc. 6th Int. Conf. Pattern Recognition, Munich, West Germany, Oct. 19-22, pp. 836-838.

Heckbert, P. [1986], Survey of texture mapping. *IEEE Computer Graphics and Applications*, 6(11), 56-67.

Heckbert, P. [1983], ‘‘Texture Mapping Polygons in Perspective’’, Technical Memo No. 13, NYIT Computer Graphics Lab.

Heckbert, P. & Moreton, H. [1991], Interpolation for polygon texture mapping and shading, In: State of the Art in Computer Graphics: Visualization and Modeling, pp. 101-111. D.F. Rogers and R.A. Earnshaw Eds. Springer-Verlag, New York, NY, U.S.A.

Henderson, T. [1983], “Efficient 3-D object representations for industrial vision systems,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, no. 6, pp. 609-617.

Herman, M. [1985], “Generating detailed scene descriptions from range images,” in Proc. Int. Conf. Robotics and Automation, St. Louis, MO, Mar. 25-28, pp. 426-431.

Heuel, S. & Förstner, W. [2001], Matching, Reconstructing and Grouping 3D Lines from Multiple Views Using Uncertain Projective Geometry, *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Hawaii, pp. 1-8.

Hoffman, R. & Jain, A. [1987], Segmentation and Classification of Range Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 9(5): pp.608-620.

Hoover, A., Jean-Baptiste, G. & Jiang, X. [1996], An Experimental Comparison of Range Image Segmentation Algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(7): p. 673-689.

Hoppe, H., DeRose, T., Duchamp, T., McDonald, J. & Stuetzle, W. [1992], Surface reconstruction from unorganized points. In: Proceedings of SIGGRAPH ’92, Chicago. Comput Graph 26(2): 71–78.

Horaud, R., Christy, S. & Dornaika, F. [1994], Object Pose: The Link between Weak Perspective, Para Perspective and Full Perspective, Techical Report RR-2356, INRIA.

Horaud, R., Dornaika, F., Lamiroy, B. & Christy, S. [1997], “Object pose: The link between weak perspective, paraperspective, and full perspective,” Int. J. Comput. Vision, vol. 22, no. 2, pp. 173–189.

Horn, B. [1987], “Closed Form Solution of Absolute Orientation using Unit Quaternions,’’ Journal of the Optical Society *A*, Vol. 4, No. 4, pp. 629–642, April.

Hourcade, J. & Nicolas, A. [1985] Algorithms for Andaliased Cast Shadows, *Computers & Graphics* 9, 3, 259-265.

Hrabacek, J. &van den Heuvel, F. [2000]. Weighted Geometric Objects Constraints Integrated in a Line-Photogrammetric Bundle Adjustment, *International Archives of Photogrammetry and Remote Sensing*, Amsterdam, Vol. XXXIII, pp. 380-387.

Huber, D. & Hebert, M. [2001], Fully automatic registration of multiple 3D data sets. In IEEE Comp. Soc. Workshop on Computer Vision. Beyond the Visible Spectrum.

Hynst S., Gervautz M., Grabner M. & Schindler, K. [2001], workflow and data model for reconstruction, management, and visualization of archaeological sites. ACM Siggraph Symposium on Virtual Reality, Archaeology and Cultural Heritage 2001 (VAST2001).

Jiang, X. & Bunke, H. [1999], Edge detection in range images based on scan line approximation. Computer Vision and Image Understanding: CVIU, 73(2): 183–199.

Kada, M., Roettger, S., Weiss, K., Ertl, T. & Fritsch, D. [2003], Real-time visualisation of urban landscapes using open-source software. Proceedings of ACRS 2003 ISRS, Busan, Korea.

Katsoulas, D. & Werber, A. [2004], Edge Detection in Range Images of Piled Box-like Objects. 17th International Conference on Pattern Recognition (ICPR 2004), 4-Volume Set, Cambridge, UK. IEEE Computer Society, ISBN 0-7695-2128-2.pp 80-84.

Kim J. & Pollefeys, M. [2004], Radiometric Self-Alignment of Image Sequences. Proceedings of CVPR, pp. 645-651

Klinec, D. & Fritsch, D. [2003], Towards pedestrian navigation and orientation. Proceedings of the 7th South East Asian Survey Congress: SEASC'03, Hong Kong, November 3-7.

