Inferring Surface Structure of Rock Piles from Range Images

William K. W. Cheung

B. Eng., (University of Sheffield, UK), 1989

Department of Mining and Metallurgical Engineering

McGill University

Montréal

June, 1992

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of
Master of Engineering

Abstract

This thesis deals with the problem of applying computer vision techniques in an underground mine environment. In particular, the problem of the localization and identification of oversized rock fragments prior to secondary breakage, following the initial drilling and blasting operation is addressed. The strategy employed is based on the methodologies developed for reconstruction and interpretation of range image data. Discrete rock pile images acquired using the NRCC/McGill laser rangefinder were used in this study. The main contribution of this thesis is the complete study of the paradigm which involves: range data acquisition, surface reconstruction, segmentation, and fitting of parametric shape models The final representation obtained from the model, describing the spatial and geometric properties of each rock fragment and can be used to control an automated rock-breaking mechanism. To support the strategy developed, a number of experimental results at different processing stages are presented.

Résumé

Cette thèse s'attelle au problème de l'application des techniques de visionique en milieu minier souterrain. En particulier, nous attaquons le problème de la localisation et de l'identification de fragments rocheux surdimensionnés, à la suite des opérations initiales de forage et d'expicsion. La stratégie employée est fondée sur les méthodes developpées pour la reconstruction et l'interprétation d'images télémétriques. Des images d'empilement de roches, obtenues à l'aide du télémètre à laser NRCC/McGill ont été utilisées pour cette étude. La contribution principale de cette thèse est l'étude complète du paradigme qui comprend: l'acquisition de données télémétriques, la reconstruction de surface, la segmentation, et le choix et l'ajustement de modèles de forme paramétriques. La représentation finale obtenue à partir du modèle décrit les propriétés spatiales et géometriques de chaque fragment rocheux et peut être utilisée pour contrôler un mécanisme automatisé de fragmentation des roches. Afin de valider notre approche, un certain nombre de résultats experimentaux des différentes étapes sont présentés dans la thèse.

Acknowledgment

First and foremost, I wish to acknowledge the support of all my supervisors, both official and unofficial. Let me begin by expressing my gratitude to John Edwards for initiating the automated rockbreaker project; without him my research work at McGill would not have been possible. I am equally indebted to Gregory Carayannis with whom I shared in many stimulating discussions during the course of my study, to Frank Ferrie for introducing me to the field of computer vision and making me realize how little I knew (and how little I know) about the subject, and last but not least to Rousso Dimitrakopoulas for making everything run smoothly in general.

Thanks are also due to Gilbert Soucy and Peter Waite for their programming support and valuable comments, to Wassim Alami, Tony Botzas, Frank DiGiuseppe and Bobby Lazar for taking the time to read and correct my English in parts of the early versions of my thesis, to Stéphane Aubry for translating the abstract into French, and to the people of the McGill Research Centre for Intelligent Machines (McRCIM) who maintained and supported an ideal research environment. Furthermore, I wish to thank McRCIM for generously providing me with its computing, printing, and laboratory facilities.

Finally, I express my warmest thanks to my beloved parents and family for the care and support that they offered me over the long period of my study both financially and morally.

Table of Contents

List of Figures

List of Tables

Index of Symbols

Chapter 1 Introduction

Machine vision has been successfully applied to industrial problems, assisting or even replacing human operators in tasks involving visual perception. Examples of such successful applications include, inspection of printed circuit boards [Hara *et al.*, 1982, Mandeville, 1985], finger-print recognition [Hrechak and McHugh, 1990], letter sorting [Mitchell and Gillies, 1989], and automatic welding [Beranek *et al.*, 1986]. However, very little has been done in transferring the existing computer vision technology to hostile environments like the ones encountered in mining. Recent advances in computer architecture, improved software reliability, and the availability of sophisticated image acquisition devices have opened up a new frontier for such novel computer vision applications. Clearly, mining tasks that require human supervision, and that are dangerous and/or tedious, are worthwhile candidates for computer vision-based automation.

This thesis deals with such an application of computer vision in mining and in particular to the problem of rock fragment identification and localization for secondary rockbreaking operations. Currently, a human operator determines the position and the geometry of each rock to be broken, and then positions and controls the breaking tool accordingly. In an automated system, spatial information and the geometry of the scene must be acquired at high rates. The "traditional" contact sensing methods, such as tactile sensing have failed in this respect, while the non-contact sensing methods, including the computer vision-based approach, provide a natural and viable framework for such a system.

The basic approach used here is quite different from those proposed in the past for use in mining, and is based on the use of a laser rangefinder rather than a standard T.V. camera, to reduce the complexity of image interpretation. Range images have a number of advantages over intensity images for inferring the 3-D structure of objects in a scene. Ideally, range information is not subjected to the changes in lighting conditions, surface reflectance and camera position. This is achieved by making the scene geometry explicit. Nevertheless, the "low-level" vision problems, such as feature detection and segmentation remain [Jain and Jain, 1990].

The strategy is based on the methodologies developed for 3-D object recognition by [Pentland, 1987, Ferrie *et al.*, 1989]. It involves the reconstruction and interpretation of range image data; stable surface properties estimation, feature recovery, part decomposition and solid modelling of three-dimensional objects. Experience gained from previous work in solid shape modelling of man-made objects, supports the application of these ideas in the rockbreaking problem [Ferrie *et al.*, 1990]. The final representation of the mine scene obtained from the modelling will be utilized to control the rockbreaking mechanism as already mentioned.

1.1 Mining automation

The mining industry is currently undergoing an extensive technological revolution. A considerable amount of research and development has recently been carried out on automating the mechanization and instrumentation of mining equipment. On the other hand, the manufacturing industry has long been experiencing the prosperity and profitability which fully or semi-automated machinery can generate. Mining industries in developed countries that are not blessed with abundant resources and cheap labour, must automate their mining operations in order to remain competitive.

One can easily identify the four main objectives of *mining automation*: (i) reduce production costs, (ii) increase productivity, (iii) enhance the working environment, and (iv) improve the safety of the workers. In order to achieve all these objectives, automation in mining requires *intelligent machines*, capable of

- carrying out actions,
- perceiving and understanding the surrounding environment,
- making intelligent decisions.

Furthermore, their behavior should adapt to changes in the environment and be based on *a priori* knowledge. Such *a priori* knowledge may include, the information about the surrounding environment, the characteristics of the machinery, and their established operating procedures.

In one of the early papers on mining automation, Salamon reviewed a general framework for mine automatic control systems, and correctly predicted the slow progress in the introduction of automation into the mining industry [Salamon, 1976]. One explanation [Kassler, 1985], is that the direct transfer of technology developed for other industries into mining has proven ineffective. A number of factors inherent in the mining machinery and environment, which may affect the automation process, should also be seriously considered during the early stages of such projects if any degree of success is to be achieved.

In general, there are two classes of machinery that can be automated, viz: (i) stationary, and (ii) mobile. A machine is considered to be *stationary*, if it is fixed to a designated location, and its work-space is well defined and restricted. A machine is said to be *mobile*, if it is capable of moving or changing its entire position without being attached to any specified location. Thereby, the workspace of a mobile machine may change according to its new location/position. The problem of automating a stationary machine is relatively straightforward when compared to that of a mobile machine. The additional mobility complicates the automation process.

Most of the robots employed in the manufacturing industry fall in the first category, with their bases fixed to the floor, performing routine "pick and place" operations. Unlike the structured robot environments that are typical in the manufacturing industry, mine environments continuously evolve over the entire production cycle, in a manner highly dependent upon the geological irregularities inherent in the randomness of nature.

Typical tasks that the mining machinery has to perform, include rock-attack, secondary breakage, ore handling and transportation. A large number of mining machines are therefore essentially mobile. For the purposes of automation, they are closely related to autonomous vehicles rather than industrial robots. The required payloads and control strategies are quite different for manufacturing automation and mining automation. Additionally, mining machinery is subjected to various abuses from their surrounding environment (e.g., dust, heat, humidity and vibration). Thus, it makes the design and development of mining automation particularly challenging.

1.1.1 Machine vision in mining

The following describes some of the mining computer vision applications discussed in the literature. Most of these applications share a common theme; the use of standard T.V. cameras as sensing devices, and registering the direct illuminance emitted from the scene. In interpreting these images, *a priori* knowledge about the characteristics of each feature to be detected is often required.

Ore deposit evaluation

The distribution of lithologic facies from an exploitation wall can be considered as an indicator for the ore deposit structure evaluation. It has been shown that, with the help of computer vision, it is possible to make a rough estimation of the mineral content in the ores, by analysing images of the exploitation wall [Bonifazi and Massacci, 1989]. Given the relationship between the chemical characteristics and the colour (spectrum) information of each ore deposit, the litho-type was determined from the differences in the grey level distribution of the images. If the exploitation is advanced according to an alignment whose coordinates are known, and the images are taken from the walls perpendicular to the feeding direction, then the orebody volumes can be computed by means of a simple geometrical calculation. This information can also be used for the geostatistical evaluation of the mine production.

Rock slope analysis

In geological surveys, one of the most common problems is the determination of the joint parameters that can be used to predict the behaviour (stability) of the rock mass. Often, direct measurements of the slope are very difficult to perform; e.g., where the joints' outcropping on rock walls is high and steep. Photogrammetric techniques have been developed as a tool for rock slope characterization and monitoring in difficult ground conditions [Baratin *et al.*, 1990].

The proposed method is based on reconstructing the three-dimensional terrain model of the rock slope. At least two stereoscopic pictures of the slope are required for the reconstruction of the digital terrain model. The latter is used to derive the geometrical parameters of the joints. Good agreement between both the calculated and the direct measured values has been reported [Baratin *et al.*, 1990].

Ore sorting

The separation of useful minerals from gangue is a classical problem that appears in every mining exploitation. This sorting process is often carried out manually depending on visual information. Vision systems have been proposed and developed in the past to help maintain the mineral content above the minimal level required for high quality industrial parts [Mäenpää *et al.*, 1983, Manana *et al.*, 1985]. These systems proceeded by identifying different rock and ore types on a moving conveyor belt based on their surface reflectance spectrums.

Experience has shown that the mineral quality achieved by the vision-based sorting system failed to surpass the quality achieved by skillful workers [Mäenpää *et al.*, 1983]. Problems occur when the rock surface is covered with dirt or the rock consists of about one-half of waste. Washing may well be required before the sorting process. A problem, however, is that washing cannot be employed in sub-zero temperatures (e.g , during the winter period).

Measurement of blast fragmentation

The knowledge of fragment size distribution is very crucial to the success of a mine production. For maximum productivity and efficiency, it is important to optimize the fragment size distribution throughout the entire production cycle; from the initial blasting to smelting. The subject of fragment size estimation via image analysis techniques, has attracted a considerable amount of interest from researchers all over the world.

Most of the proposed techniques are *edge-based*; i.e., they proceed by detecting features corresponding to discontinuities in intensity changes. Studies have shown that the size estimation can be derived, by using the edge information and the assumption that blasted fragments are spherical in shape [Carlsson and Nyberg, 1983, Hunter *et al.*, 1990]. Others include more sophisticated algorithms such as overlapping correction functions derived empirically from experiments [Maerz *et al.*, 1987] and stereo techniques to compute more reliable size distribution estimates using 3-D information [Cheung and Ord, 1990].

Attempts have also been made in the past to model rock fragments with convex polygons [Gao and Wong, 1989]. An intensity image of a rock pile, obtained from a standard T.V. camera, was pre-segmented by thresholding. The resulting binary image was then treated as a rough segmentation of the muckpile which might contain a lot of broken (weak) boundaries. The image was then passed to a second stage of segmentation in an attempt to reconstruct the missing boundaries of the rock lumps. The boundary reconstruction was based on approximating each 2-D rock profile with a convex polygon. The final segmented image shared some resemblances to that obtained by manual segmentation. Although this was judged to be adequate for computing the size and/or volume distribution [Gao and Wong, 1989], its lack of precision in localizing the "true" rock boundaries, makes it insufficient for tasks such as rock fragment localization and identification.

Roadheader automation

The outline of the orebody boundary and its geological features are essential in selecting the cutting trajectories for a roadheader machine during the excavation process [Orteu and Devy, 1991, Fuentes-Cantillana *et al.*, 1991]. This is due to the fact that the cutting sequence must change according to the geometrical distribution of the mineral deposits, especially when these deposits are highly irregular. Orteu and Devy proposed the use of a computer vision-based system to discriminate the ore types present on the cutting face [Orteu and Devy, 1991].

In the study, two colour cameras were mounted on the roadheader at different locations, in such a way that the covered areas are complementary to each other. Knowing the geometrical setup of the instruments and the roadheader, the two images were combined together to form a single image of the complete cutting face. The resulting image was then separated into different regions that corresponded to the colour spectrum of each ore type. From this segmented and labelled image, a so-called "face map" of the ore distribution is produced, which can be used to control the cutting boom of the roadheader machine.

Guidance of LHD vehicles

Load-Haul-Dump (LHD) vehicles are in widespread use for the loading and transport operations in underground mine environments. Considerable emphasis has been put both on better path control and on more efficient usage of these vehicles. A monitoring system has been developed to collect data from various sensors mounted on LHD vehicles, and this information is used to assist the mine staff in problem diagnosis and maintenance service [Baiden, 1988].

More recently, the development of automatic guided LHD vehicles based on machine vision has been reported. A prototype of such a system [St-Amant *et al*, 1991, Hurteau *et al.*, 1991] has been built and was tested in an underground mine. The results demonstrated the feasibility of employing automatic guided LHD vehicles in an underground mine, equipped with optical lines fitted on the ceiling just above the vehicle's guide-path. The basic idea is very simple; the prototype vehicle is equipped with two specially designed optical line detectors for tracking the vehicle movement with respect to the guide-path· one for the forward direction and one for the reverse direction A similar practice in Sweden is also reported [Vagenas *et al.*, 1991], however, a white-painted line was used as the optical guide-path rather than a retro-reflective ribbon.

Secondary rock breakage

It is very difficult, if not impossible, to achieve perfect blasting results, due to both technical and economic reasons. The drilling of straight blast holes that remain parallel over the long distances demanded by mining economists is difficult to achieve [Chabot *et al.*, 1989]. After blasting therefore, the resulting fragment sizes are often badly distributed, making the processing and transporting of minerals very difficult. In cases where the rock fragments are too large, a *secondary breakage* is usually required to prevent oversized material from being transported. It is very common to find that in a mine machine, breakdown results from the frequent loading of oversized fragments.

Initial studies demonstrated the potential of machine vision for the automation of the

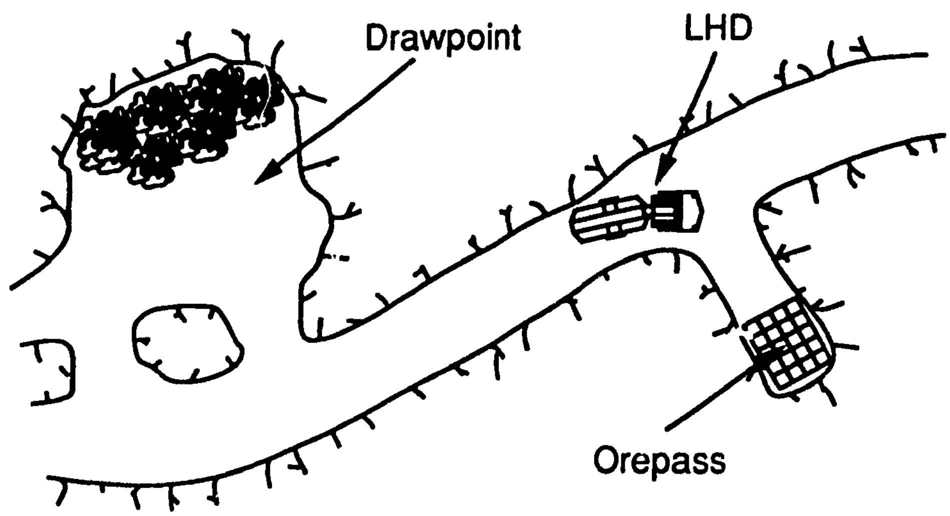


Figure 1.1: A typical plan-view of an underground hardrock mine.

secondary rockbreaking process [Hurteau *et al.*, 1989, Cheung *et al.*, 1990]. This application is the primary focus of this thesis and will be examined in more detail in the next section.

1.2 Rockbreaker automation problem

In a typical underground hardrock mining operation, after the initial drilling and blasting, the broken material, also known as *muckpile*, is loaded onto an LHD vehicle. The LHD vehicle carries the muckpile from the drawpoint and empties its load on top of a vertical *orepass* — a gateway in underground mines for transporting minerals. Despite attempts to avoid the transport of oversized rock lumps, a considerable proportion do arrive at the top of the orepass. Since the orepass can be easily blocked by large rocks (Fig. 1.1), a metal sieve structure, of perhaps 1 m^2 in mesh size and 4 m $\times$ 5 m in dimensions, known as a *grizzly* is placed on top of it to prevent oversized lumps from entering the ore transfer system. It is essential to keep the grizzly clear and free from the accumulation of rock lumps throughout the entire production cycle. A large mechanical device equipped with a jack hammer, commonly referred to as the *rockbreaker* (Fig. 1.2), is employed for clearing

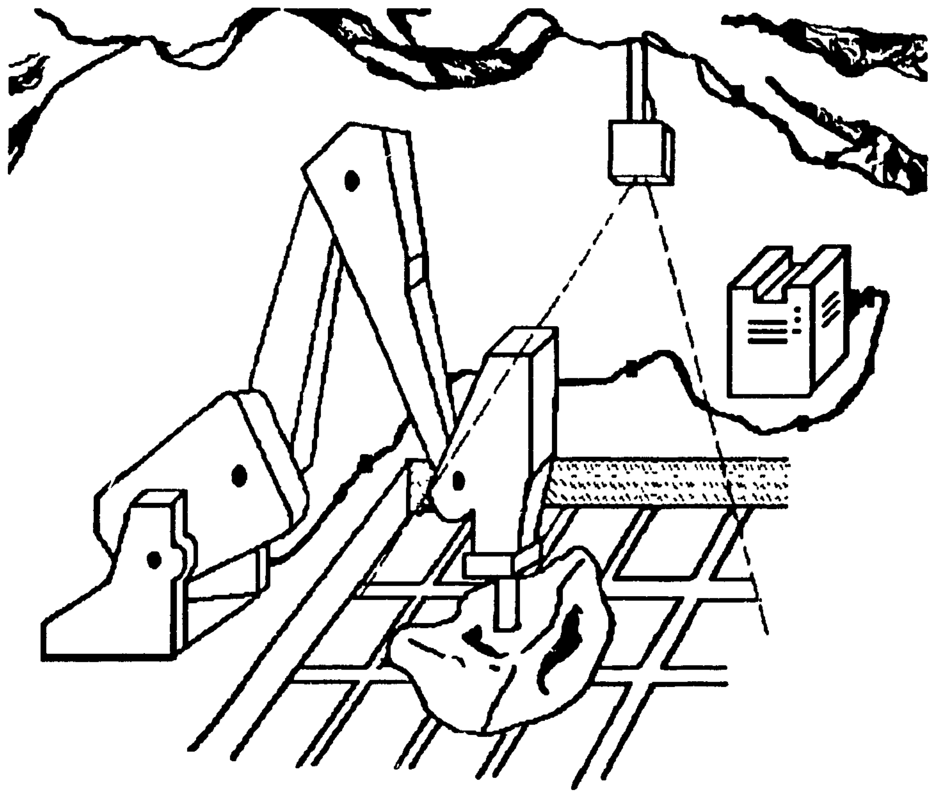


Figure 1.2: Schematic of an automated rockbreaker, consists of both sensing and processing components.

the grizzly and breaking up the remaining oversized rocks. The role of the rockbreaker is to ensure that the production is free from interruption due to blockage of the grizzly.

During a typical, manually controlled rockbreaking operation, a human operator protected inside the control cabin of the rockbreaker, manoeuvres the jack hammer of the rockbreaker directly above the muckpile to be cleared by sieving and hammering. Theoretically, using the mechanical device the operator can reach any location on the grizzly with restrictions on the orientation of the hammer. In practice, it is found to be more difficult to break rocks that are close to the near corners than those located in the middle of the grizzly because of a bad viewing angle and the geometrical configuration of the rockbreaker.

A typical rockbreaking process, starts by cleaning the front of the grizzly and then continues to the back in a sweeping fashion. Sometimes, the operator will go through each hole systematically if he/she is unable to make a decision; e.g., if the muckpile and the small fragments obstruct the view. The size of the target rock is also a very important factor for the breakage. If the rock is large, its sides will be attacked first until it is reduced to a more manageable size. If it is small, the hammer will be aimed directly at its centre. Before applying the hammering, care must be taken to ensure that the tip of the hammer is in good contact with the surface of the rock. Moreover, an appropriate amount of pressure has to be exerted by the hydraulic system during the breaking. Ideally, the operator will position the hammer perpendicular to the rock face. Any other orientation will increase the risk of breaking the hammer's tip, producing uncontrollable flying chips and undesirable movement of the rock.

In positioning the hammer, skill, experience, attention, knowledge of rock mechanics and the location of the grizzly are required on the part of the human operator, making this process an extremely slow and costly one. It may take up to 45 minutes[1] to complete the clearance of one muckpile. Furthermore, the operation of the hammer is based on the operator's vision and perception of the given rock configuration, as well as on his/her experience. Thus, it makes the performance of the rockbreaker very much dependent upon each individual operator.

The rockbreaker resembles an industrial robot in a number of ways; it is composed of articulated joints with four degrees of freedom and has a fixed base. It seems obvious that the knowledge gained from previous robotic and machine vision research can be applied to the rockbreaker automation problem.

The strategy for the rockbreaker problem used here is very similar to the work reported in [Choi *et al.*, 1990] and [Ikeuchi and Hebert, 1990], in the development of a vision-based rock sampling system for a planetary exploration mission that collects terrain samples. More will be said about their work later as the strategy closely resembles that of the present project. However, it differs significantly in the computational methodology used to perform image segmentation.

[1] The breaking time can vary substantially from mine to mine

1.3 Objectives and contributions

This thesis concentrates on the computer vision aspect of the rockbreaker automation project. Although, other issues such as trajectory planning and control are also important for the success of the rockbreaker automation, they are considered to be secondary when compared to the problem of identifying and modelling target rock lumps in the muckpile.

The main objective of this research is to investigate, study and select the existing algorithms developed in computer vision, and apply them to the problem of rock fragment localization and identification.

The main contribution of this thesis, is the successful demonstration of computer vision techniques applied to characterizing the rock shapes on a laboratory scale. More importantly, the complete image processing framework is studied; i.e., range data acquisition, data reconstruction, image segmentation, and fitting of volumetric models.

1.4 Organization of the thesis

The next chapter describes how the measurement of the scene is acquired using the NRCC/McGill laser rangefinder. In addition. alternative range imaging techniques are also briefly discussed. Chapter 3 reviews the basic framework for range data reconstruction and interpretation. Some of the methodologies developed based on standard differential geometry for surface analysis are also discussed. Chapter 4 describes the algorithms involved in our computing strategy. Chapter 5 presents muckpile images at different stages of the processing, and the final validation of the image analysis strategy proposed in this thesis. Finally, Chapter 6 summarizes the results of the research and recommends future research directions.

Chapter 2 Sensor measurements

2.1 Introduction

The ultimate objective of this research is to provide the rockbreaker with the ability to automatically identify, locate and break the oversized rocks remaining on the grizzly. It is obvious that special attention has to be given to the measurement and analysis of the environment if any degree of autonomy is to be delegated to the machine. In particular, *vision* has been regarded as the most important channel of perception. This chapter is devoted to the imaging aspect of the rockbreaker research.

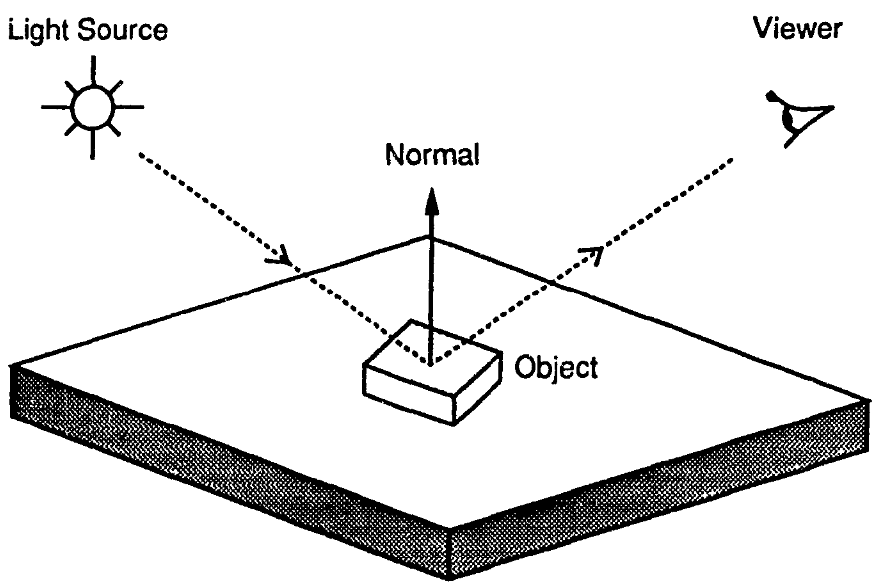


Figure 2.1: A typical image formation — a function of four variables: the position of the light source, the viewing position, the reflectance of the surface, and the geometry of the object.

Different imaging sensors have been developed to acquire various types of measurements. Therefore, one has to begin by understanding how these measurements are created. The visual images perceive in our everyday life are complex functions of four variables [Marr, 1982, Levine, 1985, Horn, 1986] (see Fig. 2.1): (i) the position of the light source(s) used to illuminate the scene, (ii) the position of the viewer, (iii) the reflectance of the surface(s), and (iv) the geometry of the objects in the scene Although, humans and animals do not seem to have any problem inferring the world's structure from visual information, the complexity of the image formation process makes the problem of scene reconstruction a very difficult one. However, researchers have shown that by exploiting additional constraints inherent in the image(s), the geometry and structure of the scene can be recovered [Ullman, 1979, Witkin, 1983, Horn, 1986]. This gives rise to the so-called "Shape from X" paradigm in computer vision, where X could be, shading, texture, motion, focus, stereo disparity, etc.

An important aspect of any automation-related research is the study of how special sensors can be used to advantage. As far as the rockbreaker automation problem is concerned, only the geometrical properties of the rocks are required for performing the breaking task. Three-dimensional imaging sensors have been developed, and can be used to obtain descriptions of the scene geometry through direct surface measurements. The majority of the so-called "range images" are created in this manner.

2.2 Range images

To understand the term "range image", one needs to understand the term "image". An $n \times m$ digital *image* $\mathcal{I}$ is defined in [Levine, 1985], as a function $g(i,j)$ of two discrete variables $i = 1,\cdots,n$ and $j = 1,\cdots,m$, where $g(i,j)$ is a grey-scale measurement. Therefore, an image can be thought of as some function of two indices i and j.

When dealing with range data one must consider the need to cover a three-dimensional surface by collecting data from several different camera positions. Therefore, the notion of an image can be extended and a *range image* $\mathcal{R}$ defined as a function of $r(\vec{c}; i,j)$, follows

$$\mathcal{R} = r(\vec{c}; i,j); \qquad i = 1,\cdots,n,\ j = 1,\cdots,m \tag{2.1}$$

where $\vec{c}$ describes the camera's position and attitude in the scene. The camera position is specified by a translation vector (X_c, Y_c, Z_c) from the scene coordinate origin, and the camera's attitude by Euler angles of rotation $(\theta_x, \theta_y, \theta_z)$ about the scene X, Y, and Z axes respectively, and r is a distance measurement from the camera to a point on the object's surface.

A special case where the image is obtained by making only *one* distance measurement at each camera position, and where the camera is always pointing in the same direction will now be considered. In this case it is the camera position that is a function of the indices

$$X_c = h_1(i,j), \tag{2.2a}$$

$$Y_c = h_2(i,j), \tag{2.2b}$$

$$Z_c = h_3(i,j) \tag{2.2c}$$

where h_1, h_2, h_3 are functions that constrain the camera to lie on a surface, for example a plane, in the scene.