Kunmar, R. & Hanson, A. [1994], Robust methods for estimating pose and a sensivity analysis. *Computer Vision Graphics and Image Processing*, 60(3), pp 313-342.

Kurazume, R., Nishino, K., Zhang, Z. & Ikeuchi, K. [2002], Simultaneous 2D image and 3D geometric model registration for texture mapping utilizing reflectance attributes, in The 5th Asian Conference on Computer Vision, pp. 454–461.

Langridge, D. [1984], Detection of discontinuities in the first derivatives of surfaces Computer Vision, Graphics Image Processing 27. 291–308.

Lavallee, S. & Szeliski, R. [1995], Recovering the Position and Orientation of Free-Form Objects from Image Contours Using 3D Distance Maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 378-390.

Lee, K., Meer, P. & Park, R. [1998], Robust Adaptive Segmentation of Range Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(2): p. 200-205.

Lee, Y. & Habib, A. [2002], Pose Estimation of Line Cameras Using Linear Features, ISPRS Symposium of PCV'02 Photogrammetric Computer Vision, Graz, Austria.

Lensch, H., Heidrich, W. & Seidel, H. [2000], Automated texture registration and stitching for real world models. In Proceedings of Pacific Graphics 2000, pages 317–326.

Leonardis, A., Gupta, A. & Bajcsy, R [1995], Segmentation of Range Images as the search for Geomatric Parametric Models. *International Journal of Computer Vision*, 14(3):253–277.

Leveau, S. & Faugeras, O. [1994], 3-d scene representation as as collection of images and fundamental matrix. Technical Report 2205, INRIA Sophia-Antipolis.

Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J. & Fulk, D. [2000], The Digital Michelangelo Project: 3D scanning of large statues. In SIGGRAPH 2000, *Computer Graphics Proceedings*, Annual Conference Series, pages 131–144. Addison Wesley.

Liu, L. & Stamos, S. [2005], Automatic 3D to 2D registration for the photorealistic rendering of urban scenes. Computer Vision and Pattern Recognition. CVPR 2005. *IEEE Computer Society*, volume 2.pp 137-143.

Liu, Y., Huang, T. & Faugeras, O. [1990], Determination of camera location from 2-d to 3-d line and point correspondences, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1).

Lu, Y. & Jain, R. [1992], Reasoning About Edges in Scale Space, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 4,pp. 450-468.

Luebke, D., Reddy, M., Cohen, J., Varshney, A., Watson, B. & Huebner, R. [2003], Level of Detail for 3D Graphics. Morgan-Kaufmann, Inc.

Marshall, D., Lukacs, G. & Martin, R. [2001], Robust segmentation of primitives from range data in the presence of geometric degeneracy; IEEE Trans. Pattern Analysis and Machine Intelligence 23(3), 304-314.

Masuda, T. & Yokoya, N. [1995], A Robust Method for Registration and Segmentation of Multiple Range Images. Computer Vision and Image Understanding, 61(3), pp. 295-307.

McMillan, L. & Bishop, G. [1995], Plenoptic modeling: An image-based rendering system. In Proc. SIGRAPH '95, pages 39-46.

Milgrim, D. & Bjorklund, C. [1980], "Range image processing: Planar surface extraction," in Proc. 5th Int. Conf. Pattern Recognition, Miami, FL, Dec. pp. 912-919.

Mitiche, A. & Aggarwal, J. [1983] "Detection of edges using range information," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 5, pp.174-178.

Mueller, K., Moeller, T., Swan, J., Crawfis, R., Shareef, N. & Yagel, R. [1998], Splatting Errors and Antialiasing. *IEEE Transactions on Visualization and Computer Graphics*. 4(2): 178-191.

Nehlig, P. & Ghazanfarpour, D. [1992], Affine Texture Mapping and Antialiasing Using Integer Arithmetic.computer Graphic Forum, volume 11 (3): pp 227-236.

Nevatia, R. & Binford, T. [1973], "Structured descriptions of complex objects," in Proc. 3rd Int.Joint Conf. Artificial Intelligence, Stanford, CA, pp. 641-647.

Niem, W. [1994], "Robust and Fast Modelling of 3D Natural Objects from Multiple Views", SPIE Proceedings "Image and Video Processing II", Vol. 2182, pp. 388-397.