However, the range measurement is not always taken along the direction in which the camera is pointing, and often needs to be transformed into local camera coordinates x, y, and z. For distance measurements that are taken at different camera positions, one must project each measurement from the camera coordinate system xyz to the scene coordinate system XYZ,

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \bar{\mathsf{R}}(\vec{\Theta}_c) \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \vec{\Theta}_c \tag{2.3}$$

where $\vec{\Theta}_c = (X_c, Y_c, Z_c)^{\mathsf{T}}$, and the rotational matrix $\bar{\mathsf{R}}(\vec{\Theta}_c)$ from the (x, y, z) coordinates to the (X, Y, Z) coordinates.

For the sake of simplicity, assume that the scene and camera coordinates are chosen to align so that the camera is always positioned in the scene's XY plane, and hence

$$x = i, \qquad y = j, \qquad z = r(i,j) \tag{2.4}$$

then

$$z = r(x, y). \tag{2.5}$$

Eq. 2.5 is also known as the *graph surface* representation. As a consequence of the indexing scheme, one can define the neighbouring sample points for each measurement. For example, the measurement of $r(i, j+1)$ is a neighbour[1] of $r(i, j)$.

In practice, the two indices i and j are often used to represent the original sampling grid of the image, and depend on the geometry of the measuring device. Range measurements in the form of Eq. 2.1 can be converted directly into the (X, Y, Z) coordinates by means of a calibration look-up table or correction function. The converse is not always true, due to the fact that it is not possible to map (X, Y, Z) from a three-dimensional parameter space back to a six-dimensional parameter space such as the one in Eq. 2.1.

Range images are also loosely referred to by many other terms depending on the context: range map, depth map, depth image, range picture, rangepic, 3-D image, $2\frac{1}{2}$D image, digital terrain map (DTM), topographic map, xyz point list, contour map, and surface height map.

2.3 Range imaging techniques

Range imaging techniques can be grouped into two classes, *passive* or *active*, depending on the sensing method. Passive techniques make use of ambient or unstructured lighting conditions and do not involve any special projections. However, the imaging process often requires *a priori* models of the objects and/or the properties of illumination of the scene. Active techniques perform range finding by projecting onto the scene *structured* or patterned signals such as light stripes, dots, pulses, special patterns or colour. An independent description of the object surface is derived without additional *a priori* knowledge about the scene.

From a practical point of view, active techniques seem to be more appropriate for the rockbreaker project. The mining environment provides an additional set of obstacles and limitations, such as dirt and lighting problems, in the use of the usual passive tech-

[1] The term "neighbour" does not necessarily correspond to the closest sample

niques, like shape from texture or shape from shading. In this section, a number of active techniques are introduced from an engineering standpoint.

2.3.1 Active techniques

Most active range imaging techniques are based on one of the following six principles [Besl, 1988]: (i) radar, (ii) moiré, (iii) lens focus, (iv) triangulation, (v) Fresnel diffraction, and (vi) holographic interferometry. Since principles (v) and (vi) demand set-ups that cannot be realistically obtained in a real-mine environment (see [Besl, 1988] for details), they are excluded from the following discussion.

Imaging radars

Animals such as bats and porpoises are equipped by nature with ultrasonic "radars" to sense their surroundings even under unfavourable lighting conditions. The underlying principle of a radar system is that the distance between the transmitter, the object, and the receiver can be expressed as a function of time. Suppose that the transmitter and the receiver are located in close proximity, and can be treated as a single transducer. Then the distance to the object can be derived from the basic time/range equation:

$$vt = 2z \tag{2.6}$$

where v is the speed of signal propagation and z is the distance from the transducer to the reflecting surface of the object. By calculating the time t taken for the signal to travel from the transmitter to the surface of the object, and from the object back to the receiver, the distance z can be recovered. Of course, Eq. 2.6 can only be true when atmospheric attenuation and other relevant physical properties of the reflecting surface are ignored.

Based on this concept, Jarvis [Jarvis, 1983] built a laser *time of flight* range scanner, capable of taking one range measurement at a time. A high speed, galvanometer controlled, scanning mirror system was employed to deflect a laser beam at a point in the scene. The operating distance for the rangefinder was between 1 and 4 meters, and 2.5 mm accuracy was achieved with 100 samples per point in the ideal situation. With

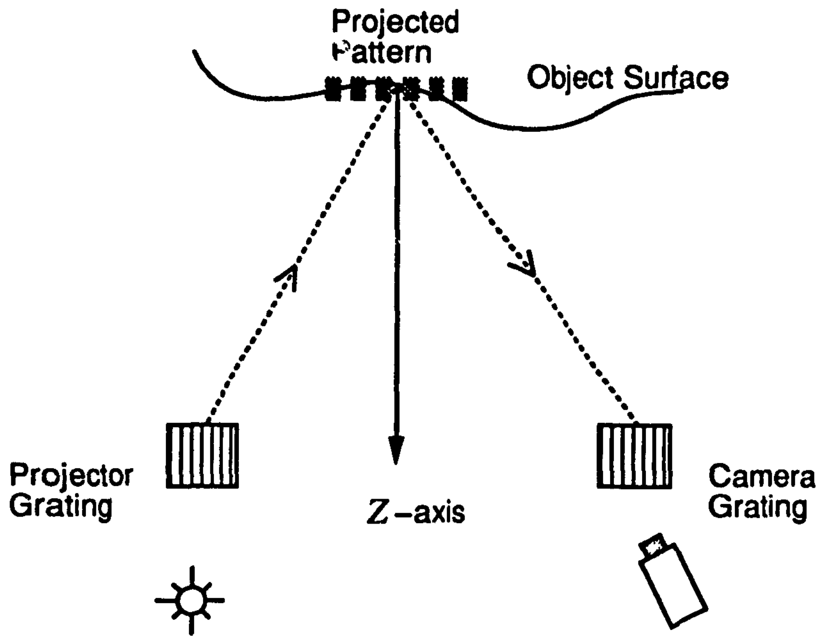


Figure 2.2: Moiré topography configuration.

the laser pulse frequency set at 10 kHz and an output power of 2.5 W, a 10 samples/point average, over a 64×64 size range image, was acquired in 4 s. Constant fraction discrimination was used to detect the return time of the pulse, independent of the pulse size. The total propagation time from "start" (laser pulsed) to "end" (laser returned) is then converted into a signal, whose amplitude is linearly proportional to this transit time.

More recently, a high performance rangefinding system using the same principle has been reported [Kaisto *et al.*, 1990]. The system can measure distance within a range from 3 m to 30 m with mm-level of accuracy. This accuracy is achieved at a maximum rate of 10×10^3 points per second.

Moiré topography

A *moiré* interference pattern can be created by illuminating a scene with the superimposed patterns of two equi-spaced gratings. The moiré projection patterns representing the contours at equal depth can be visualized when viewing the scene through an identical optical camera grating (Fig. 2.2). However, there is no sign information in the contour map to indicate the increase or decrease in depth from one contour to another. Mathematically, the interference pattern $P(x)$ from the two grating patterns P_1 and P_2 is described by:

$$P(x) = P_1 P_2 \Big\{ 1 + m_1 \cos[\omega_1 x + \phi_1 x] \Big\} \Big\{ 1 + m_2 \cos[\omega_2 x + \phi_2 x] \Big\} \tag{2.7}$$

where the P_1 and P_2 are the amplitudes of the patterns, m_1 and m_2 are the modulation indices, ω_1 and ω_2 are the spatial frequencies of the two gratings, and ϕ_1 and ϕ_2 are the spatial phase shifts.

When the signal is low-pass filtered (or blurred), resulting in $P^*(x)$, only the frequency difference and the constant terms remain:

$$P^*(x) = P_1 P_2 \Big\{ 1 + m_1 m_2 \cos\Big[(\omega_1 - \omega_2)x + \phi_1(x) - \phi_2(x) \Big] \Big\}. \tag{2.8}$$

For equally spaced (i.e., identical) gratings, only the phase difference term is left. Therefore, contour lines for different range levels can be recovered by changing the phase of the second grating while maintaining the object and the camera fixed.

Moiré range imaging techniques are suitable for measuring the *relative distance* to surface points on a smooth surface. A commercial range imaging sensor based on a single frame moiré with a reference plane is reported in [Besl, 1988]. The sensor can acquire a 480×512 range image in approximately 2 seconds, with 1 part in 4000 accuracy. Unfortunately, no data on the field of view were quoted in the article.

Lens focusing

Lens focusing techniques can also be used to determine depth. The range measurement is taken by adjusting the lens setting so that the best focus is achieved. Lenses are governed by a simple equation (see Fig. 2.3):

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{F} \tag{2.9}$$

where u is the distance from a point on the object surface to the lens, v the distance between the lens and the reference plane onto which the focused image is projected, and

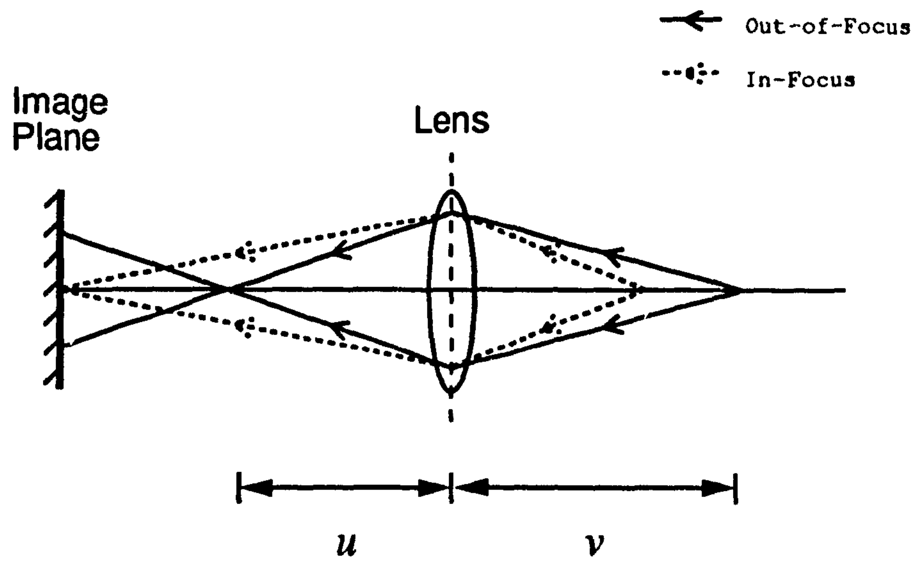


Figure 2.3: Range from focusing principle.

F is the focal length of the lens. Given the geometric setup of the apparatus, the distance v can be easily computed.

Two similar techniques [Rioux and Blais, 1986] are developed based on lens focusing and are closely related to the triangulation methodology (discussed in the next section). In the first technique, an array of point light sources is projected onto the scene, while using an annular mask (with a circular opening) in the aperture of the objective lens. A reference plane at a known distance is set in-focus by the camera focus adjustment such that, points at the reference plane will have no effect on the image, whereas points at different heights will form a blurred circle on the image. The relative distance from each point to the reference plane is computed from the radius of the blurred image in the focal (image) plane of the camera; e.g., the bigger the radius, the larger the distance. The designed sensor is capable of measuring depths of 144 points with an accuracy of ± 1 mm over a 100 mm depth of view.

The second technique requires that the scene be illuminated by multiple light stripes while using a double aperture mask in the camera lens. By following the same set-up procedure for the reference plane as in the previous one; when the light stripe is not in-focus, the camera sees the split lines. Similar to the first technique, the relative distance to the reference plane is proportional to the splitting distance. The instrument is equipped with a specially designed signal processing element to detect the maximum returned signal on the CCD (charge-coupled device) array. This information is then used to derive the line splitting distances on each scan sequence for the whole image. The sensor, also known as "BIRIS" [Blais and Rioux, 1986], is capable of capturing a 256×256 range image in less than 1 second by analyzing 10 projected lines in every 24 frames. In each frame, the projected light stripes are shifted. A resolution of 1 mm over a depth range of 25 cm is achieved.

Triangulation

Active triangulation is probably the most popular technique for acquiring range images, and many commercially available sensors are based on this principle. Fig. 2.4 shows a simple configuration of an active triangulation range sensing system. A light beam is projected onto the object in the scene and the resulting illuminated pattern is imaged by the detector. Knowledge of the spatial parameters of the instrument, the position of the image on the detector, the lateral separation b (base line) between the detector lens and the light source, and the projection angle θ of the source — allows the determination of the distance z by means of solving a simple trigonometry problem.

Most triangulation-based range sensors require structured light to illuminate the scene. One distinct feature associated with the structured light projection is the so-called *shadow effect*[2], which appears in the following locations: (i) at points on the surface where the projected light cannot reach, or (ii) at points on the surface where the projected light is occluded from the sensing elements due to the presence of intervening substance. Consequently, no data is obtained for the image element that corresponds to these locations. The absence of range data may also be due to poor surface reflectance or other artifactual responses of the sensor. The shadow problem has always been seen as a major drawback of structured light-based sensors.

An example of a commercial sensing system using this principle is the Jupiter Series,

[2]Also referred to as the "missing parts" problem

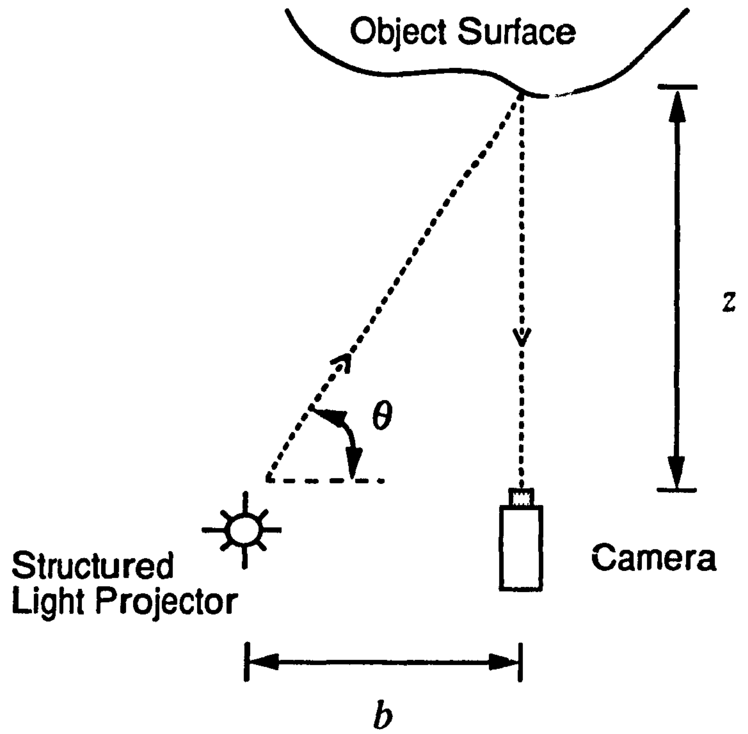


Figure 2.4: Simple triangulation range finding geometry.

which is marketed by Servo-Robot and is capable of scanning 3000 points/s in a viewing space of 1 m^3, with the volume-centre resolution of 1 mm in the x, and 0.3 mm in the z direction. A similar scanner patented by Rioux [Rioux, 1984], with a moderate acquisition rate and high resolution is used in our study. In addition to the devices mentioned above, other designs involving the projection of structured light stripes/patterns exist, and are presented in [Kanade, 1987, Besl, 1988].

2.4 NRCC/McGill laser rangefinder

The rangefinder used in this study is the result of a joint development project between the National Research Council of Canada (NRCC) and McGill University. The technique employed is based on optical triangulation using a novel geometry (synchronized scanning) invented at NRCC by Marc Rioux [Rioux, 1984]. A schematic diagram of the scanner is shown in Fig. 2.5.

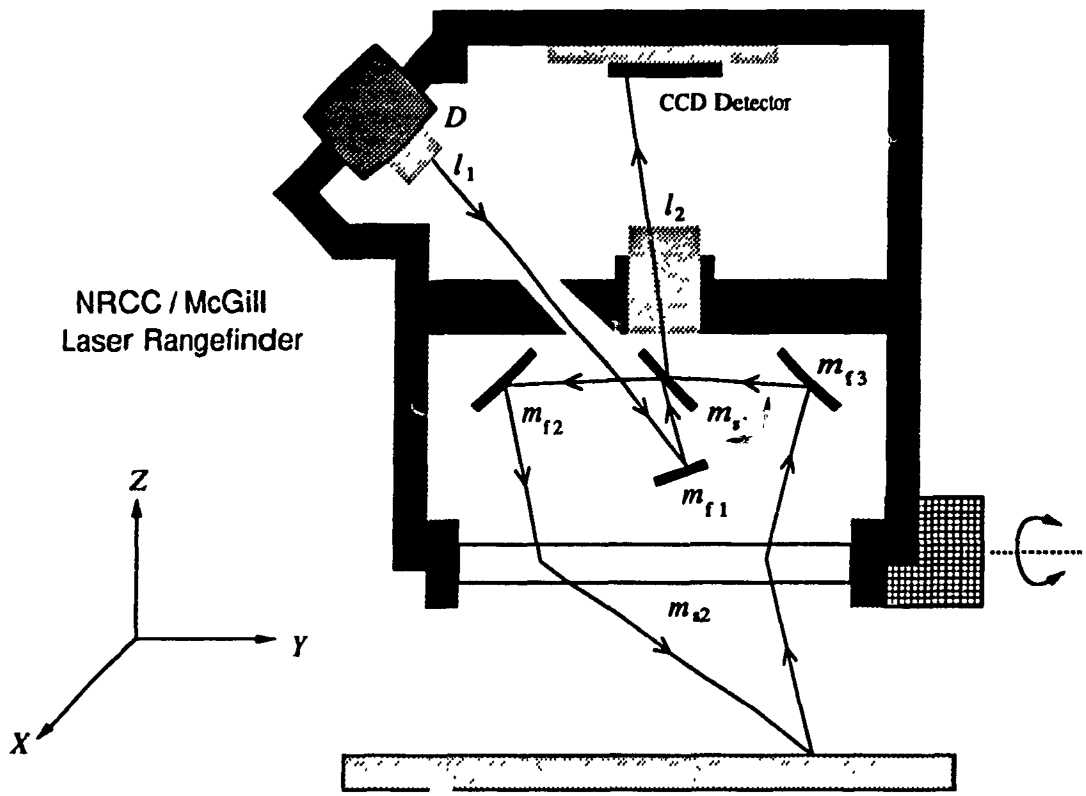


Figure 2.5: Schematic of the NRCC/McGill laser rangefinder.

In the existing laboratory prototype, an He-Ne laser is used as the energy source. The laser beam, on entering the compact optical head via a fibre-optic cable onto a fixed mirror m_{f1}, is then reflected to one side of a double-sided, coated scanning mirror m_{s1} along the X axis. From there, it reflects onto a second fixed mirror m_{f2} which directs the beam onto a second scanning mirror m_{s2} along the Y axis, which finally projects the beam to a point on the object's surface. The reflected light follows a symmetrical path back onto the opposite side of the first scanning mirror (m_{s1}) from where it is deflected onto a linear CCD array. The two scanning mirrors are driven by two galvanometers, one for the X direction and the other for the Y direction. The synchronization of the scanning geometry is maintained by a specially designed timing circuitry [Livingstone and Rioux, 1986], so that the orientation of the two scanning mirrors is acquired simultaneously. This makes random access to any pixel in the field of view possible.

Field of View:	
single axis	100 cm × 100 cm
dual axis	100 cm × 100 cm × 100 cm
Resolution:	
X axis, closest approach	0.40 mm/pixel
X axis, most distant point	4.00 mm/pixel
Y axis, closest approach	0.40 mm/pixel
Y axis, most distant point	4.00 mm/pixel
Z axis, closest approach	0.20 mm/pixel
Z axis, most distant point	1.40 mm/pixel
Acquisition Speed:	
single axis	20.0 lines/second
dual axis	15 seconds/frame (256 lines)
Approximate Dimensions:	
main body	18 cm × 13 cm × 5 cm
motor shaft protrusion	2 cm
Power Output:	Approximately 6.0 mW (final mirror)
Weight:	Approximately 1.0 kg

Table 2.1: Specifications of the NRCC/McGill prototype laser scanner based on the synchronized scanning principle.

The scanning geometry of the rangefinder is designed in such a way that, for a given orientation of the two scanning mirrors, the distance along the Z axis from the object to the scanner is largely proportional to the displacement of the returned laser signal on the CCD detector array. Thus, from the measurements of the orientations of the two mirrors and the beam deflection, the distance to any point on the surface of an object on the scan line can be determined. A look-up table [Bumbaca *et al.*, 1986, Archibald and Amid, 1989] is used to correct for the geometrical distortion and the non-linearities resulting from the optics and the scanning mechanism. One reason for using a look-up table is that high accuracy is achieved without compromising the data acquisition rate. It is worth noting that the NRCC/McGill laser scanner geometry employs a relatively small angular displacement separation between the energy source and the CCD detector, and thereby considerably reduces the occurrence of shadow effects in the range image.

The laboratory prototype has been designed with a number of objectives in mind: a minimal shadow effect, a compact and lightweight unit, high accuracy and a moderate acquisition rate. The specifications for the resulting device are listed in Table 2.1. In a research environment, it is highly desirable to have the flexibility of moving a scanner around to collect information at different viewpoints. The compact and lightweight design enables the scanner to be mounted on the gripper of an industrial robot such as a PUMA 560. This permits various views of the objects to be acquired, and therefore, provides the ability of overcoming the problems due to object occlusion, shadowing, and insufficient data. A discussion of the problem of multiple view integration from range images can be found in [Soucy, 1992].

2.5 Practical problems

The NRCC/McGill laser scanner has proven very useful in solving many laboratory scale problems in the past. However, increasing the 1 m^3 field of view of the current version of the scanner to a larger and more practical volume of perhaps 125 m^3 (5 m $\times$5 m $\times$ 5 m) for a mine environment, can create a number of technical and safety problems. For example, several hundred mW of laser power will be required in order to cover this volume. This may cause safety hazards to the workers in the proximity of the scanner unless adequate protection is provided[3]. The 125 m^3 field of view may also create some technical difficulties in the optical design of the device, but these have been viewed as secondary in light of the power requirements.

A similar design of the scanner already exists for welding applications [Beranek *et al.*, 1986]. The device is environmentally sealed, temperature stabilized, and equipped with an air jet to eliminate the disturbance caused by smoke and fumes. Therefore, it seems feasible to incorporate the same type of technology in the mine environment.

Alternatively, other range finding technologies as mentioned in this chapter, are commercially available with a large depth of view (see Table 2.2). However, not all these devices are readily available for the mine environment, and much experimentation will be required before a device could be considered viable in a mine environment.

[3]The chances of getting struck by the laser source reduce considerably when the beam is scanning.

Category	Typical Resolution	Typical Depth of View	Typical Applications	General Comments
Radar	0.1 mm – 20 cm	2 m – 100 m	Cartography, target detection and navigation.	Lack of precision, but should have sufficient accuracy for the characterization of muckpile shape. Main problem is the relative slow acquisition rate.
Moiré	1.0μm –	50 mm – 10 m	Shape analysis, inspection and assembly.	Moiré topography has been around since 1859. Capable of obtaining high accuracy measurements, but the major drawback is that only a few designs have a large field of view.
Focusing	1.0 mm –	150 mm – 10 m	Navigation and object manipulation.	High acquisition rate could be achieved with compromise on the accuracy. Has great potential in mining applications, but experimentation in an actual mine environment would be required before any real assessment could be made.
Triangulation	1.0μm –	100 mm – 10m	Navigation, assembly, inspection, object manipulation and shape analysis.	The most popular technique for range finding; commercial products are widely available Safety could be a problem if high power laser source is used for a large field of view.

Table 2.2: Brief overview of the four active range finding techniques (radar, moiré, focusing and triangulation).

Field of View:	500 cm $\times$ 500 cm $\times$ 500 cm
Resolution:	
X axis, closest approach	0.5 cm/pixel
X axis, most distant point	2.0 cm/pixel
Y axis, closest approach	0.5 cm/pixel
Y axis, most distant point	2.0 cm/pixel
Z axis, closest approach	0.2 cm/pixel
Z axis, most distant point	1.0 cm/pixel
Acquisition Speed:	$\leq$ 5 seconds/frame (512 $\times$ 512 pixels)
Approximate Dimensions:	
main body	$\leq$ 30 cm $\times$ 30 cm $\times$ 20 cm
Power Output:	Possibly within the eye safety level
Weight:	$\leq$ 5.0 kg

Table 2.3: Proposed rangefinder specifications for the rockbreaker application. These specifications are far from complete and only serve here as an illustration of what is expected from such a device.

2.6 Summary

Four types of active range finding techniques were discussed in this chapter. Out of all these techniques, no one method seems to be clearly superior. Various approaches based on the same principle may yield unequal performances and/or accuracies depending on the hardware design and instrument set-up. One technique, suitable for a particular application may well prove to be inappropriate for another.

The NRCC/McGill laser scanner is chosen in this study, partly because of its high resolution and the relatively short acquisition time. In this study, the concentration is mainly focused on deriving meaningful descriptions of the scene.

Nevertheless, in selecting or designing a rangefinder that qualifies for the rockbreaker application, one needs to consider a large number of issues both technical and non-technical. The proposed specifications for such a rangefinder are presented in Table 2.3.

Chapter 3 Reconstruction and interpretation of sensor measurements

3.1 Introduction

The information obtained from a rangefinder consists of direct distance measurements of points on the surface, and they are often very difficult to interpret. A more explicit representation of these surface measurements is therefore required before the geometric properties of the object can be inferred. In this chapter, a basic framework for analyzing range images is reviewed. The so-called "bottom-up" or "data-driven" paradigm is made-up of three levels of processes; namely — image reconstruction, image segmentation and inference of scene geometry (Fig. 3.1). In this chapter, more will be said about each of these processes, together with discussions on some of the methodologies developed in the past for range image analysis. However, higher levels of processes such as, object recognition and object manipulation also exist but they are not addressed in this thesis.

3.2 Surface reconstruction

Image reconstruction is referred to in [Blake and Zisserman, 1987], as the process that produces stable and reliable representation of the scene from discrete data samples. For range images, the image reconstruction is limited to the recovery of the geometry structure of the scene, which is best described by its surface properties. One advantage of using the surface properties is that they indicate all the essential features for the surface characterization. Thus the accuracy of the reconstructed surface is very critical to the entire image analysis process. Many surface reconstruction methodologies have been evolved over the past few years, and all are essentially regularization methods, that transform the inherently *ill-posed* problem into a *well-posed* one by imposing additional constraints [Grimson, 1983, Terzopoulos, 1983, Blake and Zisserman, 1987].

However, the term "surface reconstruction" is somewhat misleading, as its definition varies from one researcher to another. To avoid confusion, *surface reconstruction* is

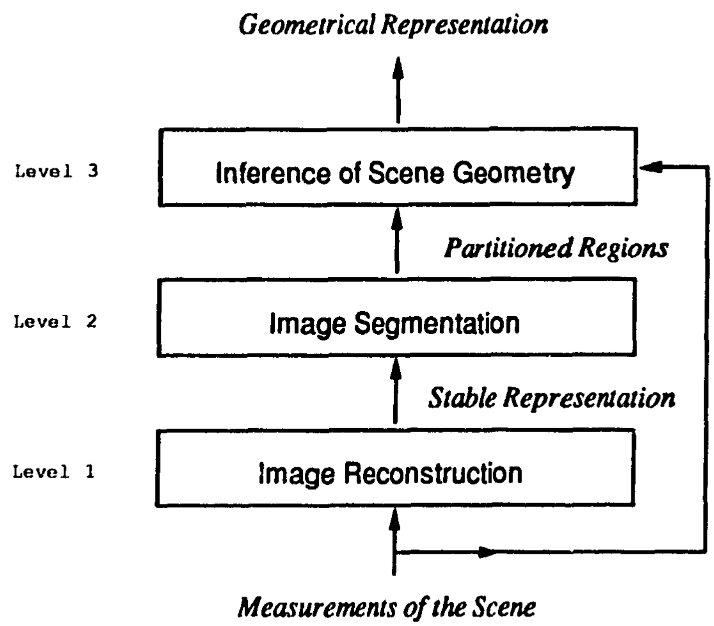


Figure 3.1: Block diagram of the "bottom-up" paradigm for image analysis. Three levels of processes are shown here, namely — image reconstruction, image segmentation and inference of scene geometry.

referred to here as the process that recovers the stable local surface structure from sensor measurements. More will be said in the next chapter about how local surface structure can be recovered stably and reliably.