Niem, W. & Broszio, H. [1995], Mapping texture from multiple camera views onto 3d-object models for computer animation. In International Workshop on Stereoscopic and ThreeDimensional Imaging Conf. Proc., Santorini, Greece.

Neugebauer, P. & Klein, K. [1999], Texturing 3D Models of Real. World Objects from Multiple Unregistered Photographic Views Proc. Eurographics'99. Volume 18.

Nguyen, V., Martinelli, A., Tomatis, N. & Siegwart, R. [2005], A comparison of line extraction algorithms using 2D laser rangefinder for indoormobile robotics. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems.

Nishihara, H. [1984], Practical real-time imaging stereo matcher. Optical Engineering, 23(5): 536– 545.

Ofek, E., Shilat, E., Rappoport, A. & Werman, M. [1997], Multiresolution textures from image sequences. *IEEE Computer Graphics and Applications*, pages 18-29.

Ortin, D. & Remondino, F. [2005],International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, part 5/W17 (CD-Rom). International Workshop of ISPRS WG V/4 '3D-ARCH 2005: Virtual Reconstruction and Visualization of Complex Architectures', 22-24, Mestre-Venice, Italy.

Paeth, A. [1986], ‘‘A Fast Algorithm for General Raster Rotation’’, *Graphics Interface ’86*, pp. 77-81.

Palmer, P., Dabis, H. & Kittler, J. [1996] A performance measure for boundary detection algorithms. CVGIP: Image Understanding, 63(3): 476-494.

Pollefeys, M., Koch, R., Vergauwen, M. & Van Gool, L. [2000], *Automated reconstruction of 3D scenes from sequences of images*, ISPRS Journal Of Photogrammetry and Remote Sensing (55) 4, pp. 251-267.

Popplestone, R., Brown, C., Ambler, A. & Crawford, G. [1975], “Forming models of plane-and-cylinder faceted bodies from light stripes,” in Proc. 4th Int. Joint Conference Artificial Intelligence, Tbilisi, Georgia, USSR, pp. 664-668.

Pulli, K., Abi-Rached, H., Duchamp, T., Shapiro, L. & Stuetzle, W. [1998], Acquisition and visualization of colored 3-D objects, in Inter. Conf. on Pattern Recognition, Australia, pp. 11-15.

Ratishauser, M., Stricker, M. & Trobina, M. [1994], Merging range images of arbitrarily shaped objects. In: Proceedings of the IEEE conference on computer vision and pattern recognition, Seattle, pp 573–580.

Remondino, F. & Niederoest, J. [2004], Generation of High-Resolution Mosaic for Photo-Realistic Texture-Mapping of Cultural Heritage 3D Models. proceedings of the 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST). K. Cain, Y. Chrysanthou, F. Niccolucci, N. Silberman (Eds), pp. 85-92.

Rocchini, C., Cignoni, P., Montani, C., Pingi, P. & Scopigno, R. [2001], A Suite of Tools for the Management of 3D Scanned Data., Proc. Workshop on 3D Digital Imaging and Modeling Applications of: Heritage, Industry, Medicine & Land.

Rocchini, C., Cignoni, P., Ganovelli, F., Montani, C., Pingi, P. & Scopigno, R. [2004], The marching intersections algorithm for merging range images. The Visual Computer, 20:149-164.

Romney, W. [1970], "Computer Assisted Assembly and Rendering of Solids," Department of Computer Science, University of Utah, TR.4.20.

Rothwell, C., Mundy, J., Hoffman, W., & Nguyen,V. [1995], "Driving Vision by Topology." Proceedings IEEE Symposium on Computer Vision, pp. 395-400.

Rusinkiewicz, S., Hall-Holt, O. & Levoy, M. [2002], Real-time 3D model acquisition. In SIGGRAPH 2002.

Sabata, B., Arman, F. & Aggarwal, J. [1990], Segmentation of range images using pyramidal data structures. In Proceedings of the ICCV, pages 662-666.

Sapidis, S. & Besl, P. [1995], "Direct Construction of Polynomial Surfaces from Dense Range Images through Region Growing," ACM Transactions on Graphics, 14, pp. 171-200.

Sappa, A. & Devy, M. [2001], “Fast Range Image Segmentation by an Edge Detection Strategy,” Proc. IEEE Conf. 3D Digital Imaging and Modeling, pp. 292-299.