3.2.1 Recovery of the local surface structure

In the past, some success have been reported using surface differential properties as tools for surface analysis, directly from the available range data [Brady *et al.*, 1985, Faugeras and Hebert, 1986, Besl and Jain, 1988]. These properties at each discrete sample are best described by using a local surface representation. This surface representation provides the essential information for the higher level processes such as image segmentation.

Local surface representation

At any point P on a *smooth surface* **S** one can define a *tangent plane* T_P that describes the local surface orientation at P. Following the definition of a "smooth surface" found in [Bennett and Hoffman, 1987], a surface is considered to be *smooth*[1], provided that every curve on the surface is a C^2 function; i.e., a function with all derivatives up to and including the second order derivative are continuous[2]. The *surface curvatures* at any point P, can be seen as the second derivatives of the parametric curves[3] at P on the surface.

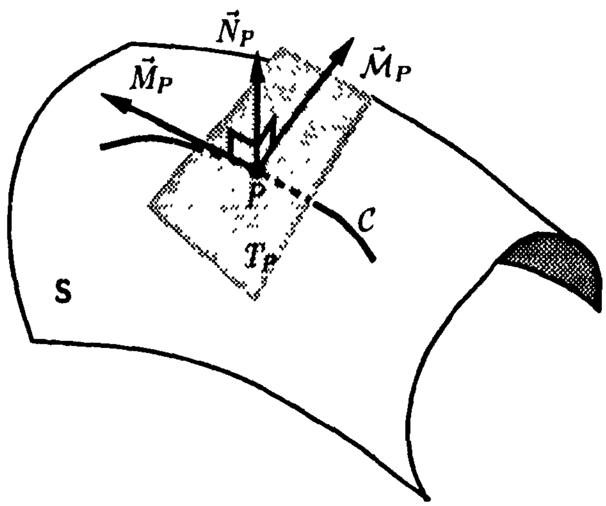


Figure 3.2: Local representation of a surface — the augmented Darboux frame $\mathcal{D}(P) = (P, \kappa_{MP}, \kappa_{\mathcal{M}P}, \vec{M}_P, \vec{\mathcal{M}}_P, \vec{N}_P)$ along the oriented curve C on a smooth surface **S**.

A simpler way to understand the notion of *surface curvature*, is to consider a series of vectors through the point P in the tangent plane T_P. Each direction of these tangent vectors (say $\vec{V}_P$) specifies a curve C on the surface. The curvature of the curve C can be measured by its tendency to "bend" out of the tangent plane; i.e., the greater of the tendency of bending away from T_P — the larger is the curvature, and vice versa. This curvature is known as the *normal curvature* κ_{nP}, in the direction specified by the vector

[1]Mathematically, a "smooth" curve is defined as one which is C^∞. This implies a function with continuous derivatives all the way up to infinite order [do Carmo, 1976].

[2]The term "continuous" is used here to refer to a function whose derivatives exist and are computable

[3]See [do Carmo, 1976] for the definition of a parametric curve

$\vec{V}_P$, and whose sign is determined by the orientation of the surface normal $\vec{N}_P$. Normal curvature forms the basis of all surface curvature measurements, and it is very useful in surface analysis. There exist two normal curvatures at every point on every curved surface, referred to as the two *principal curvatures*. One has the maximum and the other has the minimum value, and they are denoted by the symbols κ_{MP} and $\kappa_{\mathcal{M}P}$. The directions associated with these two normal curvatures are commonly referred to as the *principal directions*, and they are represented by two vectors $\vec{M}_P$ and $\vec{\mathcal{M}}_P$ respectively (see Fig. 3.2). One very important fact is that $\vec{M}_P$ and $\vec{\mathcal{M}}_P$ are always orthogonal to each other [do Carmo, 1976]. This forms the basis of the surface reconstruction algorithm employed in our study. This algorithm will be discussed in the next chapter.

The local surface properties at each sample point, can be described by an orthonormal frame referred to as the *augmented Darboux frame*[4] [Sander, 1988, Sander and Zucker, 1990]. The augmented Darboux frame $\mathcal{D}(P)$ at point P, is a collective unit of two scalars and three unit vectors (Fig. 3.2). The two scalars are the magnitudes of the two principal curvatures κ_{MP} and $\kappa_{\mathcal{M}P}$, the three unit vectors are, the surface normal $\vec{N}_P$, the two principal directions $\vec{M}_P$ and $\vec{\mathcal{M}}_P$.

3.2.2 Local estimation techniques

After the local surface representation for a smooth surface is defined, the next step is to estimate the surface properties at each sample. A number of local methods have been proposed in the computer vision literature [Flynn and Jain, 1989]. Some used the analytical fitting of local surface patches to the range data, while others estimated the surface derivatives or curvatures directly from the range measurements. Typically, a surface estimation involves two steps: (i) compute the surface normals, and (ii) estimate the principal curvatures and principal directions. This section presents a brief description of several methods used for local curvature estimation.

[4]The name "augmented Darboux frame" is after Gaston Darboux [do Carmo, 1976]. The only difference between the augmented and the original Darboux frame, is that the latter does not include information about the directions of the principal curvatures.

Analytical surface estimation techniques

One analytical method, based on the least-squares fitting of surface patch is reported in [Besl and Jain, 1986b]. The estimation of surface normal at each point is carried out by fitting a plane to its neighbouring samples. To compute the surface curvatures, parabolic quadric patches are fitted to the data. This local surface fitting is made through a series of separable convolution operations using a certain size window. Surface curvatures at the sample point are then computed directly from the approximated local surface patch.

Another analytical approach is to apply a B-spline approximation to the original range data as the surface fit [Dierckx, 1977]. The analytical expressions of the surface derivatives required for curvature estimation can be derived quite easily, once the B-spline approximated surface is established. One disadvantage of the B-spline approximation is that any jump discontinuities, including those associated with the physical edges/boundaries of the object are smoothed out by the fitting process. This however may not cause a problem in cases where the samples are relatively smooth and dense.

Curvatures at any point P on the surface, can also be determined by studying the local surface orientation changes in a small neighbourhood [Ittner and Jain, 1985]. One way of accomplishing this, is first to obtain an estimate of the surface normal at each point that best describes the local surface orientation. As mentioned before, the surface normal can be estimated simply by fitting a plane to the neighbouring points using a least-squares approximation. The normal curvatures can then be derived from the neighbouring surface normals.

Direct surface estimation techniques

The technique to be described [Fan *et al.*, 1987], is based on direct estimates of the first and second partial derivatives from the available range data. Based on the values of these derivatives, the normal curvatures at each point are computed in four directions; 0, 45, 90 and 135 degrees. The resulting directional curvatures are then combined to form the estimates of the two principal curvatures at each sample point. It is well-known in numerical analysis that derivative functions are highly sensitive to noise, in which high frequency noise is amplified. However, direct surface estimation approaches are less computationally expensive, and require fewer computing cycles than the analytical approach.

In this section a number of local methods for obtaining the surface estimates are discussed. In general, analytical approaches have shown to be more stable in estimating surface properties than direct computational approach [Flynn and Jain, 1989]. However, one must bear the additional computation expenses that the analytical approaches require.

3.3 Surface segmentation

Marr was one of the first researchers to emphasize that segmentation is a context dependent process, whose goal is often not very well defined [Marr, 1982]. *Image segmentation* has been referred to by some researchers in the past as the operation that analogous to figure to ground separation; i.e., that isolates the objects of interest from the background. This definition is somewhat ambiguous; to what extent can one distinguish the subtle difference between the objects and the background from an image? For example, given an image of an office scene, should one consider the bookshelf as an object apart from the wall, or should the books be treated as objects partible form the bookshelf? Both cases can be correct, because as the definition of an object can vary depending on the succeeding actions to be performed. For this reason, Marr ruled out a general segmentation methodology, applicable to all vision problems. Nevertheless, segmentation is required for most vision-based systems if any higher level processes are to be performed.

Typically, a range image would contain a large amount of information about the geometric structure of the scene. As a result, a direct interpretation is very difficult, and therefore it is often necessary to first partition (segment) the data into different regions/surfaces before the scene geometry can be inferred. There are many ways of decomposing surfaces; for most vision-based systems, these partitioning operations are usually application oriented. Most of the segmentation techniques found in the computer vision literature to date fall in one of the two main categories: (i) data-driven; and (ii) model-driven.

Model-driven approaches require *a priori* knowledge about the objects in the scene; e.g., knowledge of shape, colour, texture, and so on. Some of the most popular model driven segmentation techniques include histogram-based thresholding and template matching, but these methods provide little information when the image data do not conform to the, restrictive, image model assumptions. These techniques seem to work well in structured environments but were proven to be inexpedient in unstructured ones, and because of this, they are excluded from the following discussion.

Data-driven approaches can be further sub-divided into region and boundary-based.

3.3.1 Region-based techniques

A typical *region-based* approach would involve two complementary operations, merging and splitting [Rosenfeld and Kak, 1982, Levine, 1985, Horn, 1986]. *Merging* is a process that combines neighbouring pixels into regions or adjacent regions into bigger ones with similar characteristics. On the other hand, *splitting* is the process that separates one region into two or more regions. It is common that statistical and/or spatial measures of pixel-to-pixel correlation (spatial coherence), is used as an indicator for determining whether merging or splitting operation should be performed. This operation ends when no more regions can be split or merged; i.e., when the number of regions becomes stabilized.

Surface type mapping

Differential geometry [do Carmo, 1976, O'Neill, 1966] and topology have been employed by mathematicians for many decades as the basic tool for characterizing surfaces. One of the advantages of using the local surface model $\mathcal{D}(P)$ is, that it allows to form any arbitrary smooth surface, where the shape can be arbitrary complicated. Given an arbitrary shaped, smooth surface S, one can map S into regions based on the signs of the mean and Gaussian curvatures at each point P. The *mean curvature* of a surface is defined as,

$$H_P = \frac{(\kappa_{MP} + \kappa_{MP})}{2} \tag{3.1}$$

and the *Gaussian curvature* is defined as,

$$K_P = \kappa_{MP} \kappa_{\mathcal{M}P} \tag{3.2}$$

where κ_{MP} and $\kappa_{\mathcal{M}P}$ are the two principal curvatures.

Based on the signs of the Gaussian and mean curvature at each point, one can obtain a qualitative measure of the surface shape. Only eight possible combinations exist as shown in Table. 3.1, and the corresponding surface types are known as: pit, peak, valley, ridge, saddle valley, saddle ridge, planar, and minimal saddle (Fig. 3.3).

	$K_P < 0$	$K_P = 0$	$K_P > 0$
$H_P < 0$	Saddle Valley (hyperbolic)	Valley (cylindrical)	Pit (parabolic)
$H_P = 0$	Minimal Surface (hyperbolic)	Flat (planar)	*not possible*
$H_P > 0$	Saddle Ridge (hyperbolic)	Ridge (cylindrical)	Peak (parabolic)

Table 3.1: Eight fundamental surface types classified by using the signs of the mean curvature (H_P) and Gaussian curvature (K_P).

A number of techniques have been developed to segment range images, based on the $K_P H_P$ mappings. Besl and Jain were among the first, to propose the use of Gaussian and mean curvatures for range image segmentation [Besl and Jain, 1986a, Besl and Jain, 1986b, Vemuri *et al.*, 1987, Yokoya and Levine, 1988], initially segments the image into many small patches according to the eight fundamental surface types. These patches were then merged to form larger ones using an iterative region growing algorithm. Some very impressive results were obtained using this technique [Besl and Jain, 1988].

3.3.2 Boundary-based techniques

Most of the *boundary-based* techniques proposed to date are based on locating various types of discontinuities in the image that associate with the physical edges of the objects [Marr, 1982, Rosenfeld and Kak, 1982, Levine, 1985, Horn, 1986]. In practice, the direct measurement of these discontinuities from raw data, is highly prone to noise, and the

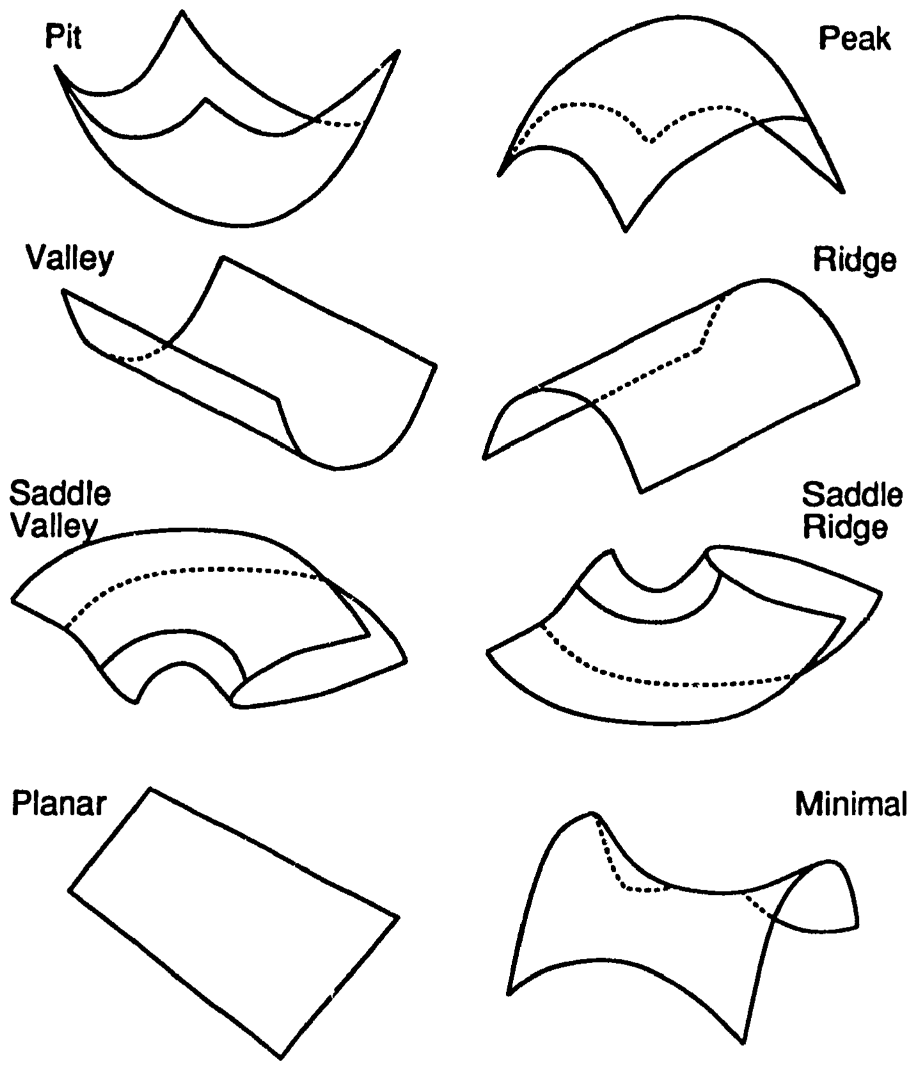


Figure 3.3: Eight fundamental surface types: pit, peak, valley, ridge, saddle valley, saddle ridge, planar, and minimal saddle.

resulting edge segments are often unreliable and picture dependent [Canny, 1986].

Partitioning Boundary detection

In the last few years, various surface boundary-based techniques have been proposed for range image segmentation, based on recovering different types of surface features [Brady *et al.*, 1985, Ponce and Brady, 1987, Fan *et al.*, 1987].

One surface decomposition theory has been put forward by Hoffman and Richards [Hoffman and Richards, 1984], in which they proposed that surfaces can be decomposed into parts by identifying the associated partitioning boundaries. In their article

[Hoffman and Richards, 1984], Hoffman and Richards began by illustrating a few line-drawing diagrams, and they subsequently demonstrated that, humans have the incredible capability of recognizing objects, even in situations when information such as shading, motion, colour and texture are absent. They argued that shape alone is sufficient for object recognition. Subsequently, they advocated that the following *transversality principal* should be used as the basis for shape decomposition.

> **Transversality principle:** An interpenetration of two arbitrarily shaped smooth surfaces results in a concave discontinuity of their tangent planes along the contour of intersection. In the context of smooth surface decomposition, the partitioning contour is located at *negative curvature minima*.

The fundamental idea embedded in this particular part decomposition strategy is, to treat each complex object as a configuration of irreducible primitives, each referred to as a *part*. Any arbitrary complex shaped object can then be made-up, using a combination of parts, and possibly with different sizes and shapes. However, each part is confined to be *convex* and *compact* in shape. (Note that concavity is created alone the intersection of two adjoined parts.) Compact is used here to refer to a surface without any dents or depressions.

Another part theory is documented earlier in [Koenderink and van Doorn, 1982]. In the article, Koenderink and van Doorn suggested part boundaries are contours on a surface where the Gaussian curvature is zero, the so-called *parabolic* contours. Such contours possess a number of nice properties. For example, parabolic contours do not intersect and always form enclosed boundaries. By using them as partitioning boundaries, Koenderink and van Doorn derived four classes of parts, namely; humps, dimples, furrows and ridges. Although, this partitioning strategy worked well on some smooth surfaces of genus zero (i.e., no holes), but clearly it poses limitations when representing complex shapes with only four classes of part primitives.

Feature aggregation

One task that boundary-based approaches have to accomplish is *feature aggregation*, which is an operation that labels and connects recovered boundary points into groups that resemble the structure of the world.

The next stage of processing involves aggregating individual boundary points into contours. The feature points along do not provide enough information for the surface decomposition, because to segment the surface into regions, it requires enclosed contours to represent each individual region. Finding the enclosed contours for each region is akin to filling in a drawing by connecting a series of dots. Where the dot density is high enough, the interpolation of the contour is fairly obvious and can be accomplished with a number of spline interpolation algorithms. However, as the density decreases, it is sometimes not obvious how to interpolate the contour without additional constraints.

3.3.3 Hybrid techniques

More recently, a hybrid technique is reported in [Gupta and Bajcsy, 1990]. In the paper, Gupta and Bajcsy purposed a paradigm for part description and segmentation that integrates various types of information obtained from different levels of processes. Three levels of processes are involved in the paradigm; the occluding contour, the surface and the volumetric levels. They argued that no single level of processes is robust enough to capture all the details of objects in the scene. The segmentation process should proceed first by obtaining the local occluding contours and surface descriptions. At the higher (global) level, a curve segmentation module and a surface segmentation module are used to refine the segmentation in a fine to coarse fashion via two feedback loops; one internal and one external. These feedback loops are controlled by a decision making module which evaluates and integrates information obtained from the curve segmentation, the surface segmentation, and the fitting of superquadric model (the latter is introduced in the next section). A number of examples are illustrated in [Gupta and Bajcsy, 1990].

3.4 Inference of scene geometry

All computer vision applications require a final description to represent the scene, which is application specific. For most practical problems such as object manipulation and object recognition, one needs a description that can capture the geometrical properties of the objects in the scene such as, position, orientation, shape and size — as well as other intrinsic properties such as, colour and texture. However, the range information obtained from a standard rangefinder makes the recovery of the latter intrinsic properties rather difficult, and they are not addressed here.

3.4.1 Parametric shape modelling

There exist many different approaches to obtain a three-dimensional representation of a scene, and the most common approach is based on the use of parametric models. In the past, a number of researchers in computer vision nave proposed the use of generalized cylinders to describe different parts of objects in a scene, by starting with the local approximation of the axis, and gradually recovering all parameters of the generalized cylinders [Binford, 1971, Marr, 1982 Brooks, 1983]. Another approach uses combinations of ellipsoidal and cylindrical models to form a coarse object representation [Ferrie and Levine, 1988]. For a more general approach to 3-D representation, a complex family of solid models is considered in [Pentland, 1987, Solina and Bajcsy, 1990]. This has the capability of modelling a large set of standard geometrical shapes, and yet is simple enough that their parameters can be solved using standard numerical methods [Press *et al.*, 1988]. Only the latter parametric solid primitive, the *superquadric*, is discussed in the following section, because it represents a large class of parametric shapes including the ones mentioned above.

Superquadric Models

The *superquadric* model was first discovered by a Danish writer and designer named Peit Hein, and has the capability of describing a wide range of three-dimensional shapes [Gardner, 1965, Gardner, 1975]. Hein used the *superellipse*, a 2-D sub-set of superquadric

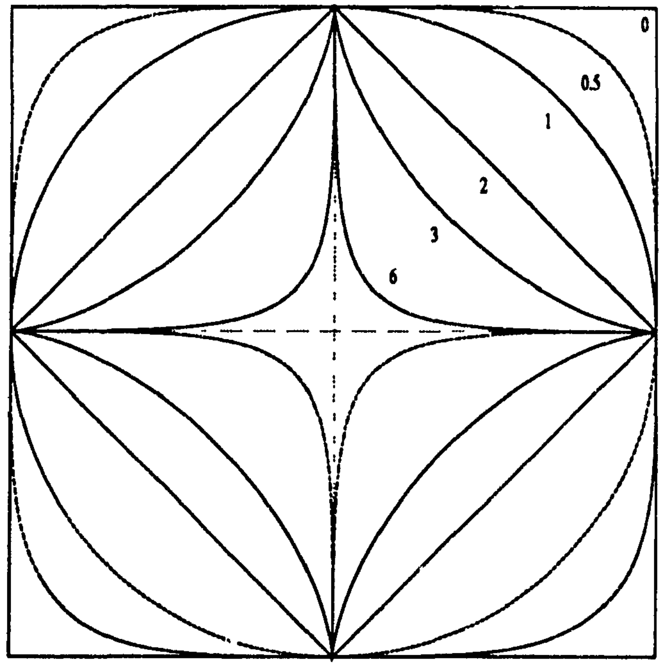


Figure 3.4: Two-dimensional parametric shapes (superellipse): $|x|^{2/\epsilon} + |y|^{2/\epsilon} = 1$ for the relative shape parameter $\epsilon = 0, 0.5, 1, 2, 3$ and 6. For simplicity, the two size parameters a_x and a_y are set to be unity.

(Fig. 3.4), and solved a city-planning problem which arose in 1959 in Sweden. During that period, he designed the outer-shape of a fountain that fitted harmoniously into a rectangular open space located at the heart of Stockholm. Since then, superquadrics have been employed by researchers from both, the computer graphics and computer vision communities for solid shape modelling [Barr, 1981, Pentland, 1987].

To better explain the characteristics of the superquadric, let us first consider a 2-D family of parametric shapes (superellipse) described by the following function,

$$f(x,y) = \left|\frac{x}{a_x}\right|^{2/\epsilon} + \left|\frac{y}{a_y}\right|^{2/\epsilon} = 1. \tag{3.3}$$

where ϵ is the *relative shape*[5] parameter, a_x and a_y are the parameters that define the superellipse *size* in x and y coordinates respectively, and they are also known as the *radial aspects*. By varying ϵ from 0 up to infinity, one can obtain a wide variety of shapes

[5] ϵ is also referred to in [Solina and Bajcsy, 1990] as the "squareness" parameter

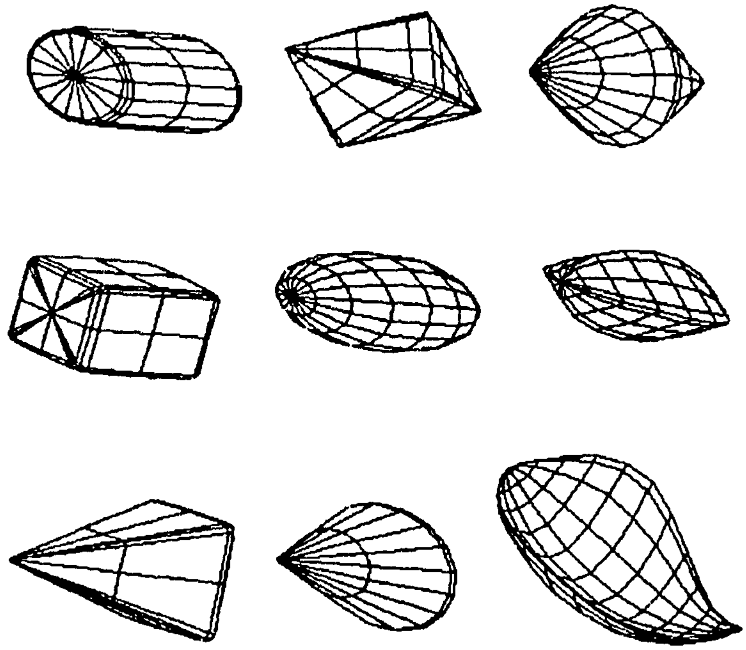


Figure 3.5: Samples of superquadric surface model by varying ϵ_1 and ϵ_2 between 1 and 2 respectively.

(see Fig. 3.4). Starting at $\epsilon = 0$, a perfect rectangular shape is produced. By increasing the relative shape parameter to 1, the squarishness of the curve gradually disappears and the shape turns into an ellipse. As ϵ increases further to 2, the shape transforms from an ellipse into a rhombus. When the relative shape parameter gets larger than 2, the shape becomes concave, and as the parameter approaches infinity, the shape turns into a "cross-like" figure with zero cross-section area.

One can expand the 2-D parametric equation Eq. 3.3 into 3-D, and derive an *implicit equation* for the surface of a superquadric model (see Fig. 3.5),

$$f(x,y,z) = \left(\left| \frac{x}{a_x} \right|^{2/\epsilon_2} + \left| \frac{y}{a_y} \right|^{2/\epsilon_2} \right)^{\epsilon_2/\epsilon_1} + \left| \frac{z}{a_z} \right|^{2/\epsilon_1}. \tag{3.4}$$

The parameters a_x, a_y and a_z, define the size of the superquadric corresponding to the x, y and z axes in the object centered coordinate system. The parameters ϵ_1 and ϵ_2 define the relative shape of the superquadric in the latitude (xz) plane and longitude (xy) plane respectively.

Eq. 3.4 is commonly referred to as the *inside-outside* function [Barr, 1981, Gross and Boult, 1988, Solina and Bajcsy, 1990], because superquadrics are mathematical solid models, and their surfaces can be divided into three distinct regions for a given point $P = [P_x, P_y, P_z]^T$ in 3-D space. These regions are defined as, if,

$f(P_x, P_y, P_z) = 1$ then P lies on the surface,

$f(P_x, P_y, P_z) > 1$ then P lies outside the superquadric,

$f(P_x, P_y, P_z) < 1$ then P lies inside the superquadric.

There exist a number of different approaches for recovering the model parameters, and they all share one thing in common — by defining a measure of the "error of fit" in their fitting functions [Pentland, 1987, Gross and Boult, 1988, Solina and Bajcsy, 1990, Whaite and Ferrie, 1991]. Pentland initially suggested to solve the model recovery problem analytically for all independent parameters [Pentland, 1986]. However, an analytical solution to the problem turned out to be very complicated for most general cases. Later, Pentland [Pentland, 1987] combined the part model recovery with segmentation, by searching through the entire superquadric parameter space for the "best" fitted model. This method has proved to be computationally expensive. In spite of the computational problems he had encountered, Pentland successfully demonstrated the power of using superquadrics in representing a wide variety of objects ranging from natural scene to man-made objects.

Another method, the so-called "minimum volume" approach, was first motivated by [Bajcsy and Solina, 1987], due to the facts that there exist situations where a set of superquadric parameters can be found, and they all fit equally well to the range data obtained from a single view-point. The basic idea behind the minimum volume approach is to select the smallest superquadric as the part model among all possible solutions. Although experiments have shown that the minimum volume model tends to produce more intuitive results, there are a number of drawbacks associated with this approach [Solina and Bajcsy, 1987, Solina and Bajcsy, 1990]. For example, if the recovered minimum volume model is wrong (i.e., too small and does not represent the corresponding part properly), then this can have disastrous effects in many cases; e.g., when a robot attempts to manipulate the part/object. Other measures of "error of fit", like the mean square value of the inside-outside function or the true Euclidean distance are also possible [Gross and Boult, 1988].

Measures of the "error of fit" are essential to the fitting process, however, all methods mentioned above failed to address the uniqueness of the fitted model when multiple solutions are available for a single part. This is very important for "higher" level processes such as, object recognition or path planning. A formal discussion on the uniqueness of the recovered parametric model is presented in [Whaite and Ferrie, 1991], along with a novel method of evaluating the uncertainty associated with the model.

Deformable parametric model

Barr was among the first to suggest that local and global deformations can be achieved by tapering, twisting, and bending of parametrized solids [Barr, 1984]. Additional expressive power of the solid modelling can be obtained by the ability to deform the superquadric model, both locally and globally into the desired shape. This is very useful for capturing the complex surface appearance of natural objects. It was later realized by Pentland [Pentland, 1987], Solina and Bajcsy [Solina and Bajcsy, 1990] in the recovery of parametric model from range data.

More recently, with the advances in the parametric solid modelling, physical and dynamic constraints are being embedded mathematically into the parametric model. In addition to the geometry recovered from the standard parametric model, the formulation of physically-based models can include simulated forces, masses, strain energies, and other physical quantities. Therefore, the physically-based model has the advantage of analyzing and predicting the motions and interactions of complex objects. Notably, there have been two physically-based modelling approaches proposed to date: (i) Pentland's "modal" analysis approach, and (ii) Terzopoulos and Metaxas' "dynamic" approach.

Pentland's "modal" analysis approach, is somewhat analogous to Fourier transform, with low-order modes to provide a description of the overall shape, and high-order modes to represent the high-frequency surface details [Pentland, 1990, Pentland and Sclaroff, 1991]. More importantly, the mathematical formulation is based on finite element method (FEM) that provides an analytic characterization of surface between nodes or pixels. For example, an object may be interpreted as a mesh of nodes, with a certain mass, damping and stiffness between the nodes. Subsequently, virtual forces are defined at each node such that the model can be deformed to fit the data samples. The final shape may be thought of as the result of pushing, pinching and pulling on a lump of elastic material (such as clay), starting with a spherical approximation. One major problem of this particular approach is the instability of modelling non-convex objects.

In Terzopoulos and Metaxas' "dynamic" approach [Terzopoulos and Metaxas, 1991], the so-called *deformable superquadrics* are governed by a set of equations of motion. They augmented the models with the local deformation capabilities of membrane splines. As a result, virtual (external) forces are permitted to deform the physically-based models globally like superquadrics in order to recover the translation, rotation, scale, three radial aspects, and two squareness parameters. In addition to this, the forces also deform the models locally like splines to reconstruct the fine structure and the natural irregular appearances from the data. The mathematical formulation of these physically-based models is rather complex, and beyond the scope of this thesis.

3.5 Hidden factors

In almost every computer vision application, one has to confront the scale, and the resolution problems. Although, *scale* and *resolution* are closely related to each other, they are treated as independent in this section.

3.5.1 Resolution factor

Resolution is considered here as a physical attribute of the image; i.e., the number of pixels in the image, and the number of bits per pixel. In other words, the resolution of a image is usually dependent upon the hardware limitation or the image format.

3.5.2 Scale factor

Scale is referred to here as the mechanism that controls the level of detail in the image for a given resolution. The scale problem in localizing feature points has been addressed by many researchers, but it is still regarded as one of the many not well understood problems in computer vision. The scale issue dates back to the early days of computer vision research on edge and curve detection [Rosenfeld and Thurston, 1971], and more recently on the introduction of the so-called *scale-space filtering* by Witkin [Witkin, 1983].

As Witkin stated in his paper [Witkin, 1983], descriptions that depend on scale can be computed in many ways. One simple way of deriving these descriptions, is the scale-space filtering approach. A family of images are obtained by convolving the original image with the Gaussian kernel $\mathcal{G}(x,y,\sigma)$ with different variances σ,

$$\mathcal{I}(x,y,\sigma) = \mathcal{I}_0(x,y) * \mathcal{G}(x,y,\sigma), \tag{3.5}$$

where

$$\mathcal{G}(x,y,\sigma) = \frac{1}{2\pi\sigma^2}\exp\left[\frac{-(x^2+y^2)}{2\sigma^2}\right], \tag{3.6}$$

where $*$ denotes convolution with respect to x and y, and $\mathcal{I}_0(x,y)$ is the original image. The smaller the value of σ, the *scale* parameter, the higher the level of detail in the image. When $\sigma = 0$, $\mathcal{I}(x,y,\sigma)$ is equal to the original image $\mathcal{I}_0(x,y)$. As σ gets larger, $\mathcal{I}(x,y,\sigma)$ becomes averaged out. In other words, σ can be used to control the amount of detail in an image.

It has been pointed out by a number of researchers, that scale-space filtering has a blurring effect equivalent to the solution of a heat conduction or diffusion equation; i.e., the inversion is an ill-posed process with a one-to-many mappings [Koenderink, 1984, Hummel *et al.*, 1987]. As Koenderink put it [Koenderink, 1984], this implies that any feature at a coarse level of resolution is required to possess a (not necessarily unique) "cause" at a finer level of resolution although the reverse need not be true.

3.6 Summary

In this chapter, the general framework for range image analysis is reviewed. Three main levels of processes are introduced; (i) image reconstruction, (ii) image segmentation, and (iii) inference of scene geometry. In addition, a selection of techniques developed for surface analysis is also briefly discussed. Although, there exist many non-surface type techniques for range image analysis, their lack of theoretical support makes them less attractive and therefore they are not included in the discussions.

This chapter also serves as the basis of our strategy for rock identification and localization, which will be discussed in the next chapter.

Chapter 4 Recovery of muckpile model from range measurements

4.1 Introduction

In this chapter, the methodologies are presented that will be employed in a strategy for rock fragments localization and identification. The paradigm is based on the framework reviewed in the previous chapter for range image analysis, which involves three main levels of processes, viz; muckpile surface reconstruction, muckpile decomposition, and muckpile modelling. One obvious reason for choosing this particular "bottom-up" approach is that very few assumptions about the scene are required, and the assumptions made are very general. This is particularly important if the strategy is to be successful when applied to an unstructured environment such as the one encountered in mining.

The problem faced is how to identify and locate highly irregular objects such as rock fragments, from the highly noisy and quantization error prone measurements such as those obtained from a rangefinder. Therefore, one must start by studying how the range measurements can be reconstructed reliably and stably.

4.2 Muckpile surface reconstruction

A novel minimization methodology for surface reconstruction was presented in Sander's doctoral thesis [Sander, 1988]. In his thesis, the emphasis was placed on the reliable recovery of surface structure from three-dimensional images acquired through magnetic resonance imaging (MRI) devices rather than from the graph surface type images like those obtained from rangefinders or standard T.V. cameras. To study Sander's methodology even further, Lagarde [Lagarde, 1990] applied the same surface reconstruction formalism to range images, and it is on Lagarde's scheme that our surface reconstruction algorithm is based.

The fundamental concept behind Sander's algorithm is *local curvature consistency* [Parent and Zucker, 1989, Sander and Zucker, 1990]; i.e., the local surface curvature at point P is assumed to be in consistency[1] with its close neighbours. Based on this assumption, a measure of the local curvature consistency is derived as the constraint to drive the minimization process[2]. In addition, the consistency constraint is subjected to the orthogonality of the augmented Darboux frame $\mathcal{D}(P)$.

The surface reconstruction process involves two steps: (i) initial surface estimation and (ii) surface estimates refinements.

4.2.1 Initial surface estimates

Initial estimates of surface properties are readily computable using a number of different local methods (§ 3.2.2). However, analytical approaches are favoured here as opposed to direct estimate approaches, trusting that the higher accuracy obtained by the initial estimates justifies the extra computational cost.

One such approach, a linear regression technique, is employed in this study for obtaining the initial estimates of $\mathcal{D}(P)$. As mentioned previously in § 2.2, a range image can be expressed as a function $z = r(i,j)$, and it can be rewritten as $z = f(u,v)$. From this function and by using standard differential geometry [do Carmo, 1976], the surface properties can be made explicit.

The initial estimation starts by fitting the local neighbourhood of P in surface S with a plane, that represents the tangent plane T_P [Ferrie and Levine, 1988]. Once T_P is found the surface normal $\vec{N}_P$ is readily available. After obtaining $\vec{N}_P$, the next step approximates the surface S with a *parabolic quadric* patch of the form:

$$\hat{f}(u,v) = a\,u^2 + b\,uv + c\,v^2. \tag{4.1}$$

The estimate $\hat{f}$ is based on samples within a particular local neighbourhood. Akin to the normal estimation, the coefficients of the second order polynomial function in Eq. 4.1 can

[1]The notion of curvature consistency is that the surface curvature at each sample point is assumed to be similar to those represented by the neighbouring local surface models; e g , in our case, the local surface model is described by a parabolic quadric patch Other surface type patches are also possible, and are discussed in [Lagarde, 1990]

[2]Typically, a minimization process is one that involves minimizing some energy functions subject to constraint satisfaction.

be computed, using a least-squares fit to the range data [Johnson and Wichern, 1982]. To simplify the analysis, a local coordinate system uvw (with the three unit vectors $\vec{e}_1, \vec{e}_2$ and $\vec{e}_3$) is defined for the estimate $\hat{f}$ with the point P at the origin and the w axis aligned with the surface normal $\vec{N}_P$.

For the curvature computation, let us consider the *second fundamental form* $\Pi_P(\vec{V}_P)$ at P in the direction pointed by the unit vector $\vec{V}_P \in T_P$, defined as:

$$\Pi_P(\vec{V}_P) = \langle -d\vec{N}_P(\vec{V}_P), \vec{V}_P \rangle. \tag{4.2}$$

A well established axiom from differential geometry states that, the value of $\Pi_P(\vec{V}_P)$ is equivalent to the normal curvature κ_{nP} of a smooth curve in surface S through point P oriented in the $\vec{V}_P$ direction [do Carmo, 1976]. $\Pi_P(\vec{V}_P)$ can be expressed in terms of the surface estimate $\hat{f}(u,v)$ as follows:

$$\Pi_P(\vec{V}_P) = l\,du^2 + 2m\,du\,dv + n\,dv^2 \tag{4.3a}$$

$$= \begin{pmatrix} du & dv \end{pmatrix} \begin{pmatrix} l & m \\ m & n \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix} \tag{4.3b}$$

$$= \left(d\vec{V}_P\right)^{\mathsf{T}} \bar{\mathsf{H}} \left(d\vec{V}_P\right) \tag{4.3c}$$

where

$$l = \frac{\hat{f}_{uu}}{\sqrt{1+\hat{f}_u^2+\hat{f}_v^2}}, \qquad m = \frac{\hat{f}_{uv}}{\sqrt{1+\hat{f}_u^2+\hat{f}_v^2}}, \qquad n = \frac{\hat{f}_{vv}}{\sqrt{1+\hat{f}_u^2+\hat{f}_v^2}}.$$

The eigenvalues of the matrix $\bar{\mathsf{H}}$ represent the maximum and minimum values that the second fundamental form can take, and are, therefore, the principal curvatures. The principal directions associated with these curvatures coincide with the eigenvectors of $\bar{\mathsf{H}}$ [do Carmo, 1976]. Note that if $\hat{f}$ is evaluated with the origin at $P = (0,0)$ in the local uv coordinates, such that $\vec{e}_1 \times \vec{e}_2 = \vec{N}_P$, and $\hat{f}_u = \hat{f}_v = 0$, then one can express $\Pi_P(\vec{V}_P)$ in terms of the polynomial coefficients by substituting Eq. 4.1 into Eq. 4.3. This results in the following Hessian matrix:

$$\bar{\mathbf{H}} = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}. \tag{4.4}$$

The principal curvatures, κ_{MP} and $\kappa_{\mathcal{M}P}$, of a smooth surface $\mathbf{S}$ at point P, are represented by the eigenvalues of $\bar{\mathbf{H}}$:

$$\kappa_{MP} = \quad a + c + \sqrt{(a-c)^2 + b^2}, \tag{4.5a}$$

$$\kappa_{\mathcal{M}P} = \quad a + c - \sqrt{(a-c)^2 + b^2}. \tag{4.5b}$$

Similarly, the two principal directions, $\vec{\theta}_{MP}$ and $\vec{\theta}_{\mathcal{M}P}$ are represented by the eigenvectors of $\bar{\mathbf{H}}$ and shown as:

$$\vec{\theta}_{MP} = \begin{pmatrix} b \\ -(a - c - \sqrt{(a-c)^2 + b^2}) \end{pmatrix}, \tag{4.6a}$$

$$\vec{\theta}_{\mathcal{M}P} = \begin{pmatrix} a - c - \sqrt{(a-c)^2 + b^2} \\ b \end{pmatrix}. \tag{4.6b}$$

Note that $\vec{\theta}_{MP}$ and $\vec{\theta}_{\mathcal{M}P}$ are explicitly expressed in the local coordinates (u, v), whereas, $\vec{M}_P$ and $\vec{\mathcal{M}}_P$ are expressed in the world coordinates (x, y, z).

4.2.2 Surface estimates refinement

This section describes the second step, the minimization stage, used to obtain a stable reconstruction of surface S. This follows the initial estimation of $\mathcal{D}(P)$ from the discrete range data. Notably noise and quantization error have always been a problem in inferring image structures both from intensity and range images. Similarly, for range images this

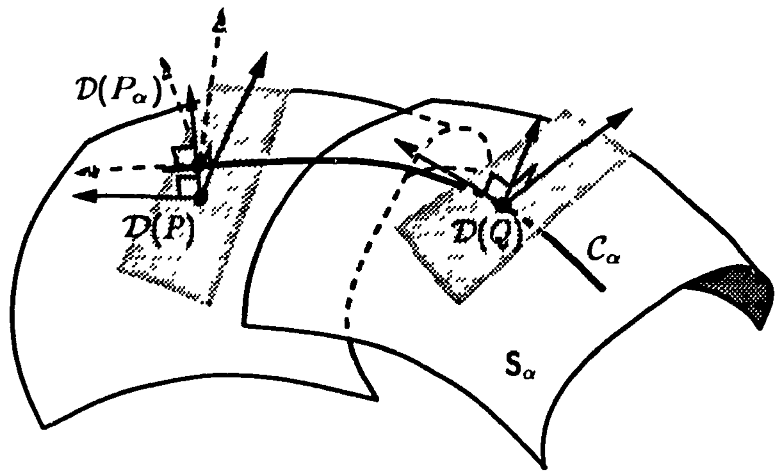


Figure 4.1: Obtaining an updated frame of $\mathcal{D}(P)$, $\mathcal{D}(P_\alpha)$, by local extrapolation. The augmented Darboux frame $\mathcal{D}(Q)$ is transported along the curve $\mathcal{C}_\alpha$ on the parabolic quadric surface patch $\mathbf{S}_\alpha$, and arrives at its neighbour P with a different orientation as $\mathcal{D}(P_\alpha)$.

problem can cause unstable or corrupted surface estimates from the local methods, especially in the estimation of the two principal directions $\vec{M}_P$ and $\vec{\mathcal{M}}_P$ [Besl and Jain, 1986a]. However, these directional properties are crucial to the inference of discontinuities and part boundaries for surface decomposition.

Iterative refinement process

From now on, Sander's surface reconstruction algorithm will be referred to as the *curvature consistency* algorithm. The algorithm operates in the following way: it refines the local surface representation at each point iteratively until it becomes consistent with its neighbouring sample points. The main reason for adopting the curvature consistency algorithm is that the complete minimization[3] is specified in terms of local surface properties, and it results in a less ambiguous (i.e., more stable) representation of the surface.

The iterative scheme can be explained with the aid of the model shown in Fig. 4.1. Since the augmented Darboux frame $\mathcal{D}(P)$ at point P is described by a parabolic quadric patch, one can consider extrapolating outwards along the quadric patch from its neigh-

[3]In theory, the minimization process can be carried out independently at each sample point, therefore some sort of parallelization in the computational process is possible.

bouring point Q to P, to extract a notion of what the surface at P should look like according to its neighbour at Q. In other words, an estimate of the frame $\mathcal{D}(P)$ at P, $\mathcal{D}(P_\alpha)$, is obtained through the assumption that the curvature is locally consistent within the neighbourhood of the two points, P and Q. By repeating this extrapolation for all neighbours of P, $\{Q_\alpha\}$, such as $\alpha = 1, \cdots, n$ and n is the number of neighbours, one can obtain a collection of augmented Darboux frames $\{\mathcal{D}(P_\alpha)\}$, each representing an estimate of $\mathcal{D}(P)$. More formally, the extrapolation along the surface is equivalent to transporting the augmented Darboux frame $\mathcal{D}(Q)$ along the surface patch $\mathbf{S}_\alpha$ to obtain $\mathcal{D}(P_\alpha)$.

Curvature consistency implementation

The functional minimization of the curvature consistency algorithm is restricted by the orthogonality constraints inherent in the frame $\mathcal{D}(P)$, which are given by [Ferrie *et al.*, 1989, Ferrie *et al.*, 1990],

$$\langle \vec{N}_P,\, \vec{N}_P \rangle = 1 \quad \langle \vec{M}_P,\, \vec{M}_P \rangle = 1 \quad \langle \vec{M}_P,\, \vec{N}_P \rangle = 0. \tag{4.7}$$

The same formulation as in [Sander, 1988, Sander and Zucker, 1990] is followed, for the augmented Darboux frame at P, $\mathcal{D}(P) = (P, \kappa_{MP}, \kappa_{\mathcal{M}P}, \vec{M}_P, \vec{\mathcal{M}}_P, \vec{N}_P)$, and its estimate $\mathcal{D}(P_\alpha) = (P_\alpha, \kappa_{MP\alpha}, \kappa_{\mathcal{M}P\alpha}, \vec{M}_{P\alpha}, \vec{\mathcal{M}}_{P\alpha}, \vec{N}_{P\alpha})$ on the surface patch $\mathbf{S}_\alpha$. Each component of the augmented Darboux frame $\mathcal{D}(P)$ to be updated is found independently, except the component $\vec{\mathcal{M}}_P$. Note that $\vec{\mathcal{M}}_P$ is given by the cross product of $\vec{M}_P$ and $\vec{N}_P$. Three *energy equations* [4] are derived for the least-squares error measure of the estimates. The first two are given as:

$$E_1 = \min_{\kappa_{MP}, \kappa_{\mathcal{M}P}} \sum_{\alpha=1}^{n} (\kappa_{MP} - \kappa_{MP\alpha})^2 + (\kappa_{\mathcal{M}P} - \kappa_{\mathcal{M}P\alpha})^2, \tag{4.8a}$$

$$E_2 = \min_{\vec{N}_P} \sum_{\alpha=1}^{n} \|\vec{N}_P - \vec{N}_{P\alpha}\|^2 \tag{4.8b}$$

[4]Energy equation – some sort of error measure function.

By applying the standard minimization procedures to solve Eq. 4.8 (see [Sander, 1988] for details), one can derive the following updating formulae first for $\vec{N}_P$,

$$\vec{N}_P^{(i+1)} = \frac{(\sum_{\alpha=1}^{n} N_{xP_\alpha}^{(i)}, \sum_{\alpha=1}^{n} N_{yP_\alpha}^{(i)}, \sum_{\alpha=1}^{n} N_{zP_\alpha}^{(i)})}{\sqrt{(\sum_{\alpha=1}^{n} N_{xP_\alpha}^{(i)})^2 + (\sum_{\alpha=1}^{n} N_{yP_\alpha}^{(i)})^2 + (\sum_{\alpha=1}^{n} N_{zP_\alpha}^{(i)})^2}}, \tag{4.9}$$

then for P, κ_{MP} and $\kappa_{\mathcal{M}P}$,

$$P^{(i+1)} = \sum_{\alpha=1}^{n} \frac{P_\alpha^{(i)}}{n}, \qquad \kappa_{MP}^{(i+1)} = \sum_{\alpha=1}^{n} \frac{\kappa_{MP\alpha}^{(i)}}{n}, \qquad \kappa_{\mathcal{M}P}^{(i+1)} = \sum_{\alpha=1}^{n} \frac{\kappa_{\mathcal{M}P\alpha}^{(i)}}{n} \tag{4.10}$$

where the superscript i indicates the iteration number.

The third energy equation requires a special treatment because there is a 180° ambiguity in the directions associated with $\vec{M}_P$ and $\vec{\mathcal{M}}_P$. To avoid this ambiguity, $\vec{M}_P$ is expressed in terms of the following tangent plane coordinates [Lagarde, 1990, Ferrie *et al.*, 1989],

$$\vec{M}_P(\psi) = \vec{e}_1 \cos\psi + \vec{e}_2 \sin\psi; \qquad \psi \in (0, 2\pi) \tag{4.11}$$

such that, $\vec{e}_1$ and $\vec{e}_2$ satisfy,

$$\vec{e}_1, \vec{e}_2 \in T_P; \tag{4.12a}$$

$$\|\vec{e}_1\| = \|\vec{e}_2\| = 1; \tag{4.12b}$$

$$\langle \vec{e}_1, \vec{e}_2 \rangle = 0. \tag{4.12c}$$

Then

$$E_3 = \min_{\psi} \sum_{\alpha=1}^{n} \left[1 - \langle \vec{M}_P(\psi), \vec{M}_{P\alpha} \rangle^2\right]. \tag{4.13}$$

The value of ψ that minimizes Eq. 4.13, is substituted into Eq. 4.11 to obtain $\vec{M}_P^{(i)}$. Similarly, the standard minimization procedures are applied to derive the following updating functional for ψ:

$$\psi^{(i+1)} = \tan^{-1}\left[\frac{(A_{22} - A_{11}) + \sqrt{(A_{11} - A_{22})^2 + 4A_{12}^2}}{2A_{12}}\right], \tag{4.14}$$

where

$$A_{ij} = \sum_{\alpha=1}^{n} \langle \vec{M}_{P\alpha}, \vec{e}_i \rangle \langle \vec{M}_{P\alpha}, \vec{e}_j \rangle; \qquad i = 1,2;\ j = 1,2. \tag{4.15}$$

A qualitative measurement of the updated surface estimates at iteration i, is given by the sum of the *residual errors* $R^{(i)}$, with

$$R^{(i)} = \sum_j R\Big(\mathcal{D}(P_j)^{(i)}, \mathcal{D}(P_{j\alpha})^{(i)}\Big) \tag{4.16a}$$

$$= \sum_j E_{1j}^{(i)} + E_{2j}^{(i)} + E_{3j}^{(i)}; \qquad P_j \in \mathbf{S}. \tag{4.16b}$$

The convergence of the minimization process can be tracked, by taking the derivative of the residual errors; i.e., the residual errors converge to a minimum as the derivative approaches zero. This can be used to control the number of iterations until the difference of two consecutive residual errors, $|R^{(i)} - R^{(i-1)}|$, is below a specified threshold. Note that $R^{(i)}$ is measured entirely by the sum of local differences computed over the surface. Experiments have shown that the algorithm converges quite rapidly, generally within 10 iterations [Lagarde, 1990].

4.3 Muckpile decomposition

The objective of surface decomposition here is to decompose the muckpile surface into regions, such that each region corresponds to the surface of an individual rock. This enables us to infer the spatial and geometrical properties of each rock independently. Ultimately, this information will be passed onto higher level processes; e.g., to derive a control strategy for the breaking mechanism [Nilsson and Lindberg, 1989].

From observations, it is found that the majority of the rock fragments resulting from blasting are largely *convex* in shape; i.e., when ignoring the small scale geometric irregularities inherent in the rock surface. Based on these findings, one can consider each rock fragment as a convex entity/part. The principle of transversality, states that when two arbitrarily shaped convex parts come into contact, they meet in a contour of concave discontinuity.

Based on the convex and compact assumption for the rock shape, and the transversality principle, one can design an algorithm to detect the features required for surface decomposition. The important features that are sought can be expressed as functions of the surface properties recovered from the curvature consistency algorithm.

4.3.1 Feature recovery

Two types of features are required for the muckpile surface decomposition; (i) *jump discontinuities* that correspond to the occluding contours of the muckpile, (ii) *negative curvature minima* that correspond to the contact boundaries between individual rocks. These feature points are best illustrated by the simple diagram shown in Fig. 4.2.

The jump discontinuities caused by occlusions of the muckpile are first recovered for the surface decomposition. Since the oversized rock fragments are located directly above the grizzly structure, if the range measurements are acquired vertically above the grizzly structure, a simple depth thresholding of the z component will suffice in identifying points that correspond to the occluding contours. However, if the measurements are taken from a different viewpoint (camera position), a transformation of the range data will be required, such that the xy plane is parallel to the grizzly structure. In the cases, where there are no overlapping or touching rocks, the surface decomposition becomes straight forward, as the regions obtained through the depth thresholding should directly correspond to each individual rock surface.

Let us consider an augmented Darboux frame $\mathcal{D}(P)$, the local surface representation at point P, with $\kappa_{\mathcal{M}P}$ and $\vec{M}_P$ representing the stable estimates of minimum curvature and its associated principal direction at P. Point P is defined as the *critical point* or *trace point* T [Ferrie *et al.*, 1989, Ferrie *et al.*, 1990] (i.e., with a negative local minimum

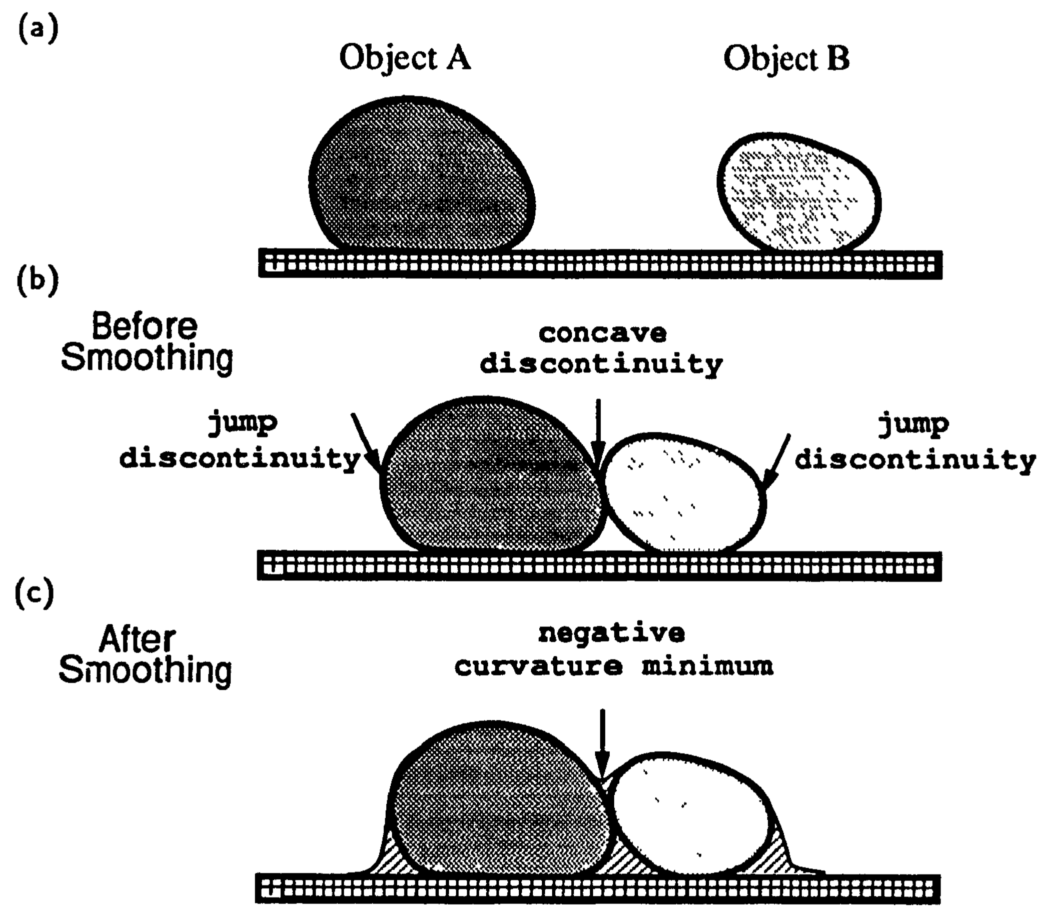


Figure 4.2: Feature recovery for part decomposition. (a) Showing two convex and complex objects A and B. (b) Before smoothing: the concave discontinuity corresponds to the contact boundary, and the jump discontinuity corresponds to the occluding contour. (c) After smoothing: the concave discontinuity is transformed into a negative curvature minimum.

curvature), if and only if the following is true,

$$\kappa'_{nP}\Big|_{\mathcal{M}P} = 0 \qquad \text{AND} \qquad \kappa_{\mathcal{M}P} < 0 \tag{4.17}$$

where $\kappa'_{nP}|_{\mathcal{M}P}$ is the directional derivative of the normal curvature κ_{nP} in the direction specified by the vector $\vec{\mathcal{M}}_P$.