Sappa, A., Restrepo-Specht, A. & Devy, M. [2001], Range image registration by using an edge-based representation, in: Proceedings of the Ninth International Symposium on Intelligent Robotic Systems (SIRS’01), Toulouse, France.

Sedlaczek, B. [2000], Petra, Art and Legend. MP Graphic Formula, Roma, Italy, pp.4-20.

Sequeira, V., Ng, K., Wolfart, E., GonÁalves, J. & Hogg, D. [1999],"Automated Reconstruction of 3D Models from Real Environments", ISPRS Journal of Photogrammetry and Remote Sensing (Elsevier), vol. 54, pp. 1-22.

Sequeira, V. & Gonçalves, J. [2002], Reality Modelling: Photo-Realistic 3D Models of Real World Scenes, Proc 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT 2002), Padova, Italy, pp 776-783, June 19-21.

Sgrenzaroli, M. & Wolfart, E. [2002], Accurate texture-mapped 3D models for documentation, surveying and presentation purposes, CIPA-Corfu 2002.

Sethi, K. & Jayaramamurthy, S. [1984], “Surface classification using characteristic contours,” in Proc. 7th Int. Con5 Pattern Recognition, Montreal, P.Q., Canada, pp. 438-440.

Skarbek, W. & Koschan, A. [1994] , Colour image segmentation – a survey –. Tech. rep., Institute for Technical Informatics, Technical University of Berlin, October.

Soucy, M. & Laurendeau, D. [1992], ‘Surface modeling from dynamic integration of multiple range views,’’ Proc. ofthe l l t h Intl. Conj on Pattern Recognition, pp. 449-452.

Soucy,M. & Laurendeau, D. [1995], A General Surface Approach to the Integration of a Set of Range Views, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 344-358, Apr. 1995.

Stamos, I. & Allen, P. [2003], 3-D model construction using range and image data. In Cornpurer Vision & Pattern Recognition ConJ. (CVPR), pages 531-536.

Stamos, I. & Leordeanu, M. [2003], Automated feature-based range registration of urban scenes of large scale. In International Conference of Computer Vision and Pattern Recognition, volume 2, pages 555–561.

Stamos, I. & Leordeanu, M. [2004], Efficient Model Creation of Large Structures based on Range Segmentation. 2nd International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT 2004), Thessaloniki, Greece. IEEE Computer Society 2004,pp 447-454.

Sutherland, I., Sproull, R. & Schumacker, R. [1974], A Characterization of Ten Hidden-Surface Algorithms”, Computing Surveys, Vol 6, No 1.

Sze, C., Liao, H., Hung, H., Fan, K. & Hsieh, J. [1998], “Multiscale edge detection on range images via normal changes,” *IEEE Transition. Circuits System*. II, vol. 45, pp. 1087–1092.

Szeliski, R., Tonnesen, D. & Terzopoulos, D. [1993], “Modeling surfaces of arbitrary topology with dynamic particles,” Proc. of IEEE Conj. onCVPR, pp. 82-87.

Tanimoto, S. & Pavlidis, T. [1975], ‘‘A Hierarchical Data Structure for Picture Processing’’, *Computer Graphics and Image Processing*, Vol. 4, No. 2, pp. 104-119.

Tomita, F. & Kanade, T. [1984], “A 3D vision system: Generating and matching shape descriptions in range images,” in Proc. Int. Conf. Roborics, IEEE Comput. Soc., Atlanta, GA,pp.186- 191.

Tsai, R. [1987], A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the shelf TV cameras and lenses. *IEEE J. Robotics and Automation*, 3(4): 323-344.

Turk, G. & Levoy, M. [1994], Zippered polygon meshes from range images, in: Proceedings of the ACM SIGGRAPH’94, Orlando, FL, pp. 311-318.

Ulama, M. [1996], Jerash , A unique Example of A Roman City, publishers Al-Ulmam, third edition.

Vanco, M. & Brunnett**,** B. [2002], Direct Segmentation for Reverse Engineering, accepted for: *Cyberworlds 2002*, Tokio, Japan.

Vanden Wyngaerd, J., Van Gool, L., Koch, R. & Proesmans, M. [1999], Invariant-based registration of surface patches. IEEE Int.Conf. on Computer Vision, Kerkyra, Greece, September 20-27,pp. 301-306.