Since we only have discrete measurements, $\kappa'_{nP}|_{\mathcal{M}P}$ can be approximated by a difference equation. To simplify the computation, a simple comparison between neighbouring curvatures is sufficient for identifying the negative local curvature minimum. Point P is bound to be a local curvature minimum, if

$$\kappa_{\mathcal{M}P} < \kappa_{nP^+}\Big|_{\mathcal{M}P} \qquad \text{AND} \qquad \kappa_{\mathcal{M}P} < \kappa_{nP^-}\Big|_{\mathcal{M}P}. \tag{4.18}$$

If P is a point that satisfies Eq. 4.18 and its minimum curvature $\kappa_{\mathcal{M}}(P)$ is less than 0, then P has a local negative curvature minimum; i.e., P is a trace point. P^+ and P^- are the closest sample points from P in the directions of $\vec{\mathcal{M}}_P$ and $-\vec{\mathcal{M}}_P$, and $\kappa_{nP^+}\,|_{\mathcal{M}P}$ and $\kappa_{nP^-}\,|_{\mathcal{M}P}$ are the normal curvatures at P^+ and P^- in the $\vec{\mathcal{M}}_P$ direction.

For various practical reasons, it is very difficult to apply the transversality principle directly on range images. First we face the problem of obtaining smooth surface data from the sensor. Inevitably there will be noise added to the data during the acquisition process, even with the measurements taken from a scene that is fully occupied by objects with no dents and depressions. Even if the data are assumed to be perfect; i.e., the range measurements match the exact physical appearances of the objects, there is still a problem in recovering the surface properties from the discrete data, which will provide reliable surface estimates. The curvature consistency algorithm is developed precisely for this purpose; i.e., provides a stable surface representation even with the presence of additive noise in the data.

Second, and more important is the problem of tuning into a correct range of scales. We are living in a world of objects consisting of different levels of structure. Humans, for example, have the extraordinary ability to tune into the right level of scale depending upon an *a priori* knowledge of the scene and/or the types of the objects. Unfortunately, the process of inferring different levels of scales remains not well understood. Not surprisingly, this scale problem turns out to be a very significant one for the muckpile surface decomposition. Rock surfaces are very textured, and contain many small irregularities due to the litho-structure of the deposits. Direct inference of rock surface properties at the very fine scales, would result in a problem of localizing the "true" partitioning boundaries. More details of this problem will be given in the next chapter.

4.3.2 Partitioning contour aggregation

By referring back to the transversality principle, objects (muckpiles) are said to be concatenated or adjoined together by a number of convex and compact parts (rocks). If the contours of negative minima are recovered from the stable surface estimates, then the trace points $\{T_q\}$ should be equivalent to the contact boundaries between rocks.

However, in practice the trace points corresponding to the contact boundaries do not always form enclosed partitioning boundaries required for the surface decomposition. Different features must be recovered and combined together in a cooperative manner, such that the image can be decomposed into meaningful regions for the later visual processing stages. As mentioned in the previous section, two types of features associated with the partitioning boundaries have to be recovered. A reliable way of combining this information together must be derived in order to perform the surface decomposition. Experience has shown that the contact boundaries between rocks are most likely to be terminated along the occluding contours of the muckpile, therefore one can use this as a constraint to derive a partitioning boundary detection scheme for identifying the partitioning boundary for each rock fragment.

Given a set of trace points $\{T_q\}$ (negative curvature minima) corresponding to the contact boundaries between different rocks, and a set of points $\{J_g\}$ (jump discontinuities) corresponding to occluding contours of the muckpile, the method for connecting the contact boundaries with the occluding contours is the following. First *end points* from the recovered trace points $\{T_q\}$ are defined.

> **End point:** a trace point with only *one* connectivity; contextually this implies trace point that has one and only one neighbouring point directly intact to one of its eight directions, and that neighbouring point is also a trace point.

After the end points are identified from the contact boundaries, the contour connection is constrained between the end points and the occluding contours; i.e., the end points can be extended to intersect at the closest distance on the occluding contours. Because of the high density of these feature points, straight line interpolation should suffice in locating the enclosed contours for each of the individual rocks. Regions can then be identified using a standard clustering algorithm which labels connected points as comprising a unique region [Ballard and Brown, 1982].

4.4 Muckpile modelling

Having partitioned the surface acquired by the rangefinder into regions corresponding to different rocks, the next task is to infer the three-dimensional shape of each rock. Our approach is to model each rock using a volumetric primitive with sufficient degrees of freedom to account for the expected range of shapes. In particular, the superquadric model is chosen for our rock modelling. The reasons that the superquadric is chosen for the modelling are, (i) its capability of representing a large range of shapes, (ii) the fact that spatial and geometric information (e.g., position, size, shape, centre of mass, surface orientation, etc.) about the objects can be easily recovered from the model.

4.4.1 Fitting of superquadric model

Given a set of partitioned surfaces $\mathbf{S}_l$ from the surface $\mathbf{S}$, such that $\mathbf{S} = \bigcup_l \mathbf{S}_l$ and $l = 1, \ldots, n$, where n is the total number of parts in the surface $\mathbf{S}$.

One can obtain a set of volumetric primitives $\mathcal{L}$ from the superquadric model, by varying $\mathbf{a} = (\epsilon_1, \epsilon_2, a_x, a_y, a_z)$ in the parameter space. Now, the task is to determine the parameter of a volumetric element $\mathcal{V}_l$, $\mathcal{V}_l \in \mathcal{L}$ that best characterizes $\mathbf{S}_l$ minimizing the expression,

$$\Big|\mathcal{V}_l(x, y, z) - \mathbf{S}_l(x, y, z)\Big|. \tag{4.19}$$

The fitting algorithm for the superquadric model is very similar to those mentioned in [Gross and Boult, 1988]. It makes use of a nonlinear minimization technique to recover the required model parameters, starting with an initial estimate. The fitting procedure is an iterative one, akin to [Solina and Bajcsy, 1990], and begins with a good ellipsoidal ($\epsilon_1 = \epsilon_2 = 1$) approximation that provides an initial estimate of the rotation and translation parameters for the part to be modelled. Experience has shown that the initialization of the relative shape parameters, ϵ_1 and ϵ_2 are not very critical to the final model $\mathcal{V}_l$ After the ellipsoidal approximation is achieved, the translation parameter is initialized by locating the centroid of the points in the surface $\mathbf{S}_l$. The next step is to find the initial rotation parameters by aligning the axes of the ellipsoid along the principal moments of inertia of $\mathbf{S}_l$ about the centroid.

A nonlinear least-squares technique is used for the minimization — the Levenberg-Marquardt method [Press *et al.*, 1988], to minimize the error of fit between a superquadric surface $\mathcal{V}_l$ and a surface patch of range data $\mathbf{S}_l$. The fitting procedure is repeated with a new set of parameters **a** at each iteration, until the sum of the least-squares error measure stops converging. The final set of parameters **a**, represents the best fitted model for the corresponding part that is to be described.

One might argue that the superquadric is an overly complex model for the purpose of the rockbreaker problem. However, in addition to breakage, the operation also requires manipulation of the rock mass for which the pose information provided by the model is indeed useful. The two additional shape parameters, which are potentially useful in the inference of other physical properties, come at a modest increase in computational expense. As a general characterization of rock shape, the superquadric is clearly limited However, the general impression is that for the problem typified by the rockbreaker, the shape approximation by the superquadric model in practice is adequate.

4.5 Discussion

As mentioned in Chapter 1, the basic strategy for muckpile identification and localization is very similar to that proposed in [Ikeuchi and Hebert, 1990] and [Choi *et al.*, 1990]. Both the paradigms involve similar processes; i.e., range data acquisition, surface reconstruction, surface segmentation and fitting of superquadric models. However, the computational approaches used to perform the surface reconstruction and the segmentation differ significantly. The approach used here is based on curvature analysis and the inference of 3-D curves [Sander and Zucker, 1990, Ferrie *et al.*, 1989]. Whereas Ikeuchi and Hebert, and Choi *et al.* applied deformable models (e.g , "snakes" and "deformable surface") similar to those proposed in [Terzopoulos *et al.*, 1987, Kass *et al.*, 1988]. Aside from the theoretical issues, there are practical considerations for the choice of approach. One consideration is that, the analysis involves a minimum number of "hidden" parameters. At present, there is a single scale parameter (see next chapter). Another consideration is the active contours or "snakes". Examples have shown that they have worked well for relatively smooth surfaces [Ferrie *et al.*, 1989], but for the highly textured surfaces associated with rock fragments, they have proven to be more difficult (especially in estimating the process parameters).

On the other hand, the surface segmentation scheme is somewhat similar to [Ponce and Brady, 1987, Fan *et al.*, 1987, Fan, 1990], in detecting features for partitioning boundaries. However, the present method is more stable for two reasons. First is that the surface representation obtained through the curvature consistency algorithm is stable, which is critical to the part decomposition. Second is that the transversality principle is the natural basis for identifying the partitioning boundaries.

However, the basic strategy here is very different from the one proposed in [Lim, 1990] for rock recognition, with graph models to denote geometric salient features. This is understandable, since the objective in the latter case is quite different.

4.6 Summary

In this chapter, the methodologies employed in our strategy for the rock identification and localization are presented. Three levels of processes are discussed.

First, the stable representation of the surface is recovered from the range measurements by applying a surface reconstruction process. The process involves two processing steps: (i) initial estimates of surface properties, and (ii) surface estimates refinement. The resulting stable surface estimates, make the surface features required for surface decomposition explicit.

In the second level of process, the transversality principle is presented as the basis of our surface decomposition strategy, for partitioning muckpile surface into regions corresponding to different rocks. Two type of features are considered in our surface segmentation scheme, (i) jump discontinuities corresponding to occlusion boundaries between muckpiles, and (ii) negative local curvature minima corresponding to contact boundaries between rocks.

The final processing step is to find a volumetric model that best characterizes the 3-D shape of each partitioned surface. Each individual rock is represented by a parametric primitive, and this final representation can be used by other higher levels of processes, such as for controlling the breaking mechanism of a rockbreaker.

The purposed strategy is fairly independent of the origin of data; i.e., we are not committing ourselves to any particular type of range sensor. Since hardware development is advancing in a much faster pace with respect to the progress in algorithm design, although this is largely a personal intuition, we are likely to see faster and more accurate range sensors becoming available in the near future.

Chapter 5 Results

5.1 Laboratory setup

This chapter describes the results obtained by using the previously described methodology in a laboratory environment. A scaled down version (50 cm $\times$ 40 cm) of a grizzly was built in our laboratory with a grid size of 10 cm $\times$ 10 cm. Nickel and copper ore samples with diameters[1] ranging from 4 to 8 cm were used for the experiments. The samples were relatively clean, i.e., free from accumulated mud and dirt. Images of rock piles were captured by the NRCC/McGill laser scanner. The scanner was mounted on the end-effector of a PUMA 560 industrial robot to accommodate different viewing positions. Fig 5 1 shows the laboratory setup.

A highly flexible software tool has been developed on a Silicon Graphics (Personal Iris) workstation, which is directly linked to a Local Area Network; the networked machines include a wide range of computers with different architectures, various types of imaging acquisition devices and robotic controllers The complete operation of image acquisition from the laser scanner, the position of the robot, the transfer of the data over the networ and the rendering of the range data is entirely abstracted from the user level. The advantage of having such a software system is that the user can concentrate on a particular problem, without worrying about the technical details of hardware implementation.

5.2 Processing sequences

The data acquisition time for a full size (256 $\times$ 256) range image was approximately 15 seconds A typical set of raw range data obtained from muckpile samples located on top of the grizzly model, is shown in Fig. 5.2.

[1]Typically, rock fragment size is measured by a sphere that has an equivalent volume

(a)

(b)

Figure 5.1: Laboratory setup for the rockbreaker pilot study. (a) An overall view of the PUMA 560 robot and the NRCC/McGill laser scanner. (b) A close-up view of the grizzly model and ore samples.

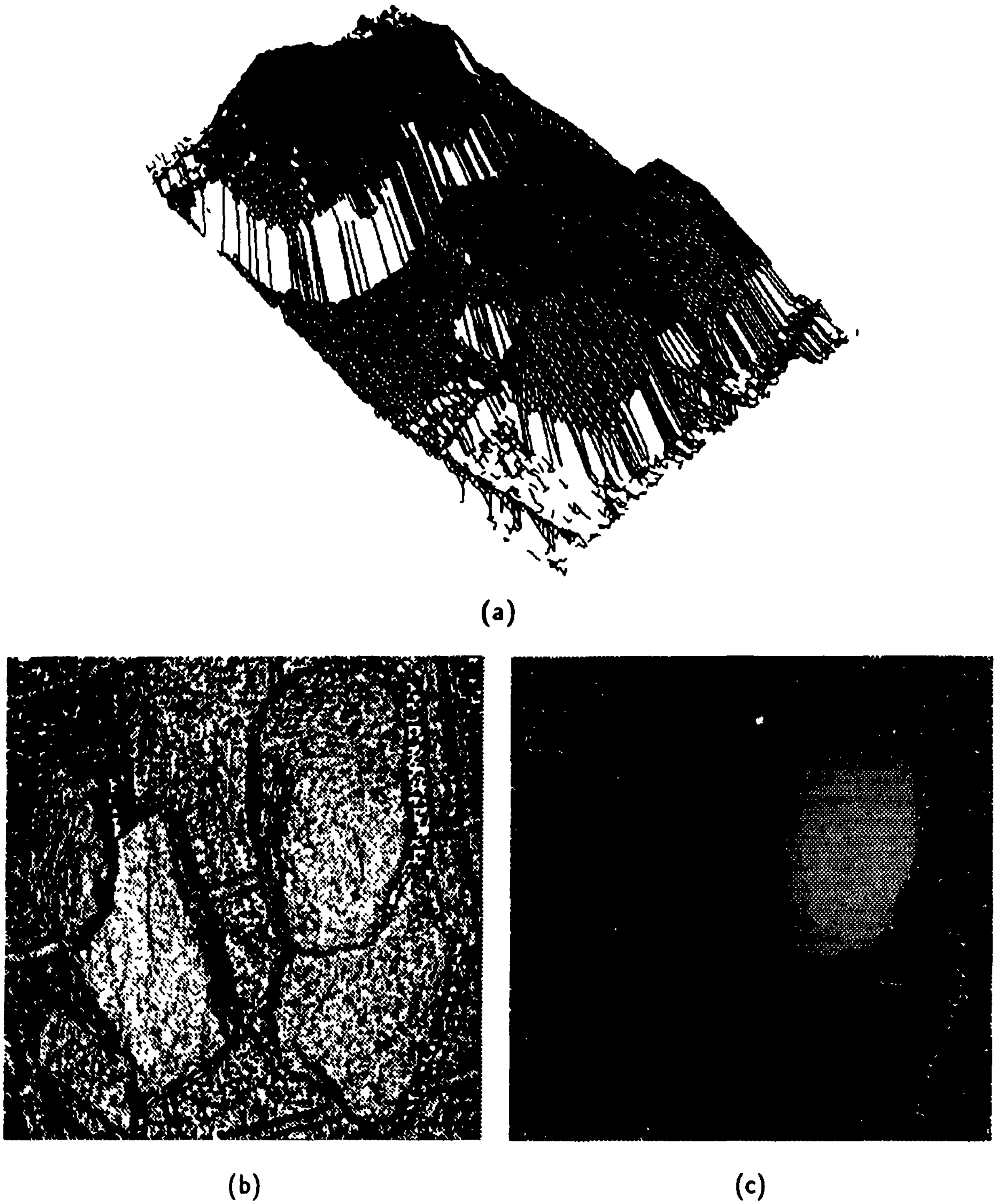


Figure 5.2: Raw range measurements of muckpiles. Two piles are shown here, one consists of two and the other of three rock fragments respectively. The muckpiles were placed on top of the grizzly model. (a) 3-D plot of the range data. (b) Shaded image. (c) Grey-scale image (darker pixels represent points further away from the scanner).

5.2.1 Figure and ground separation

Because of the chosen view-point[2] (directly above the grizzly model) and the calibrated field of view of the laser scanner, background/object separation after data acquisition can be performed by a simple depth thresholding based upon an *a priori* knowledge of the height of the grizzly. The result is a binary image, which shows muckpile separated from the grizzly model (Fig. 5.3)

(a) (b)

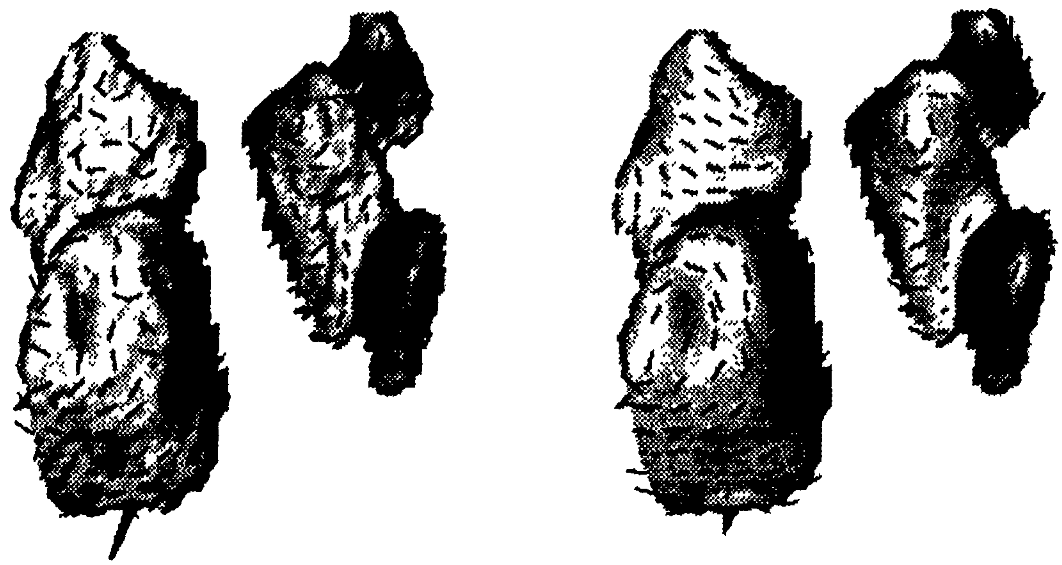
Figure 5.3: Figure/ground separation (a) Range image of rock pile after figure/ground separation with grey-scale representing the range. (b) The associated binary image.

5.2.2 Rock surface reconstruction

The next processing stage is the reconstruction of the muckpile surface from the range data. First, we find initial estimates of the surface properties (i e., surface normal, principal curvatures and principal directions at each sample point) using local methods Then we refine the initial estimates iteratively, using the curvature consistency algorithm. In

[2]Transformation is required for other camera position, such that the xy plane is always parallel to the grizzly model.

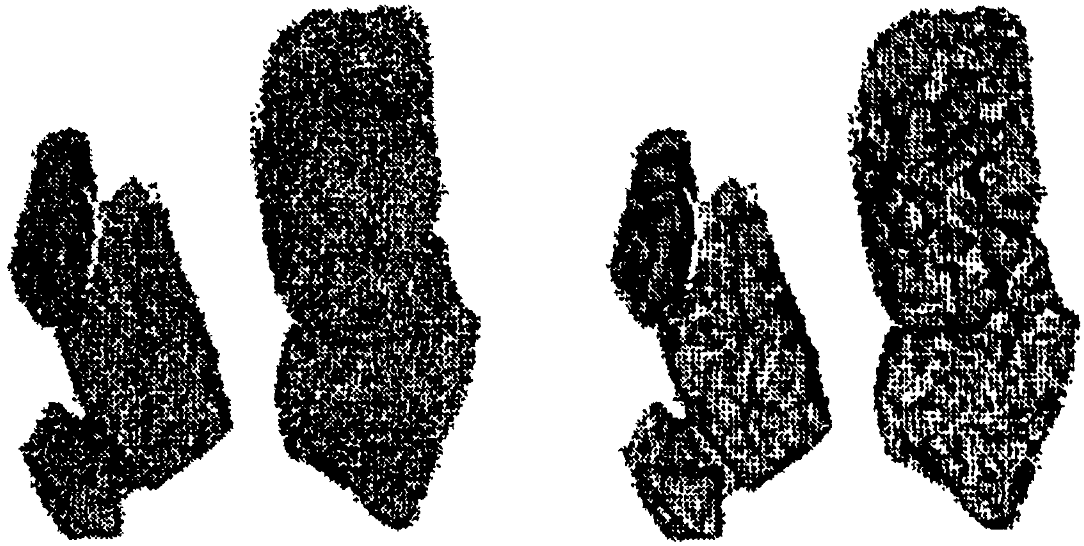
general, the solution converges rapidly after five iterations. Fig. 5.4 shows a comparison of the principal directions (with needles representing the principal directions). One can easily note that the principal directions obtained after applying five iterations of curvature consistency are far more stable when compared to the initial estimates obtained through local methods



(a) (b)

Figure 5.4: Needle map of principal directions $\vec{\mathcal{M}}_P$. (a) Initial estimates using local method. (b) Refinements after applying 5 iterations of curvature consistency algorithm

To compare the curvatures of the muckpile surface, the $K_P H_P$ curvature-sign map that characterizes the muckpile surface based on the eight surface primitives defined in § 3 3.1 is computed (Fig 5 5) Although, the surface representation obtained from the minimization process is much more stable than the one obtained from the initial estimates, the problem of inferring the surface boundaries remains (see Fig 5.5b) Ideally, for the surface decomposition, it is desirable to obtain a surface representation that describes the general geometric structure of the muckpile, rather than a detailed representation which includes the unwanted and highly textured rock surface features.



(a) (b)

Figure 5.5: Comparison of $K_P H_P$ curvature-sign map (a) Initial estimates using local method. (b) Refinements after applying 5 iterations of curvature consistency algorithm.

Scale-space filtering

As mentioned in the previous chapters, to infer the partitioning boundaries for the part decomposition, only the overall structure of the muckpile at certain scales are significant To obtain the surface representation of the muckpile at a particular scale, Gaussian filtering is first applied to the original range data before the surface reconstruction A set of stable surface descriptions at different scales can be obtained by varying the scale parameter σ (Fig. 5.6).

5.2.3 Rock decomposition

Similarly, the surface features (negative curvature minima) at different scales for the partitioning boundary detection can be recovered. These critical points should correspond to the contact boundaries between rocks if the appropriate scale is chosen. Fig 5 7 shows critical points recovered at different scales after applying five iterations of the curvature consistency algorithm. By combining these critical points together with the

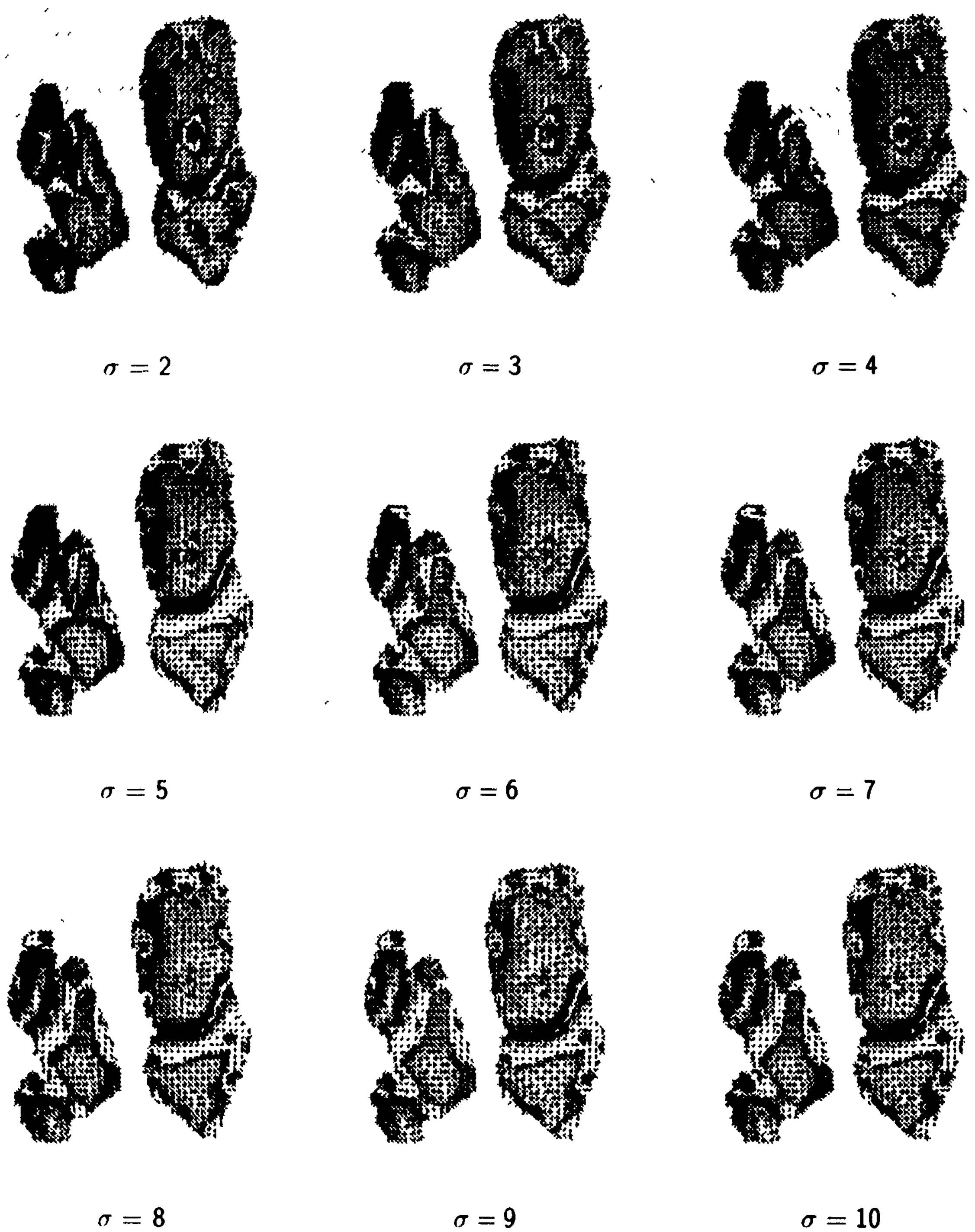


Figure 5.6: K_pH_p curvature-sign map at different scales.

occluding contours obtained from the feature/ground separation, the enclosed contours each isolating an individual rock can be recovered. Because of high density of these feature points, and the availability of *a priori* knowledge of how the contact boundaries should end, a simple straight line interpolation between end points and occluding boundaries, in practice is often sufficient for finding the rock boundaries However, if the density of these feature points decreases, it might not clear how the interpolation should proceed The resulting enclosed regions are then labelled using a standard clustering algorithm A set of segmentation results based on the critical points in Fig 5.7 is shown in Fig 5.8.

5.2.4 Fitting of superquadric model

The final processing stage for the rock localization and identification problem is to recover the spatial properties for each individual rock. These properties include the centre of gravity, the surface orientation, the approximated shape and size The wire frames corresponding to the final models for each rock are shown in Fig. 5.9b To get a more qualitative appreciation of these results, the wire frames are rendered as shaded images and plotted against the original set of range data (Fig. 5.9).

5.2.5 Processing time

Table 5.1 shows the approximated time required for each processing stage on a Sun Sparcstation 1.

Surface Reconstruction	20 mins.
Surface Decomposition	30 secs
Fitting of Superquadric Model	5 mins

Table 5.1: Approximated time required for each processing stage. surface reconstruction, surface decomposition, and fitting of superquadric model.

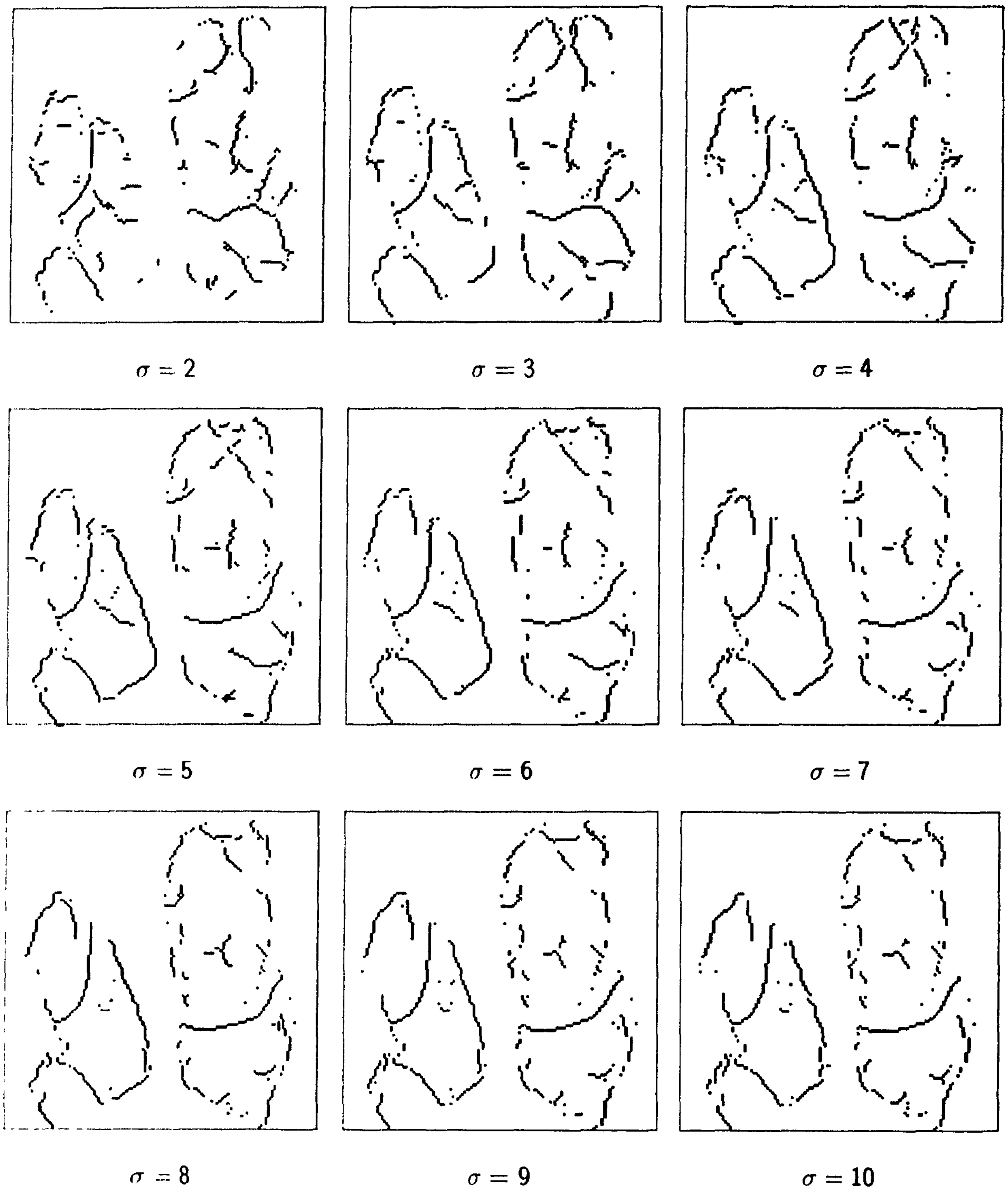


Figure 5.7: Surface feature points at different scales.

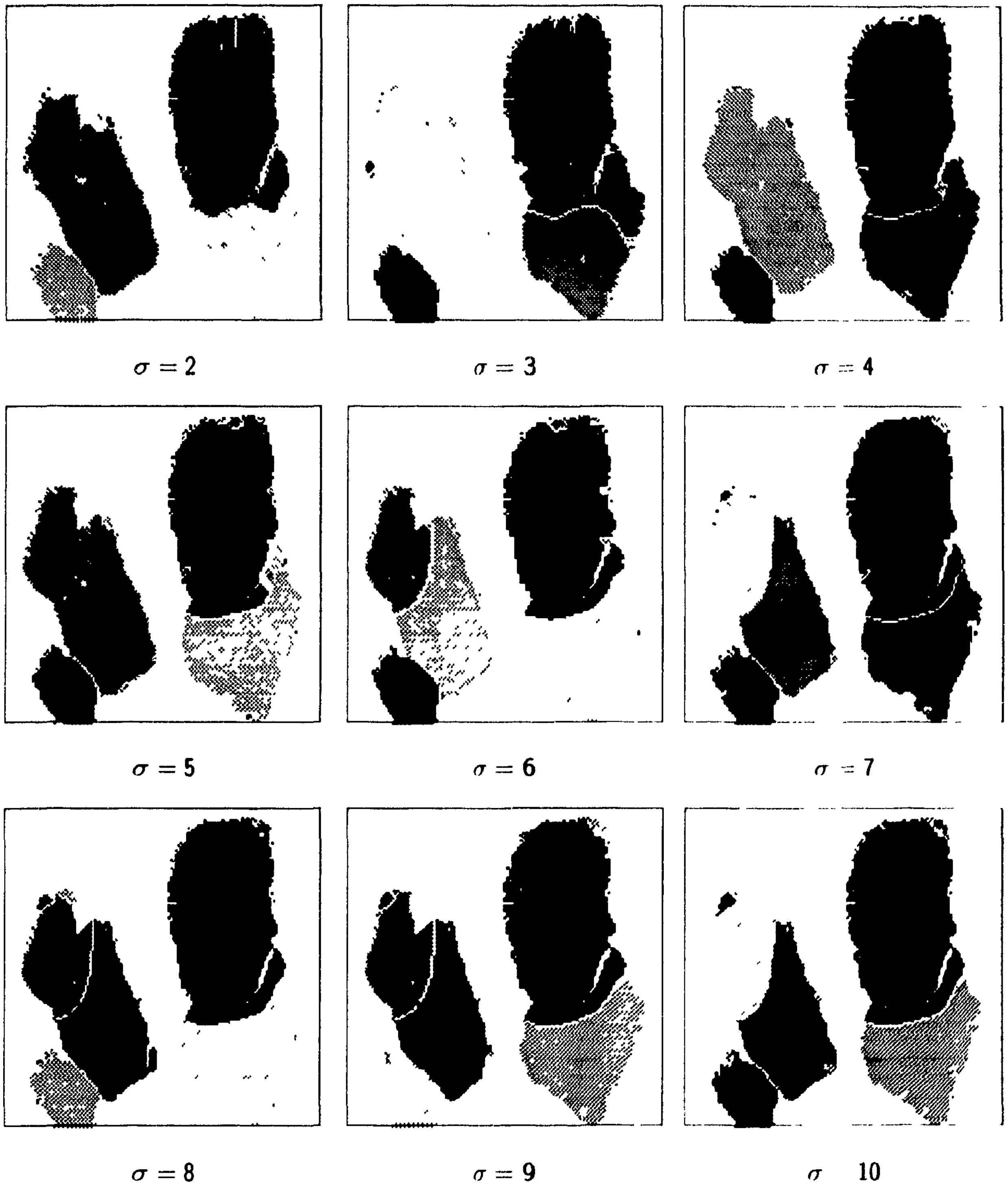
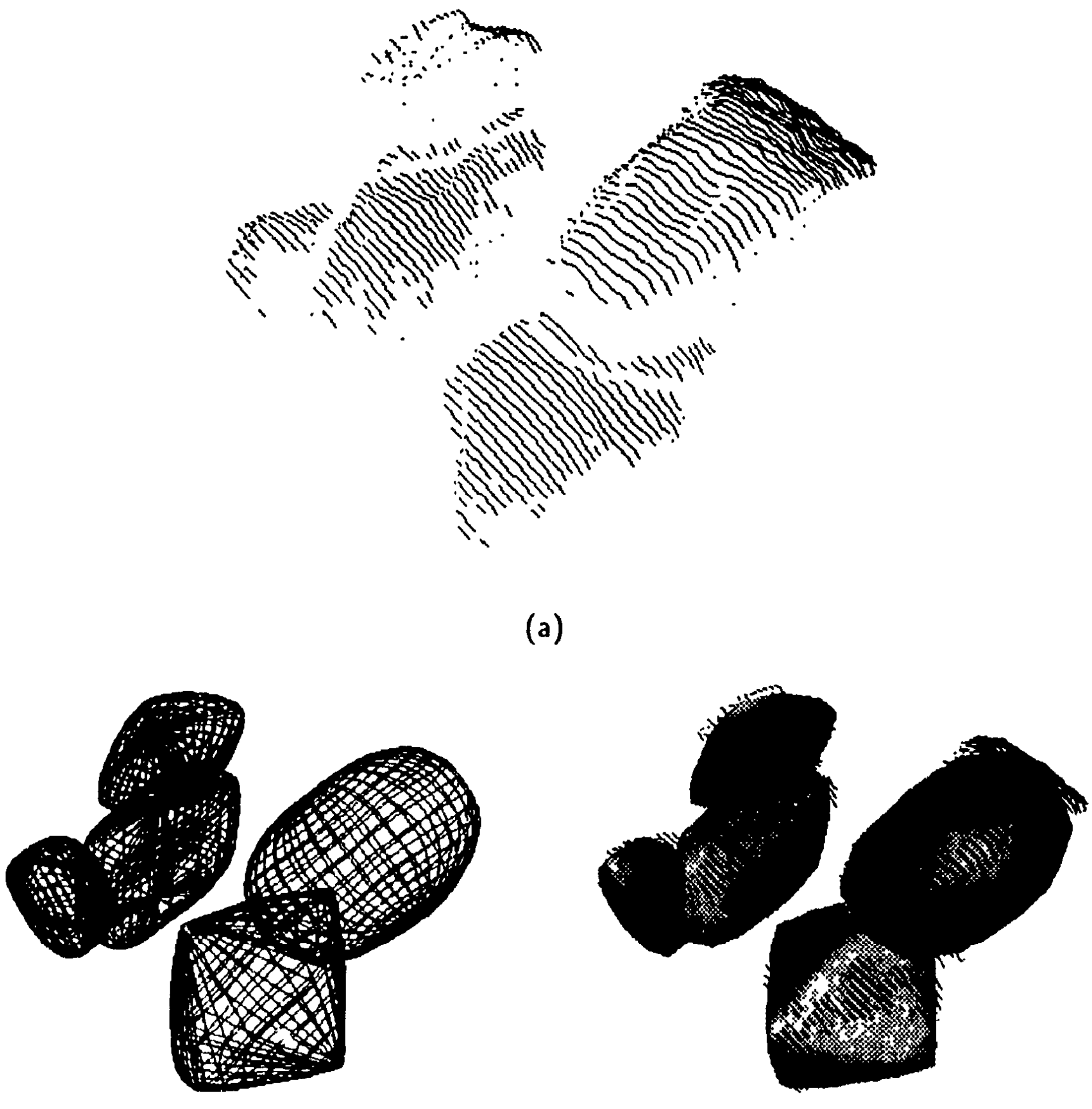


Figure 5.8: Surface decomposition at different scales



(a)

(b) (c)

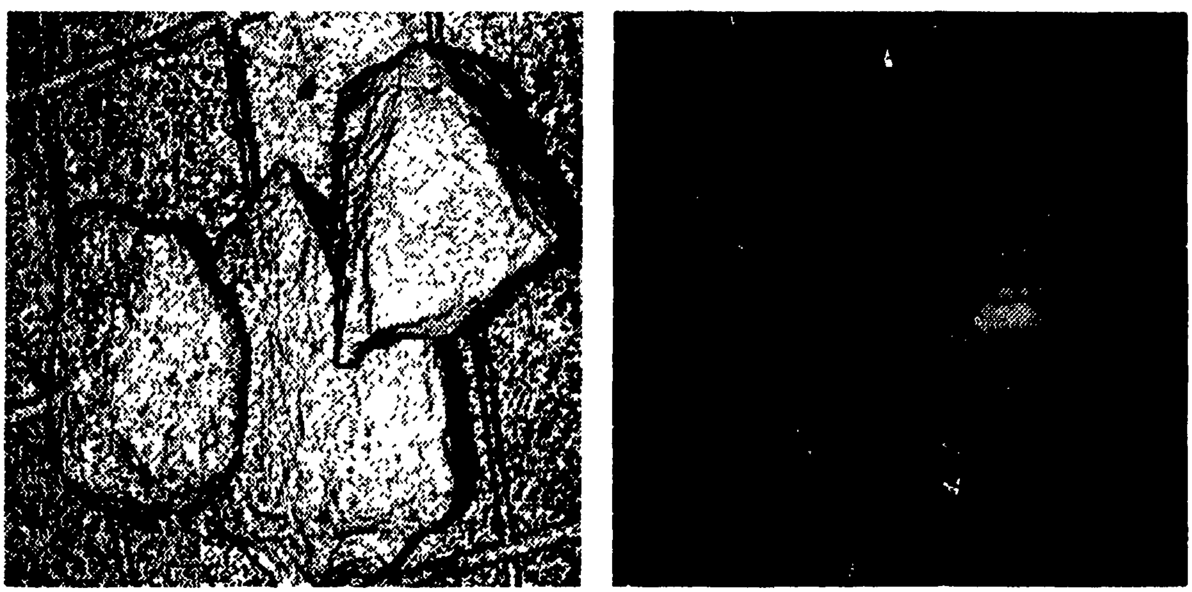
Figure 5.9: (a) Original range data partitioned according to the region map $\sigma = 6$ in Fig 5.8. (b) Superquadric (wire frame) model fitting of the rock-pile. (c) Shaded image of the superquadric model.

5.3 Case studies

Five case studies are presented in the following section, to evaluate the strategy proposed. The computation of surface reconstruction and part decomposition was performed on a Sun Sparcstation 1+. For the superquadric modelling, each rock was modelled individually The rendering of the muckpile models was done on a Silicon Graphics 4D/35 (Personal Iris) workstation.

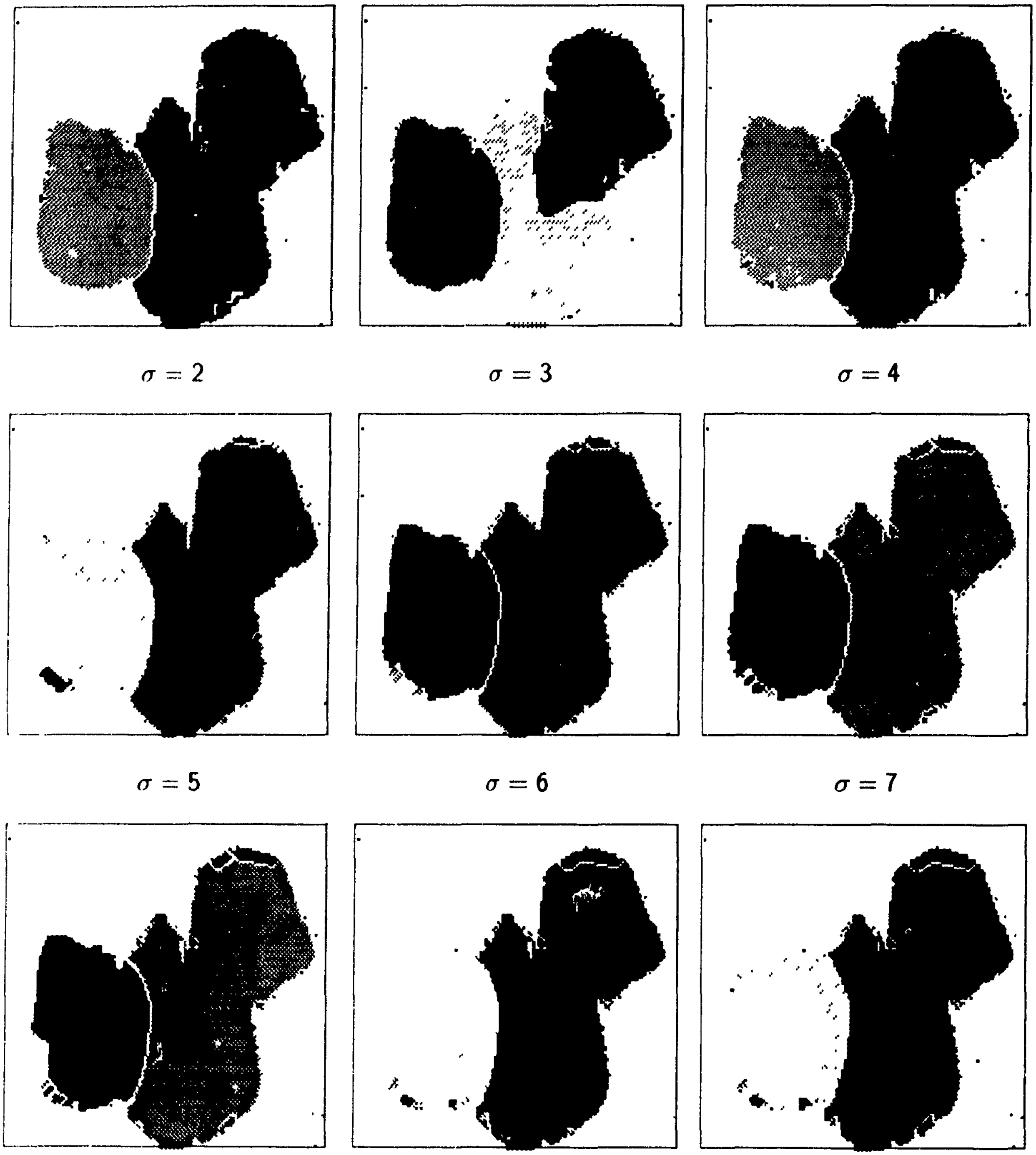
Example one

Fig. 5.10 shows a shaded image and a range image (intensity represents range) of a muckpile consisting of three rock fragments. Using the same procedures mentioned in this chapter, the muckpile surface is decomposed at different scales (Fig 5 11) The final muckpile model is shown in Fig 5 12.



(a) (b)

Figure 5.10: Raw range measurements of muckpiles. Three rock fragments are shown here, with two rocks overlapping a bigger rock The muckpiles were placed on top of the grizzly model. (a) Shaded image. (b) Grey-scale image (darker pixels represent points further away from the scanner).

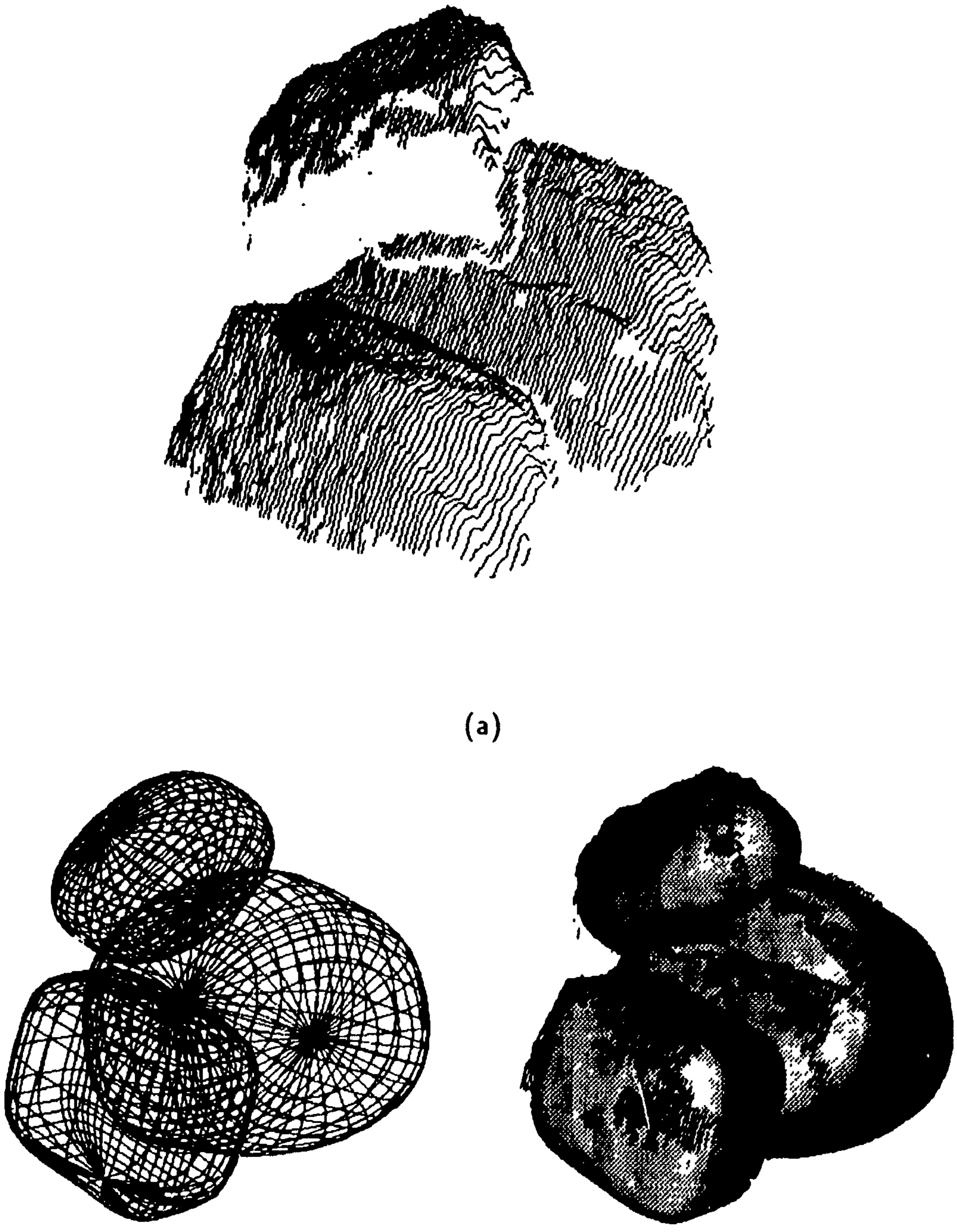


$\sigma = 2$ $\sigma = 3$ $\sigma = 4$

$\sigma = 5$ $\sigma = 6$ $\sigma = 7$

$\sigma = 8$ $\sigma = 9$ $\sigma = 10$

Figure 5.11: Surface decomposition at different scales.



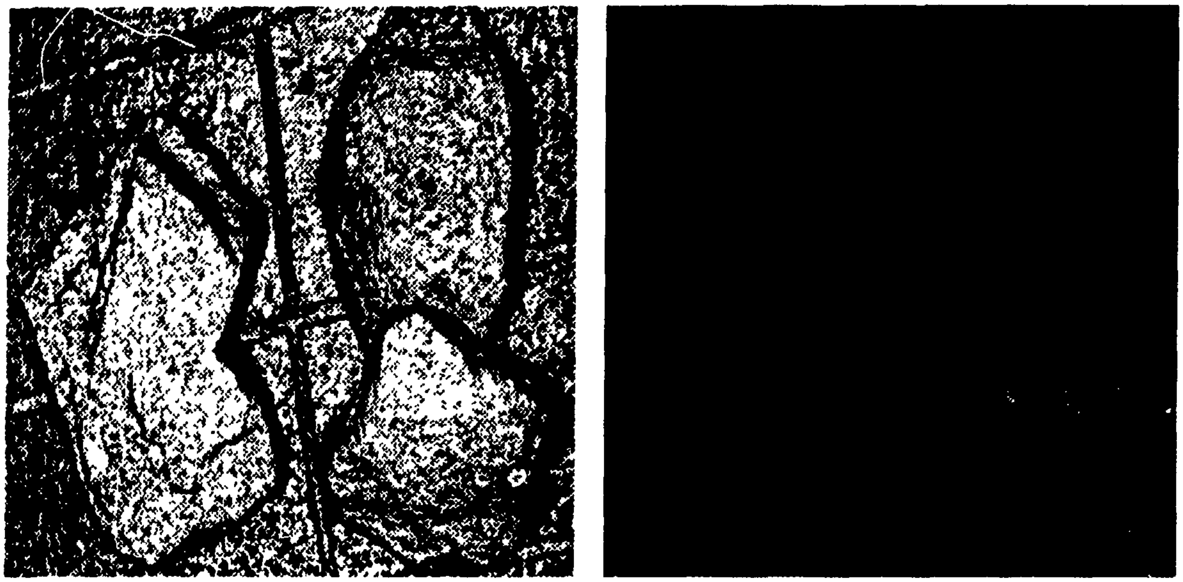
(a)

(b) (c)

Figure 5.12: (a) Original range data partitioned according to the region map Fig. 5.11. (b) Superquadric (wire frame) model fitting of the rock-pile. (c) Shaded image of the superquadric model.

Example two

The second example shows three rock samples (one isolated and two overlapping rocks) in Fig. 5 13 The surface decomposition at different scales is shown in Fig. 5.14. The superquadrics corresponding to this muckpile model is shown in Fig. 5.15.



(a) (b)

Figure 5.13: Raw range measurements of muckpile. Three rock samples are shown here, one isolated and two overlapping fragments. (a) Shaded image. (b) Grey-scale image (darker pixels represent points further away from the scanner)

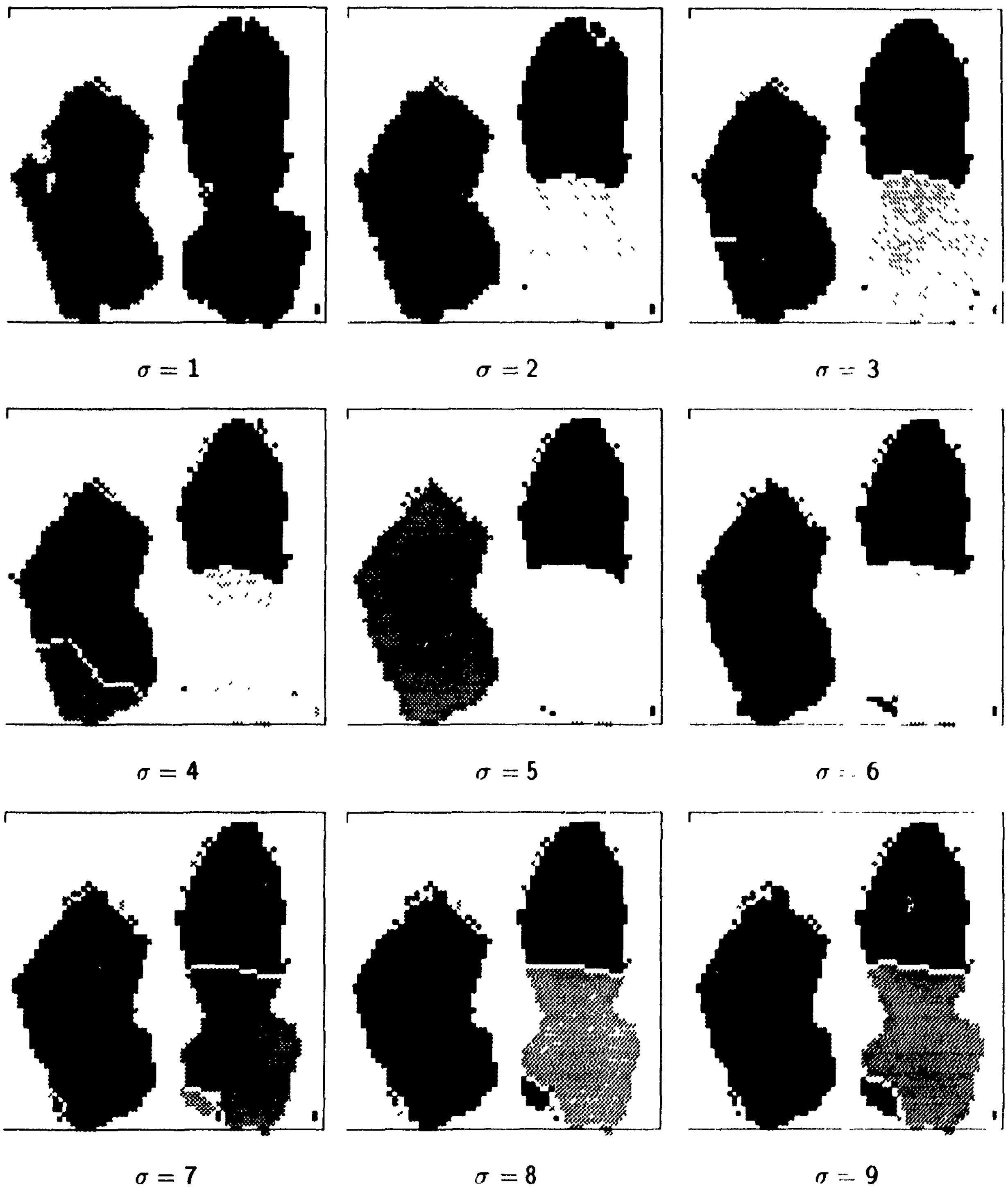


Figure 5.14: Surface decomposition at different scales.