Vemuri, B., Mitiche, A. & Aggarwal, J. [1986], “Curvature-based representation of objects from range data,” Image and Vision Computer. Vol. 4, no. 2, pp. 107-114.

Visnovcova, J., Gruen, A. & Zhang, L. [2001], Generating a 3D Model of a Bayon Tower Using Non-metric Imagery, Proc. Intern. Workshop Recreating the Past– Visualization and Animation of Cultural Heritage, pp. 30-39, Ayutthaya, Thailand.

Wang, Z. & Jepson, A. [1994], A new closed-form solution for absolute orientation. In: CVPR94, pp. 129–134.

Wang, X., Totaro, S., Taillandier, F., Hanson, A. & Teller. S. [2003], Recovering façade texture and microstructure from real-world images. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. (34) 3A, pp. 381–386.

Wani, M. & Batchelor, B. [1994], Edge Region Based Segmentation of Range images. IEEE Transactions on Pattern Analysis and Machine Intelligence.16(3): p. 314-319.

Watson, F. [1981], "Computing the N-Dimensional Delaunay Tessellation with Apllication to Voronoi Polytopes" The Computer Journal (24).

Weinhaus, M. & Devarjan V. [1997], “Texture mapping 3D models of real-world scenes”, ACM Computer Survey, 29(4), pp. 325-365.

Werner, T. & Zisserman, A. [2002], New technique for automated architectural reconstruction from photographs. Proc. 7th European Conf. Computer Vision, Copenhagen, May.

Williams, L. [1983], ‘‘Pyramidal Parametrics’’, *Computer Graphics*, (SIGGRAPH ’83 Proceedings), Vol. 17, No. 3, pp. 1-11.

Woo, A., Poulin, P. & Fournier, A. [1990], A survey of shadow algorithms. IEEE Computer Graphics & Applications 10, 6 (November), 13–32.

Wünstel, M. & Röfer, T. [2004], Feature Based Registration of Range Images in Domestic Environments In: International Conference on Computer Vision and Graphics 2004 (ICCVG 2004), Computational Imaging and Vision, Kluwer.

Yokoya, N. & Levine, M. [1989], "Range image segmentation based on differential geometry: A hybrid approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*,vol. 11, pp. 643–649.

Yu, Y. [1999], Efficient visibility processing for projective texture mapping. Computer and Graphics , Volume 23 (2): 245-253.

Yu,Y. [2000] Andras Ferencz, Jitendra Malik. Compressing Texture Maps for Large Real Environments. SIGRAPH'2000.

Zhang, H., Manocha, D., Hudson, T. & Hoff, K. [1997]. Visibility culling using hierarchical occlusion maps. Proc. of ACM SIGGRAPH.

Zhang, Y. [1997], Evaluation and Comparison of Different Segmentation Algorithms, Pattern Recognition Letters, vol. 18, No. 10, pp. 963-974.

Zhang, Z. [1994], Iterative Point Matching for Registration of Free-Form Curves and Surfaces, International Journal. Computer Vision, vol. 13, no. 2, pp. 119-152.

Zhao, H. & Shibasaki, R. [2003], Reconstructing a textured CAD model of an urban environment using vehicle-borne laser range scanners and line cameras. *MVA*, 14(1): 35–41.

Zitova, B. & Flusser, J. [2003], Image registration methods: a survey," Image and Vision Computing 21(11), pp. 977-100

Curriculum Vitae

Name	Yahya Alshawabkeh
Feb. 23, 1975	Born in Amman, Jordan
1980-1986	Primary School, Amman, Jordan
1986-1990	Secondary School, Amman, Jordan
1990-1993	High School, Amman, Jordan
1993	The General Secondary Education Examination
1993-1998	B.Sc. in Civil Engineering, Jordan University of Science and Technology
1998-2001	M.Sc. in Civil Engineering, Jordan University of Science and Technology
	Teaching Assistant in Civil Enginering Department
2001-2003	Working at Hashemite University of Jordan as a Surveying lab. Super-visor
Since March 2003	Institute of Photogrammetry, University of Stuttgart, Stuttgart, Germany
	(Terrestrial Positioning System and Computer Vision Group)