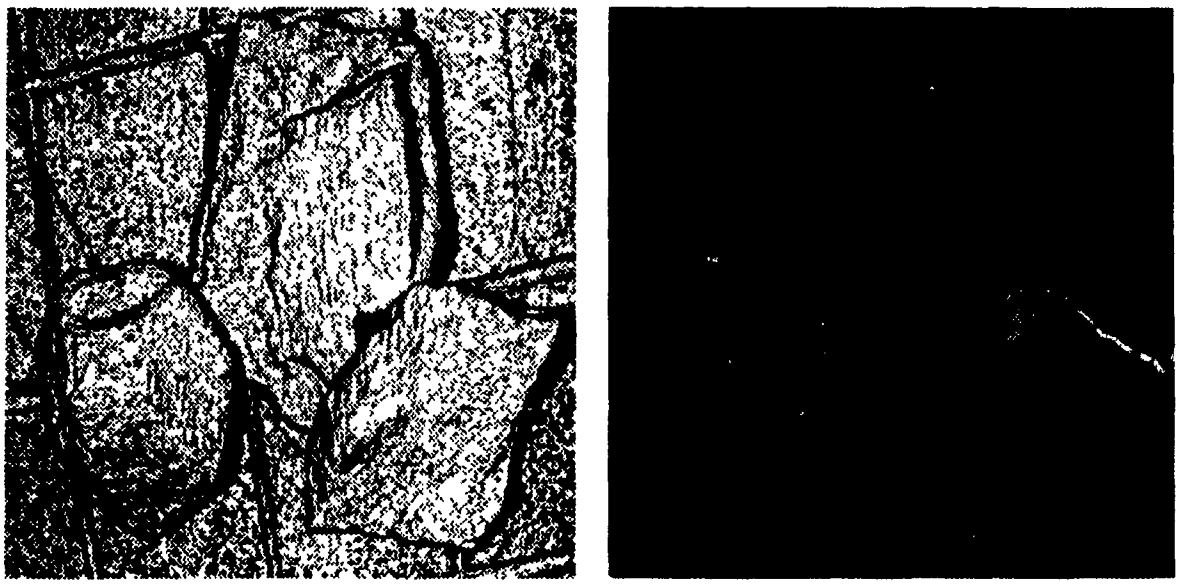
(a)

(b) (c)

Figure 5.15: (a) Original range data partitioned according to the region map Fig. 5.14. (b) Superquadric (wire frame) model fitting of the rock-pile. (c) Shaded image of the superquadric model.

Example three

Fig. 5.16 shows another example of three overlapping rock samples. In particular, the rock fragments were arranged in such a way, that they only slightly overlap each other; i.e., the depth discontinuities along the contacting boundaries are relatively small The surface decomposition worked well even under these situations (see Fig. 5 17) The final muckpile model is shown in Fig. 5.18.



(a) (b)

Figure 5.16: Raw range measurements of muckpile. Three rocks are shown here, with two small rock fragments slightly overlapping a bigger fragment. (a) Shaded image. (b) Grey-scale image (darker pixels represent points further away from the scanner).

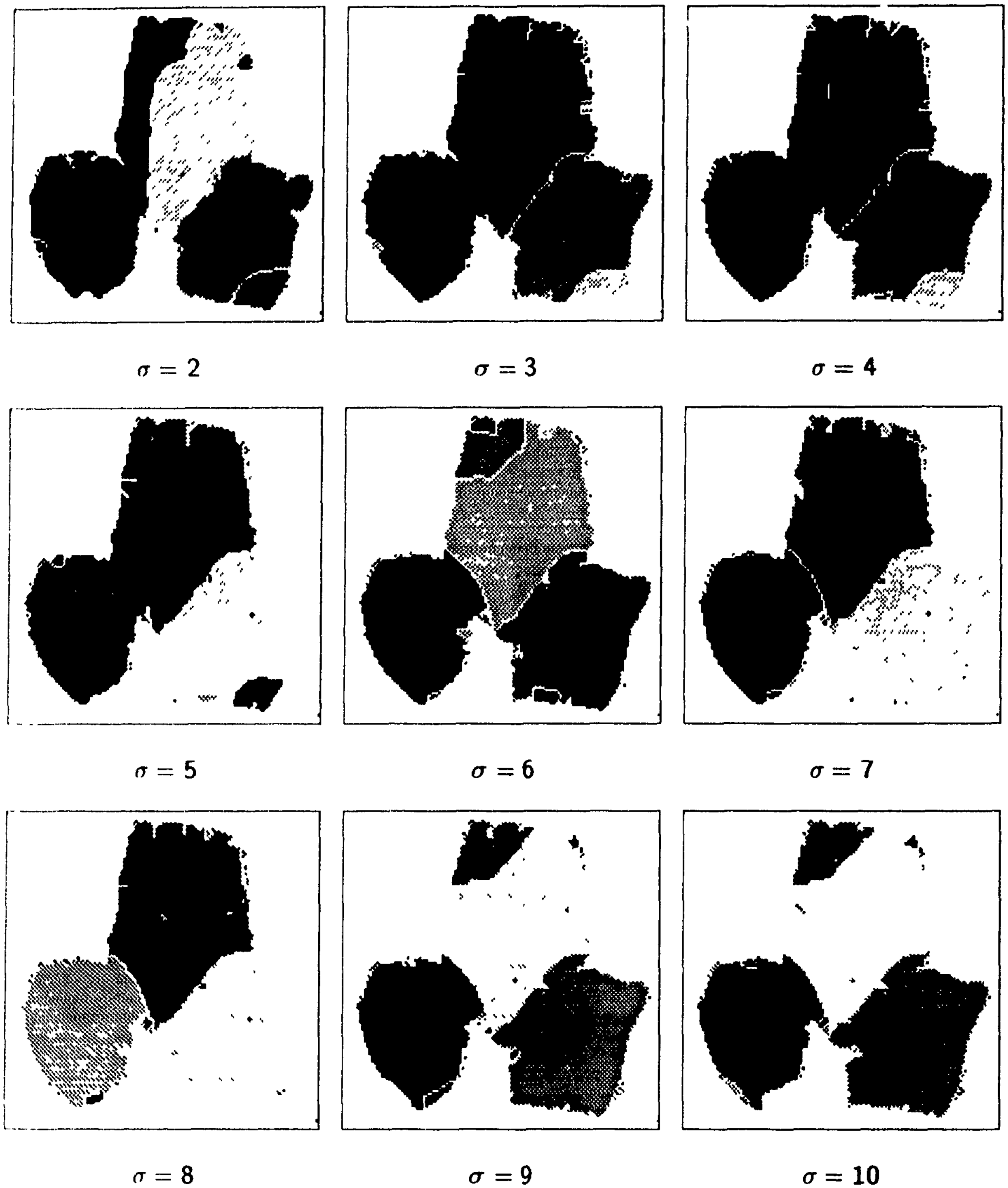
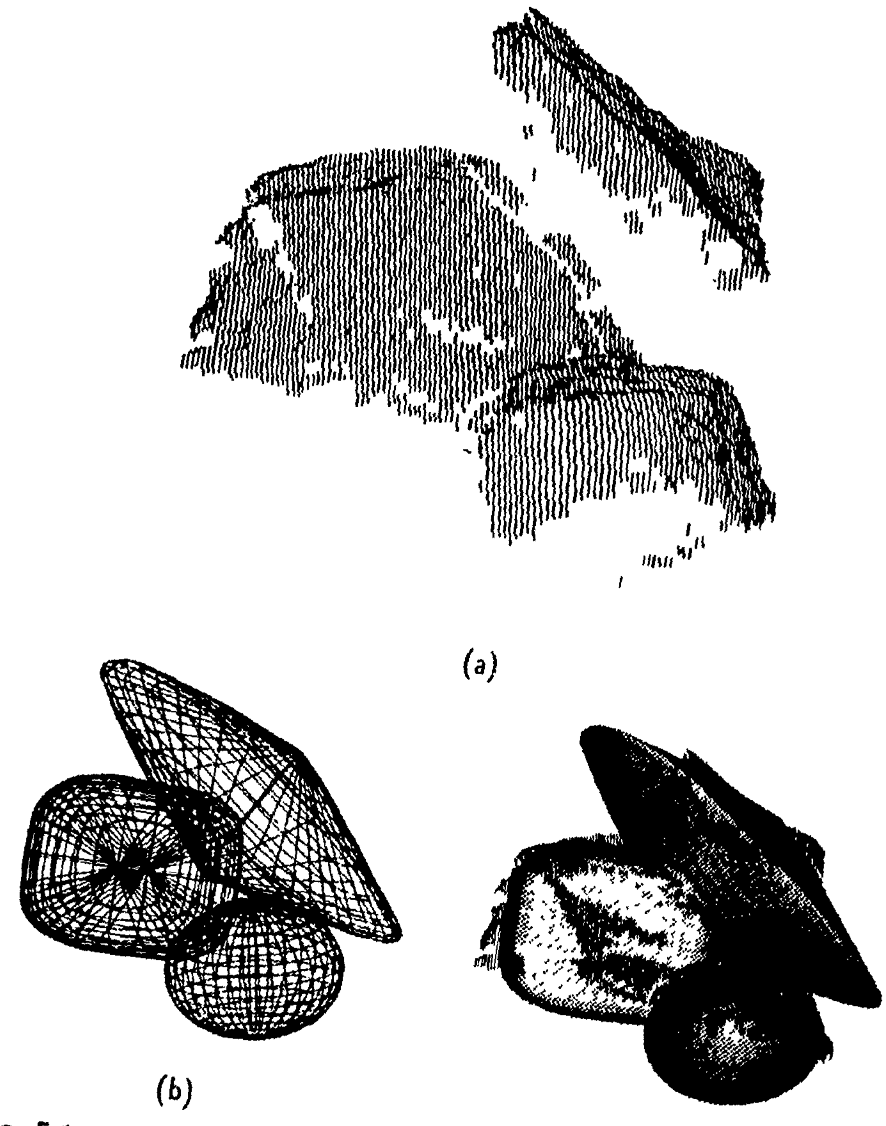


Figure 5.17: Surface decomposition at different scales.

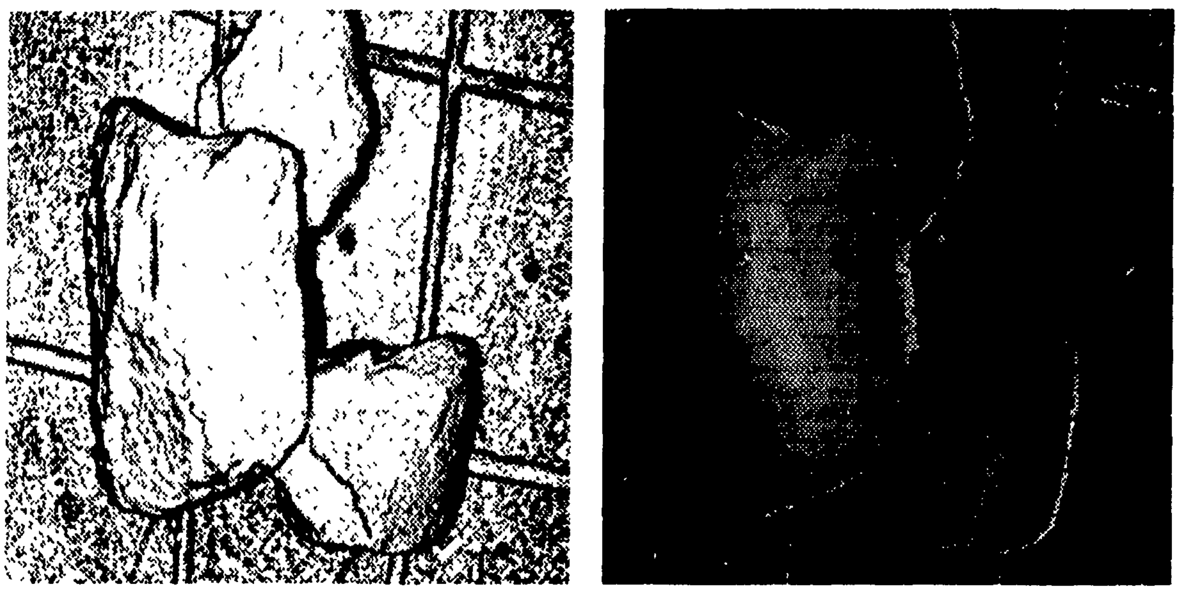


(a) (b) (c)

Figure 5.18: (a) Original range data partitioned according to the region map Fig. 5.17. (b) Superquadric (wire frame) model fitting of the rock-pile (c) Shaded image of the superquadric model.

Example four

The fourth example is of a muckpile sample consisting of three rock fragments (Fig. 5.19). In this particular example, a bigger rock fragment was lying on top of two smaller rock fragments Note that this case is unlikely to happen in a real mine, since small fragments tend to stay on top of the bigger ones after being unloaded from the LHD vehicles. Fig 5.20 shows the surface decomposition at different scales. The final muckpile model is shown in Fig 5.21



(a) (b)

Figure 5.19: Raw range measurements of muckpile. Only one muckpile is shown here, with one big rock fragment on top of two smaller fragments. (a) Shaded image. (b) Grey-scale image (darker pixels represent points further away from the scanner).

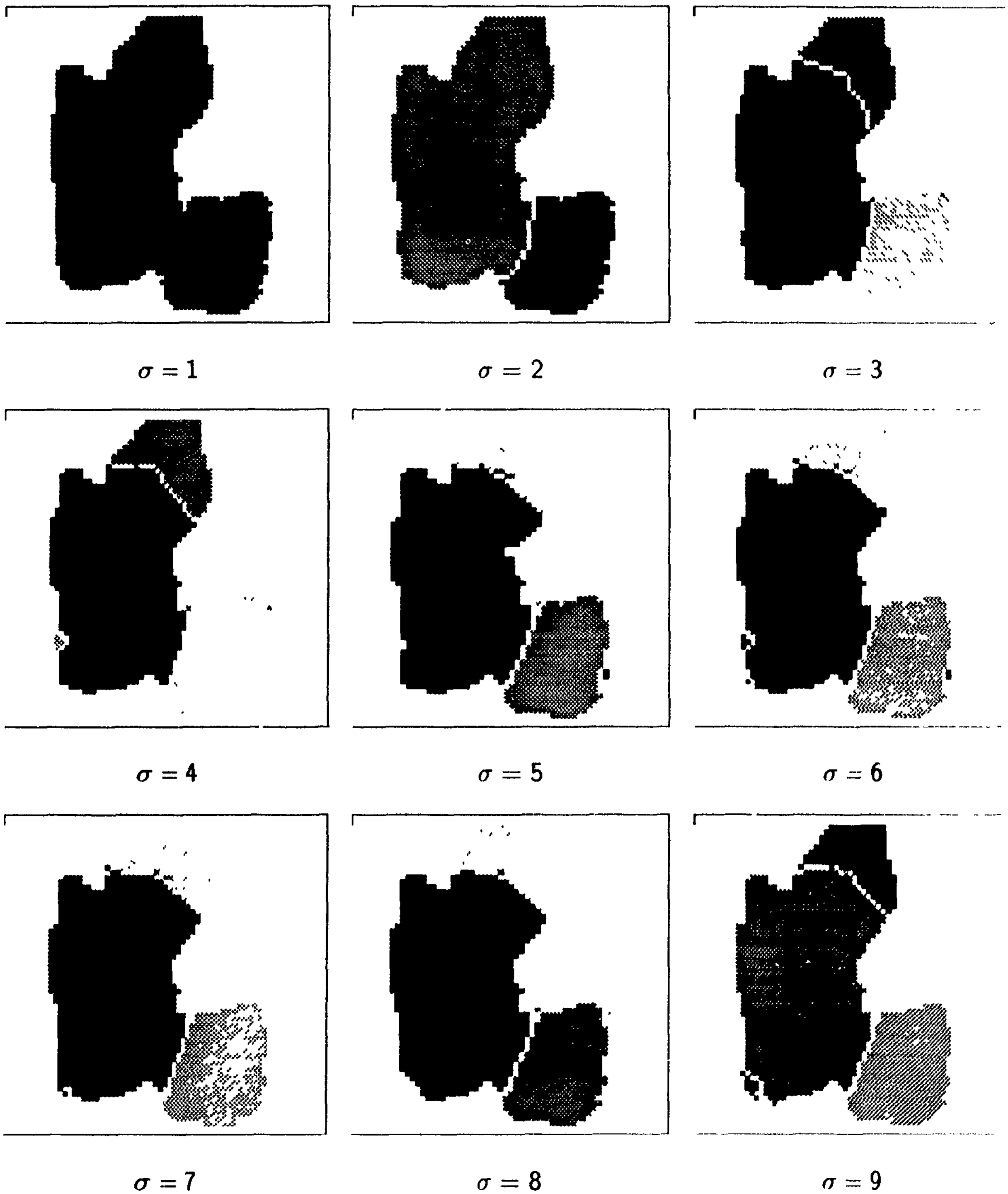


Figure 5.20: Surface decomposition at different scales.

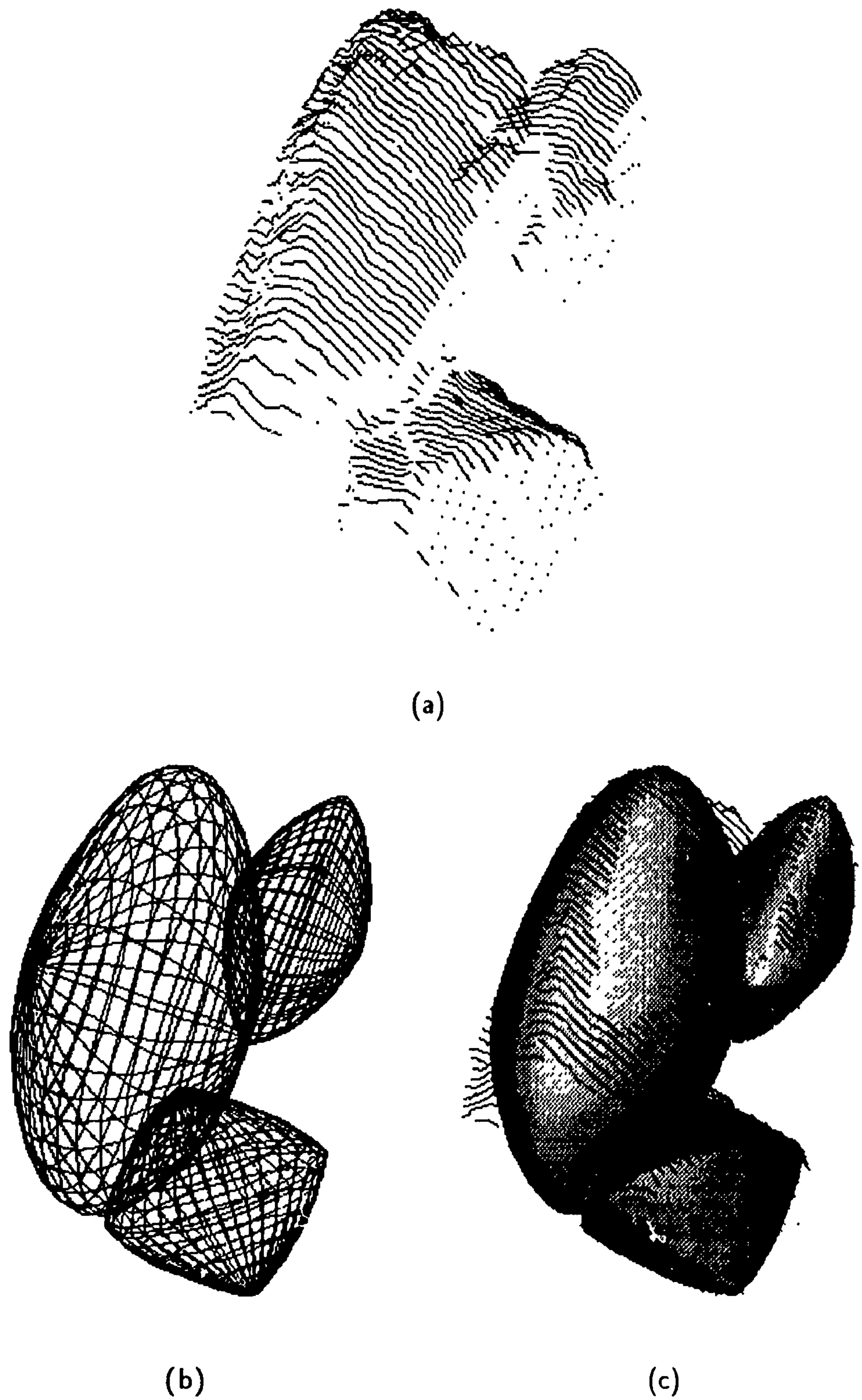


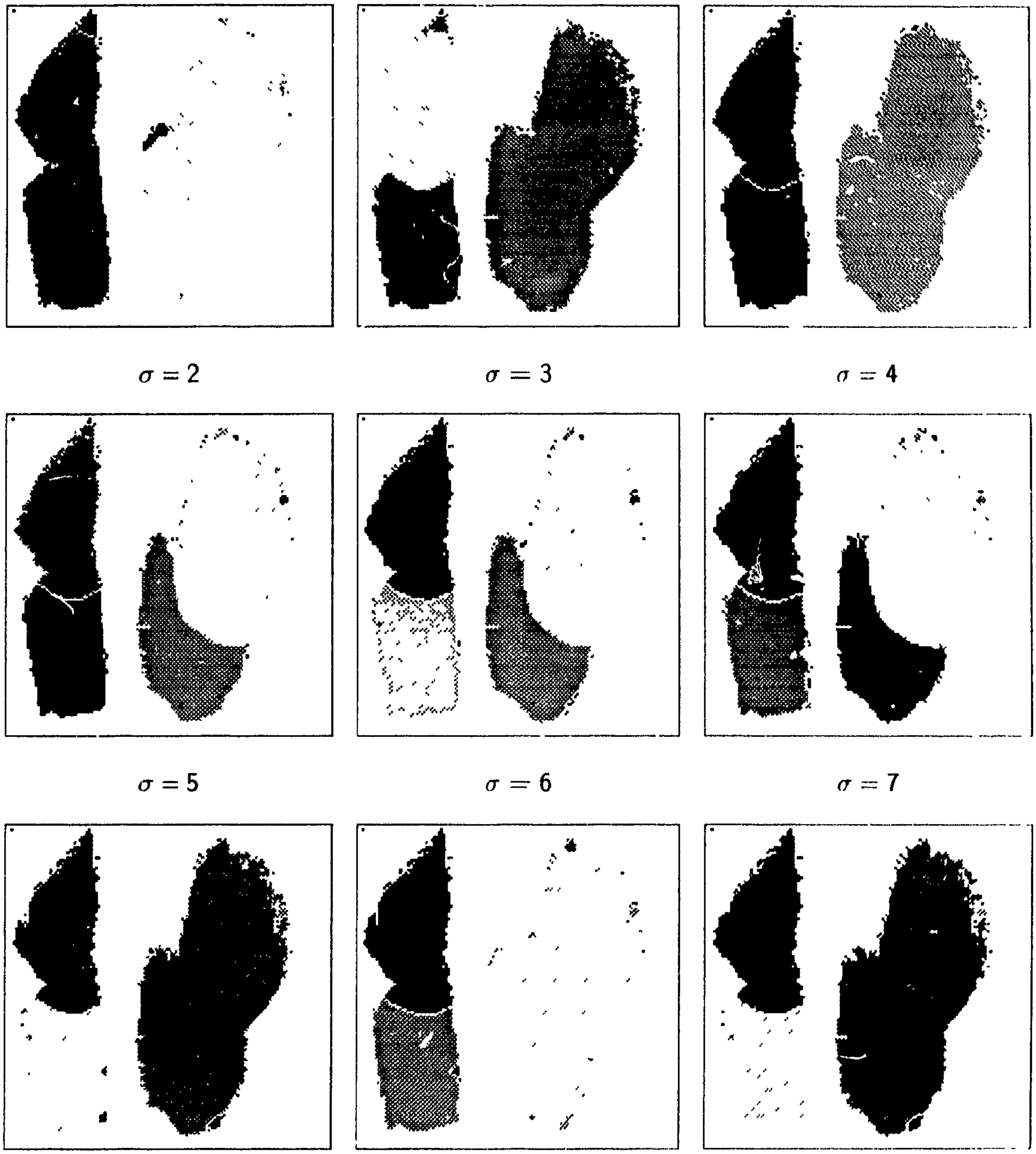
Figure 5.21: (a) Original range data partitioned according to the region map Fig. 5.20. (b) Superquadric (wire frame) model fitting of the rock-pile. (c) Shaded image of the superquadric model.

Example five

Fig 5.22 shows four rock fragments. The surface decomposition at different scales is shown in Fig. 5.23. Fig 5.24 shows the final muckpile model.

(a) (b)

Figure 5.22: Raw range measurements of muckpiles. Two piles are shown here, each with two rock fragments. (a) Shaded image. (b) Grey-scale image (darker pixels represent points further away from the scanner).

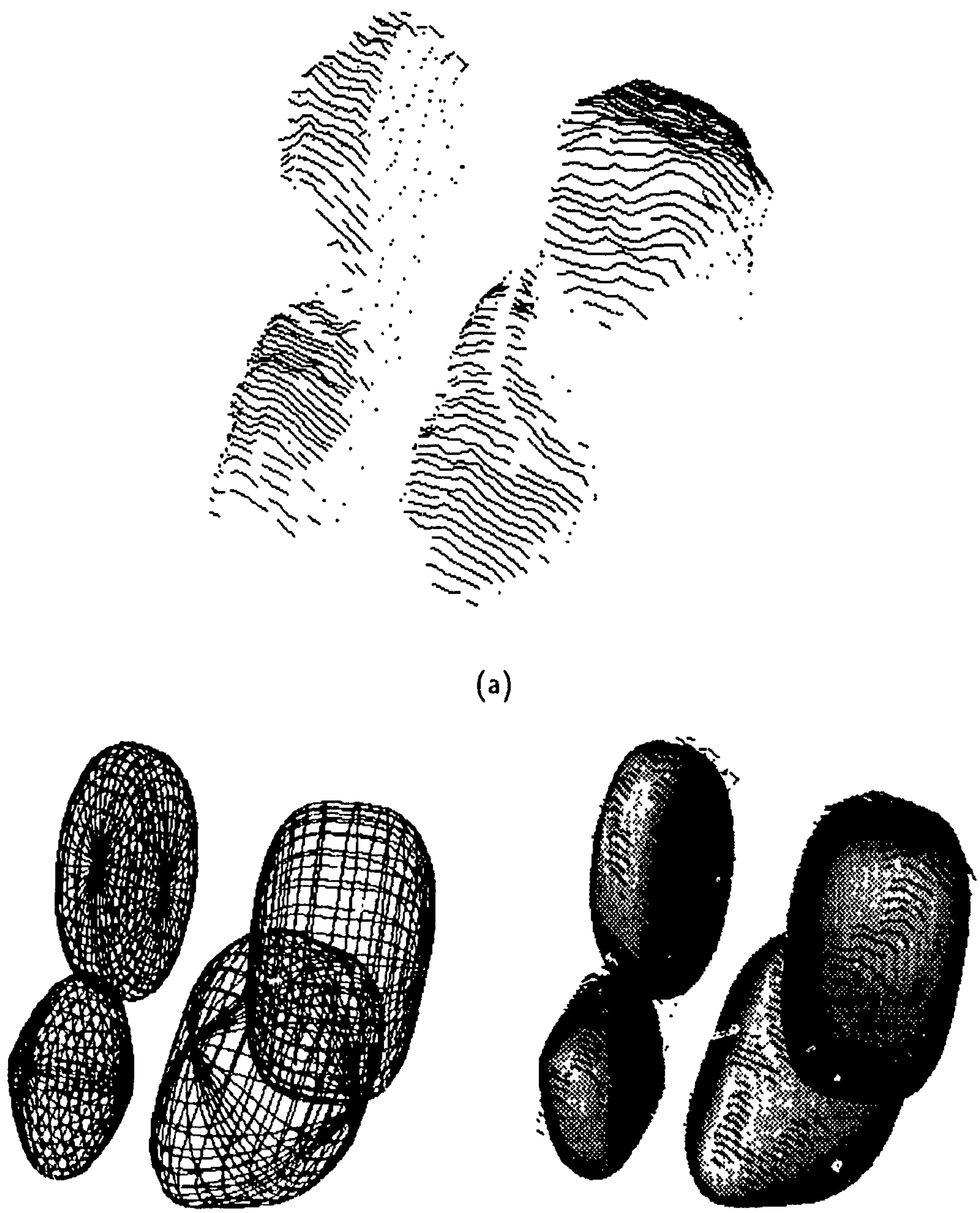


$\sigma = 2$ $\sigma = 3$ $\sigma = 4$

$\sigma = 5$ $\sigma = 6$ $\sigma = 7$

$\sigma = 8$ $\sigma = 9$ $\sigma = 10$

Figure 5.23: Surface decomposition at different scales.



(a) (b) (c)

Figure 5.24: (a) Original range data partitioned according to the region map Fig. 5.23. (b) Superquadric (wire frame) model fitting of the rock-pile. (c) Shaded image of the superquadric model.

5.4 Discussion

Six sets of results have been presented in this chapter, including five from the case studies It is very difficult to assess the reliability and robustness of the strategy proposed, based on the limited number of examples. However, among the total of more than twenty test cases that had been studied, the success rate was approximately 80 percent. This promising figure suggests that the strategy is indeed adequate for identifying and locating rock fragments from the configurations simulated in the laboratory

The main problem of this particular approach is the muckpile decomposition, and especially the determination of the scale parameter (σ) The experience gained from this exercise has shown that once the appropriate range of scales is established, based on trial-and-error, the rock lumps can be identified correctly. The scale parameter is highly correlated to the physical attributes of the equipment and the objects in the scene For example, the larger the rock size and/or the greater the surface irregularity, the larger the scale parameter (see Table 5.2)

	$\sigma = 3$	$\sigma = 4$	$\sigma = 5$	$\sigma = 6$	$\sigma = 7$
$\frac{\text{Average Rock Size}}{\text{Field of View}}$ (m^{-2})	5.6×10^{-6}	6.0×10^{-6}	6.4×10^{-6}	7.2×10^{-6}	8.0×10^{-6}
$\frac{\text{Average Size of Surface Irregularity}}{\text{Average Rock Size}}$	0.3×10^{-1}	0.4×10^{-1}	0.8×10^{-1}	1.1×10^{-1}	1.5×10^{-1}

Table 5.2: The correlation of scale parameter (σ) between the average rock size (4 – 8 cm diameter), the field of view ratio (1 m^3), and the average size of surface irregularity (0.2 – 1.5 cm)

The partitioning boundaries identified by muckpile decomposition algorithm do not always correspond to the "true" rock boundaries This is partly because of the smoothing effect from the scale-space filtering. Results have shown that the errors introduced are negligible, and are compensated for during the modelling fitting procedure

There are also some problems encountered in the experiment for modelling the muckpile, when the rock shape is highly irregular or flat These problems are due to the loose constraint in inferring the rock size with insufficient data and the incapability of superquadric for modelling complex rock shapes. Additional constraints can be embedded in the modelling fitting procedure, such that each overlapped rock fragment cannot be modelled by the parametric primitive that grows below the fragment(s) identified underlay

Chapter 6 Conclusions

In this thesis, steps towards the goal of rock fragment localization and identification required for the secondary rock-breakage automation, using computer vision-based techniques have been presented. In particular, the study reported in this thesis has concentrated on the inference of the muckpile surface structure from the range measurements obtained from a laboratory scale model. The basic concepts employed are borrowed from standard differential geometry for analyzing 3-D curves and surfaces.

The proposed strategy involves three main steps: (i) surface reconstruction, (ii) surface decomposition, and (iii) volumetric modelling. Sander's curvature consistency algorithm has proven to be very useful in recovering stable surface structure even with highly noisy measurements Subsequently, the part decomposition of muckpile is made possible because of the stable surface estimates. The final muckpile model was obtained by fitting superquadric primitives to the identified rock surfaces This gives a qualitative description of the position, orientation, size and shape of each rock fragment. This description is very useful for the "high level" operations, not just limited to rock-breaking application, but can also be applied to many applications of similar nature; such as ore analysis, blasting assessment, materials transportation, etc.

Not surprisingly, the scale factor turned out to be very important because of the highly textured surface and geometrical irregularity inherent in the litho-structure. In spite of the fact that, the scale factor is somewhat related to the resolution of the measurements and the physical size of the objects, the exact relationship remains not well understood. A number of questions remain to be answered — how to chose the range of scales, and how to combine information across different scales.

Although encouraging results have been obtained, based on the proposed strategy, much of the work remains to be refined and enhanced before it can be applied in the real mine environment. The following is a list of items from the strategy proposed which can be improved:

- For the surface reconstruction aspect, the current version of the curvature consistency algorithm does not have the capability of preserving surface discontinuities, however, this information is very important for the part decomposition process. A new enhanced version of the curvature consistency algorithm will be required, for example to include a mechanism that guides the update of surface estimates based on the error measurement.

- For the muckpile surface decomposition aspect, the boundary features such as negative local curvature minima and jump discontinuities have proven to be very useful. A more robust decomposition strategy will be required, such as combining the boundary information with the region information, say, within a certain neighbourhood size, and identifying the partitioning boundaries more reliably.

- For the modelling aspect, physically-based deformable models seem to be more promising than the existing "pure" superquadric models. They provide better local surface structure, which is very useful if the breaking tool is to be positioned precisely on the rock surface. Moreover, the physically-based model provides the basic dynamic elements which would be used to simulate and predict the physical responses of the object due to external forces. This would be very useful if one can derive a physical model from the ore samples, and use this model to predict how much force is required to break each of the oversized rock lumps and to subsequently predict how well the rock fragments will react.

Other practical problems:

(i) The viewing problem due to bad view angle or occlusion of small fragments. To solve this particular problem integration of range measurements at different viewpoints will be required.

(ii) Sensor fusion — integration of different sensor measurements. An aspect of the overall problem that has not been investigated here is the use of the rockbreaker hammer itself as a feedback in the rock discrimination process. Hypotheses for rock location identified by the vision system, could be tested by trial shifting of apparently separate rocks to ensure that they are indeed separate/exist before the breakage operations are initiated. Such a combination of see-and-touch techniques is of course common amongst intelligent creatures exploring new environments.

(iii) The computing time problem has not been studied in this thesis, but it is very important to the automation process. Therefore, it should be given a high priority in the future research

(iv) The safety problem, such as providing the worker with sufficient sheltering and protection if a high power laser is used as the energy source for the sensor.

The results obtained from the experiments are a successful demonstration of how computer vision techniques can be applied in locating and identifying highly irregular objects such as rock fragments. To conclude this thesis, more research will be required in the future to investigate the problems addressed here, and to build a one-to-one scale test-bed so that more in-depth assessment of the project can be made.

References

[Archibald and Amid, 1989] C Archibald and S Amid, "Calibration of a wrist-mounted range profile scanner," in *Proceedings Vision Interface '89* [Can, 1989], pp 24–28

[Baiden, 1988] G Baiden, "INCO'S LHD maintenance assistant system development," in *Proceedings, 3RD Canadian Symposium on Mining Automation*, (Montréal), pp 75–82, Sep 14-16 1988

[Bajcsy and Solina, 1987] R Bajcsy and F Solina, "Three dimensional object representation revisited," in *Proceedings, 1ST International Conference on Computer Vision* [Com, 1987], pp 231–240

[Ballard and Brown, 1982] D Ballard and C Brown, *Computer Vision* Englewood Cliffs, New Jersey Prentice-Hall, Inc , 1982

[Baratin *et al*, 1990] L Baratin, F Crosilla, and P Paronuzzi, "Image processing for determining joint parameters in difficult rock slope conditions," in *Proceedings, Close-Range Photogrammetry Meets Machine Vision*, vol 1395, (Zurich, Switzerland), pp 878–885, SPIE – The International Society for Optical Engineering, Sep 3-7 1990

[Barr, 1981] A Barr, "Superquadrics and angle preserving transformations," *IEEE Computer Graphics and Applications*, vol 1, pp 11–23, January 1981

[Barr, 1984] A Barr, "Global and local deformations of solid primitives," *Computer Graphics*, vol 18, pp 21–30, July 1984

[Bennett and Hoffman, 1987] B Bennett and D Hoffman, "Shape decompositions for visual recognition the role of transversality," in *Image Understanding 1985-86* (W Richards and S Ullman, eds), ch 8, pp 215–256, Norwood, New Jersey Ablex Publishing Co , 1987

[Beranek *et al*, 1986] B Beranek, J Boillot, and F Ferrie, "Laser sensor for adaptive welding," in *Proceedings, Optical Techniques for Industrial Inspection* [SPI, 1986], pp 195–199

[Besl and Jain, 1986a] P Besl and R Jain, "Segmentation through symbolic surface description," in *Proceedings, IEEE Conference Computer Vision and Pattern Recognition*, (Miami Beach, Florida), pp 77–85, Computer Society of the IEEE, IEEE Computer Society Press, June 22-26 1986

[Besl and Jain, 1986b] P Besl and R Jain, "Invariant surface characteristics for 3D object recognition in range images," *Computer Vision, Graphics, and Image Processing*, vol 33, pp 33–80, January 1986

[Besl and Jain, 1988] P Besl and R Jain, "Segmentation through variable-order surface fitting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 10, pp 167–192, March 1988

[Besl, 1988] P Besl, "Active, optical range imaging sensors," *Machine Vision and Applications*, vol 1, no 2, pp 127–152, 1988

[Binford, 1971] T Binford, "Visual perception by a computer," in *Proceedings, IEEE Conference on Systems and Controls*, (Miami, Florida), Dec 1971

[Blais and Rioux, 1986] F Blais and M Rioux, "BIRIS A simple 3-D sensor," in *Proceedings, Optics, Illuminations, and Image Sensing for Machine Vision*, vol 728, (Cambridge, Massachusetts), pp 235–242, SPIE – The International Society for Optical Engineering, Oct 26-31 1986

[Blake and Zisserman, 1987] A Blake and A Zisserman, *Visual Reconstruction* Cambridge, Massachusetts MIT Press, 1987

[Bonifazi and Massacci, 1989] G Bonifazi and P Massacci, "Ore deposit structure evaluation by image processing of exploitation walls," in *Proceedings, 6TH IFAC Symposium on Automation in Mining, Mineral and Metal Processing*, (Buenos Aires, Argentina), pp 39–43, The International Federation of Automatic Control, Sep 4-8 1989

[Brady *et al*, 1985] M Brady, J Ponce, A Yuille, and H Asada, "Describing surfaces," *Computer Vision, Graphics, and Image Processing*, vol 32, pp 1–28, October 1985

[Brooks, 1983] R Brooks, "Model-based three-dimensional interpretation of two-dimensional images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 5, pp 140–150, March 1983

[Bumbaca *et al*, 1986] F Bumbaca, F Blais, and M Rioux, "Real-time correction of three-dimensional nonlinearities for a laser rangefinder," *Optical Engineering*, vol 25, pp 561–565, April 1986

[Can, 1989] Canadian Image Processing and Pattern Recognition Society, *Proceedings Vision Interface '89*, (London, Ontario), June 19-23 1989

[Canny, 1986] J Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 8, pp 679–698, November 1986

[Carlsson and Nyberg, 1983] O Carlsson and L Nyberg, "A method for estimation of fragment size distribution with automatic image processing," in *Proceeding, 1ST International Symposium on Rock Fragmentation by Blasting*, (Luleå, Sweden), pp 333–345, Aug 1983

[Chabot *et al*, 1989] L Chabot, J Edwards, M Mazandarni, J Peck, and G Potvin, "Computer simulation studies of longhole drill deviation," in *Proceedings du Colloque CRM sur les équipements miniers*, (Rouyn, Québec), pp 233–266, Nov 7-11 1989

[Cheung and Ord, 1990] C Cheung and A Ord, "An on line fragment size analyser using image processing techniques," in *Proceedings, 3RD International Symposium on Rock Fragmentation by Blasting*, (Brisbane, Australia), pp 233–238, Aug 26-31 1990

[Cheung *et al*, 1990] W Cheung, M Lin, F Ferrie, G Carayannis, and J Edwards, "Pilot studies on computer vision techniques for an automated mine environment," in *Proceedings, 3RD National Conference on Robotics*, (Melbourne, Australia), pp 142–153, The Australian Robot Association, June 4-6 1990

[Choi *et al*, 1990] T Choi, H Delingette, M DeLusie, Y Hsin, M Hebert, and K Ikeuchi, "A perception and manipulation system for collecting rock samples," in *Proceedings, 4TH Annual Space Operations, Applications, and Research Symposium SOAR 90*, (Albuquerque, NM), June 1990

[Com, 1987] Computer Society of the IEEE, *Proceedings, 1ST International Conference on Computer Vision*, (London, U K), IEEE Computer Society Press, June 8-11 1987

[Dierckx, 1977] P Dierckx, "An algorithm for least-squares fitting of cubic spline surfaces to functions on a rectilinear mesh over a rectangle," *Journal of Computational and Applied Mathematics*, vol 3, no 2, pp 113–129, 1977

[do Carmo, 1976] M do Carmo, *Differential Geometry of Curves and Surfaces* Englewood Cliffs, New Jersey Prentice-Hall, Inc , 1976

[Fan *et al*, 1987] T Fan, G Medioni, and R Nevatia, "Segmented descriptions of 3-D surfaces," *IEEE Journal of Robotics and Automation*, vol RA-3, pp 527–538, December 1987

[Fan, 1990] T Fan, *Describing and Recognizing 3-D Objects Using Surface Properties* New York Springer-Verlag, 1990

[Faugeras and Hebert, 1986] O Faugeras and M Hebert, "The representation, recognition, and locating of 3-D objects," *The International Journal of Robotics Research*, vol 5, pp 27–52, Fall 1986

[Ferrie and Levine, 1988] F Ferrie and M Levine, "Deriving Coarse 3D Models of Objects," in *Proceedings, IEEE Conference Computer Vision and Pattern Recognition*, (Ann Arbor, Michigan), pp 345–353, Computer Society of the IEEE, IEEE Computer Society Press, June 5-9 1988

[Ferrie *et al*, 1989] F Ferrie, J Lagarde, and P Whaite, "Darboux frames, snakes, and super-quadrics. Geometry from the bottom-up," in *Proceedings IEEE Workshop on Interpretation of 3D Scenes*, (Austin, Texas), pp 170–176, Computer Society of the IEEE, IEEE Computer Society Press, Nov. 27-29 1989

[Ferrie *et al*, 1990] F Ferrie, J Lagarde, and P Whaite, "Recovery of volumetric object descriptions from laser rangefinder images," in *Proceedings, 1ST European Conference on Computer Vision*, (Antibes, France), pp 387–396, Springer-Verlag, Apr 23-27 1990

[Flynn and Jain, 1989] P Flynn and A Jain, "On reliable curvature estimation," in *Proceedings, IEEE Conference Computer Vision and Pattern Recognition*, (San Diego, California), pp 110–116, Computer Society of the IEEE, IEEE Computer Society Press, June 4-8 1989

[Fuentes-Cantillana *et al*, 1991] J Fuentes-Cantillana, J Catalina, A Rodriguez, J Orteu, and

D Dumahu, "Use of computer vision for automation of a roadheader in selective cutting operation " in *Proceedings International Symposium on Mine Mechanization and Automation* [ISM, 1991], pp 15–1–15–10

[Gao and Wong, 1989] Q Gao and A Wong, "Rock image segmentation," in *Proceedings Vision Interface '89* [Can, 1989], pp 125–133

[Gardner, 1965] M Gardner, "The "superellipse" a curve that lies between the ellipse and the rectangle," *Scientific American*, vol 213, pp 222–234, September 1965

[Gardner, 1975] M Gardner, "Piet Hein's superellipse," in *Mathematical Carnival*, ch 18, pp 240–254, New York Alfred A Knopf, Inc , 1975

[Girod, 1990] B Girod, ed , *Proceedings, Sensing and Reconstruction of Three-Dimensional Objects and Scenes*, vol 1260, (Santa Clara, California), SPIE – The International Society for Optical Engineering, Feb 15-16 1990

[Grimson, 1983] W Grimson, "Surface consistency constraints in vision," *Computer Vision, Graphics, and Image Processing*, vol 24, pp 28–51, October 1983

[Gross and Boult, 1988] A D Gross and T E Boult, "Error of fit measures for recovering parametric solids," in *Proceedings, 2ND International Conference on Computer Vision*, (Tampa, Florida, USA), pp 690–694, Computer Society of the IEEE, IEEE Computer Society Press, Dec 5-8 1988

[Gupta and Bajcsy, 1990] A Gupta and R Bajcsy, "Parts description and segmentation using contour, surface and volumetric primitives," in Girod [Girod, 1990], pp 203–214

[Hara *et al* , 1982] Y Hara, N Akiyama, and K Kavasaki, "Automatic inspection system for printed circuit boards," in *Proceedings, IEEE Workshop on Industrial Applications of Machine Vision*, (Silver Spring, MD), pp 62–70, May 3-5 1982

[Hoffman and Richards, 1984] D Hoffman and W Richards, "Parts of recognition," *Cognition*, vol 18, pp 65–96, 1984

[Horn, 1986] B Horn, *Robot Vision* Cambridge, Massachusetts MIT Press, 1986

[Hrechak and McHugh, 1990] A Hrechak and J McHugh, "Automated fingerprint recognition using structural matching," *Pattern Recognition*, vol 23, pp 893–904, 1990

[Hummel *et al*, 1987] A Hummel, B Kimia, and S Zucker, "Deblurring Gaussian blur," *Computer Vision, Graphics, and Image Processing*, vol 38, pp 66–80, April 1987

[Hunter *et al* , 1990] G Hunter, C McDermott, N Miles, A Singh, and M Scoble, "A review of image analysis techniques for measuring blast fragmentation," *Mining Science and Technology*, vol 11, pp 19–36, July 1990

[Hurteau *et al*, 1989] R Hurteau, P Corbeil, and A Piché, "Automatic positioning of a rockbreaker using vision and a tactile sensor," in *Proceedings International Workshop on Sensorial Integration for Industrial Robots Architecture and Applications*, (Zaragoza, Spain), pp 334–336, Nov 22-24 1989

[Hurteau *et al*, 1991] R Hurteau, M St-Amant, Y Laperriere, G Chevrette, and A Piché, "Optically Guided LHD A Demonstration Prototype," in *Proceedings International Symposium on Mine Mechanization and Automation* [ISM, 1991], pp 6–11–6–20

[Ikeuchi and Hebert, 1990] K Ikeuchi and M Hebert, "Task oriented vision," in *Proceedings Darpa Image Understanding Workshop*, (Pittsburgh, PA), pp 497–507, September 1990

[Int, 1985] The International Federation of Automatic Control, *Proceedings, 1ST IFAC Symposium on Automation for Mineral Resource Development*, (Brisbane, Australia), July 9-11 1985

[ISM, 1991] *Proceedings International Symposium on Mine Mechanization and Automation*, (Golden, Colorado, USA), June 10-13 1991

[Ittner and Jain, 1985] D Ittner and A Jain, "3-D surface discrimination from local curvature measures," in *Proceedings, IEEE Conference Computer Vision and Pattern Recognition*, (San Francisco, CA), pp 119–123, Computer Society of the IEEE, IEEE Computer Society Press, June 19-23 1985

[Jain and Jain, 1990] R Jain and A Jain, *Analysis and Interpretation of Range Images* New York Springer-Verlag, 1990

[Jarvis, 1983] R Jarvis, "A laser time-of-flight range scanner for robotic vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 5, pp 505–512, September 1983

[Johnson and Wichern, 1982] R Johnson and D Wichern, *Applied Multivariate Statistical Analysis* Englewood Cliffs, New Jersey Prentice-Hall, inc , 1982

[Kaisto *et al*, 1990] I Kaisto, J Kostamovaara, I Moring, and R Myllylä, "Laser rangefinding techniques in the sensing of 3-D objects," in Girod [Girod, 1990], pp 122–133

[Kanade, 1987] T Kanade, ed , *Three-Dimensional Machine Vision* Norwell, MA Kluwer Academic Publishers, 1987

[Kass *et al*, 1988] M Kass, A Witkin, and D Terzopoulos, "SNAKES active contour models," *International Journal of Computer Vision*, vol 1, pp 321–332, January 1988

[Kassler, 1985] M Kassler, "Robots and mining The present state of thought and some likely developments," in *Proceedings, 1ST IFAC Symposium on Automation for Mineral Resource Development* [Int, 1985], pp 31–33

[Koenderink and van Doorn, 1982] J Koenderink and A van Doorn, "The shape of smooth objects and the way contours end" *Perception*, vol 11, pp 129–137, 1982

[Koenderink, 1984] J Koenderink, "The structure of images," *Biological Cybernetics*, vol 50, pp 363–370, August 1984

[Lagarde, 1990] J Lagarde, "Constraints and their satisfaction in the recovery of local surface structure," Master's thesis, Dept Elect Eng , McGill University, Montréal, Québec, Canada, 1990

[Levine, 1985] M Levine, *Vision in Man and Machine* New York McGraw-Hill, 1985

[Lim, 1990] W Lim, "Qualitative 3-D models," in Girod [Girod, 1990], pp 180–190

[Livingstone and Rioux, 1986] F Livingstone and M Rioux, "Development of a large field of view 3-d vision system," in *Proceedings, Optical Techniques for Industrial Inspection* [SPI, 1986], pp 188–194

[Mäenpäa *et al*, 1983] I Mäenpäa, P Malinen, and R Söderström, "A computer system for photometric mineral sorting," in *Proceedings, 4TH IFAC Symposium on Automation in Mining, Mineral and Metal Processing*, (Helsinki, Finland), pp 499–509, The International Federation of Automatic Control, Aug 22-25 1983

[Maerz *et al*, 1987] N Maerz, J Franklin, L Rothenburg, and D. Coursen "Measurement of rock fragmentation by digital photoanalysis," in *Proceedings, 6TH International Congress of Rock Mechanics*, (Rotterdam, The Netherlands), pp 687–692, Aug 30 - Sep 3 1987

[Manana *et al*, 1985] R Manana, J Artieda, and J Catalina, "Ore sorting and artificial vision," in *Proceedings, 1ST IFAC Symposium on Automation for Mineral Resource Development* [Int 1985], pp 235–240

[Mandeville, 1985] J Mandeville, "Novel method for analysis of printed circuit images," *IBM Journal of Research and Development*, vol 29, pp 73–86, January 1985

[Marr, 1982] D Marr, *Vision a computational investigation into human representation and processing of visual information* San Francisco Freeman, 1982.

[Mitchell and Gillies, 1989] B Mitchell and A Gillies, "A model-based computer vision system for recognizing handwritten ZIP codes," *Machine Vision and Applications*, vol 2, no 4, pp 231–243, 1989.

[Nilsson and Lindberg, 1989] B Nilsson and C Lindberg, "Investigation of control problems in rockbreaker automation," Tech Rep 1989 140E, University of Luleå, Luleå, Sweden, September 1989

[O'Neill, 1966] B O'Neill, *Elementary Differential Geometry* San Diego, California Academic Press, Inc , 1966

[Orteu and Devy, 1991] J Orteu and M Devy, "Application of computer vision to automatic selective cutting with a roadheader in a potash mine," in *Proceedings, 5TH International Conference on Advanced Robotics*, (Pisa, Italy), pp 385–392, June 19-22 1991.

[Parent and Zucker, 1989] P Parent and S Zucker, "Trace inference, curvature consistency, and curve detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 823–839, August 1989

[Pentland and Sclaroff, 1991] A Pentland and S Sclaroff, "Closed-from solutions for physically based shape modeling and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 13, pp 715–729, July 1991

[Pentland, 1986] A Pentland, "Perceptual organization and the representation of natural form," *Artificial Intelligent*, vol 28, pp 293–331, May 1986

[Pentland, 1987] A Pentland, "Recognition by parts," in *Proceedings, 1ST International Conference on Computer Vision* [Com, 1987], pp 612–620

[Pentland, 1990] A. Pentland, "Automatic extraction of deformable part models, *International Journal of Computer Vision*, vol 4, pp 107–126, March 1990

[Ponce and Brady, 1987] J Ponce and M Brady, "Toward a surface primal sketch," in Kanade [Kanade, 1987], pp 195–240

[Press *et al*, 1988] W Press, B Flannery, S Teukolsky, and W Vetterling, *Numerical Recipes in C – The Art of Scientific Computing* Cambridge Cambridge University Press, 1988

[Rioux and Blais, 1986] M Rioux and F Blais, "Compact three-dimensional camera for robotic applications," *Journal of the Optical Society of America A*, vol 3, pp 1518–1521, September 1986

[Rioux, 1984] M Rioux, "Laser range finder based on synchronized scanners," *Applied Optics*, vol 23, pp 3837–3844, November 1984

[Rosenfeld and Kak, 1982] A Rosenfeld and A Kak, *Digital Picture Processing*, vol 1 & 2 New York Academic, second ed , 1982.

[Rosenfeld and Thurston, 1971] A Rosenfeld and M Thurston, "Edge and curve detection for visual scene analysis," *IEEE Transactions on Computers*, vol C-20, pp 562–569, May 1971

[Salamon, 1976] M D G Salamon, "Modern control systems in mining," in *Proceedings, 2ND IFAC Symposium on Automation in Mining, Mineral and Metal Processing*, (Johannesburg, South Africa), pp 1–7, The International Federation of Automatic Control, Sep 13-17 1976

[Sander and Zucker, 1990] P Sander and S Zucker, "Inferring surface trace and differential structure from 3-D images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 12, pp 833–854, September 1990

[Sander, 1988] P Sander, *On Reliably Inferring Differential Structure from Three-Dimensional Images* PhD thesis, Dept Elect Eng , McGill University, Montréal, Québec, Canada, 1988

[Solina and Bajcsy, 1987] F Solina and R Bajcsy, "Range image interpretation of mail pieces with superquadrics," in *Proceedings AAAI-87 Sixth National Conference on Artificial Intelligence*, (Seattle, Washington), pp 733–737, The American Association for Artificial Intelligence, July 13-17 1987

[Solina and Bajcsy, 1990] F Solina and R Bajcsy, "Recovery of parametric models from range images The case for superquadrics with global deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 12, pp 131–147, February 1990

[Soucy, 1992] G Soucy, "View correspondence using curvature and motion consistency," Master's thesis, Dept Elect Eng , McGill University, Montréal, Québec, Canada, 1992 in preparation

[SPI, 1986] SPIE – The International Society for Optical Engineering, *Proceedings, Optical Techniques for Industrial Inspection*, vol 665, (Québec City, Québec), Jun 2-6 1986

[St-Amant *et al*, 1991] M St-Amant, Y Laperriere, R Hurteau, and G Chevrette, "A simple robust vision system for underground vehicle guidance," in *Proceedings International Symposium on Mine Mechanization and Automation* [ISM, 1991], pp 6–1–6–10

[Terzopoulos and Metaxas, 1991] D Terzopoulos and D Metaxas, "Dynamic 3D models with local and global deformations Deformable superquadrics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol 13, pp 703–714, July 1991

[Terzopoulos *et al*, 1987] D Terzopoulos, A Witkin, and M Kass, "Symmetry-seeking models and 3D object reconstruction," *International Journal of Computer Vision*, vol 1, pp 211–221, October 1987

[Terzopoulos, 1983] D Terzopoulos, "Multilevel computational processes for visual surface reconstruction," *Computer Vision, Graphics, and Image Processing*, vol 24, pp 28–51, October 1983

[Ullman, 1979] S Ullman, *The interpretation of visual motion.* Cambridge, Massachusetts: MIT Press, 1979

[Vagenas *et al*, 1991] N Vagenas, H Sjoberg, and S. Wikstrom, "Application of remote-controlled/automatic Load-Haul-Dump system in Zinkgruvan, Sweden," in *Proceedings International Symposium on Mine Mechanization and Automation* [ISM, 1991], pp. 6–21–6–30.

[Vemuri *et al.*, 1987] B Vemuri, A. Mitiche, and J Aggarwal, "3-D object representation from range data using intrinsic surface properties," in Kanade [Kanade, 1987], pp 241–266

[Whaite and Ferrie, 1991] P Whaite and F Ferrie, "From Uncertainty to Visual Exploration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp 1038–1049, October 1991.

[Witkin, 1983] A Witkin, "Scale-space filtering," in *Proceedings, 8TH International Joint Conference on Artificial Intelligence*, (Karlsruhe, West Germany), pp 1019–1022, Aug 8-12 1983.

[Yokoya and Levine, 1988] N. Yokoya and M. Levine, "A hybrid approach to range image segmentation," in *Proceedings, 9TH International Conference on Pattern Recognition*, (Rome, Italy), pp. 1–5, Computer Society of the IEEE, IEEE Computer Society Press, Nov 14-17 1988.