INCREMENTAL RECOVERY OF VOLUMETRIC MODELS

Santanu Roymoulik

Department of Electrical Engineering
McGill University, Montréal

May 1996

A Thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements of the degree of
Master of Engineering

Abstract

This research investigates the use of a deformable model representation in the context of an active exploration process. The exploration process is used to acquire data, fuse measurements from different viewpoints, and construct shape models. Till now approximate shape models, e.g., superellipsoids have been used to represent objects in an active exploration strategy. Although these models provide an adequate characterization of 3D shape for many tasks, a more detailed representation is needed when it comes to tasks that require discrimination of similar objects.

In this thesis, we explore the use of a class of deformable models, developed by Pentland and Sclaroff, to futher refine the initial coarse superellipsoid description to provide a richer description of the scene. The autonomous exploration strategy used is the one developed by Whaite and Ferrie. Along the way a number of specific issues will be addressed, namely how the models are initialized, the stability of the model parameters throughout the exploration process, and the attachment of data points to the model surface. This work demonstrates the utility of the deformable models for shape interpolation and as general representations for 3D shape.

Acknowledgements

First and foremost, I would like to thank my supervisor, Frank P. Ferrie, for his support and his continuous encouragement throughout my research work. Thanks go out to some individuals in my group, whose help was essential towards the completion of this thesis. Thanks to Peter Whaite, Tal Arbel, Gilbert Soucy, Shailendra Mathur for their technical expertise and support. A special thank you to Gilbert Soucy whose help in technical support is greatly appreciated.

Thanks to my family and friends for their support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

1. Introduction

Recovering the shapes and motions of 3D objects from their images is one of the major challenges facing Computer Vision. The problems of shape reconstruction and object recognition have preoccupied researchers for years [**33**]. To achieve this goal, one must be able to build stable representations of visual models which can integrate conflicting and diverse constraints. These visual models should be able to improve upon conventional representations that impose simplifying assumptions about shape and motion. Deformable models are one such desirable class of models. This thesis explores the use of these models to represent objects imaged using an active exploration strategy.

The real world is three-dimensional, obeys the laws of physics, and is often nonrigid. Therefore, to predict and analyse an object's behavior we require models with which we can easily and succinctly describe dynamic and nonrigid three-dimensional behavior [**34**]. These models should be dynamic in the sense that they should evolve over time. Until now, most of the research on recovering global abstractions or representations is based on the concept of rigid motion. But it is nonrigid motion that the human visual system copes with routinely - cheeks bulging and stretching, body parts twisting and bending, trees swaying among the many things we can think of [**22**]. Clearly, we need models that can accomodate deformation, nonplanarity, inexact symmetry and a whole range of localized irregularities [**33**].

Reconstructed models should be accurate, thus the recovery process must be able to incorporate models of sensor error as well as a priori knowledge of it's environment [**15**]. Model parameters should be stable, so that changes in viewpoint, sampling, or sensor noise do not affect the resulting 3D description. Descriptions should be unique, so that object recognition can proceed by simply comparing recovered model parameters. For this purpose volumetric models such as spheres, cylinders, and prisms seem appropriate, since they can decompose composite shapes into natural parts that are compactly expressible using a small set of parameters. The problem of shape modeling can be approached by simultaneously satisfying the requirements of reconstruction and recognition and trying to provide a smooth transition between these two important aspects of vision.

Until now, most of the high-level computer vision systems have been unsuccessful at developing representations which are general and stable enough to be useful in encoding a variety of images. There has been a clear division between representations which provide the necessary abstraction - parametric models [**33**] - and those which provide the necessary expressiveness - finite element-based, deformable models [**24**]. Clearly we should be better off with a hybrid model, one which can handle both representations. The parametric models used in this thesis are superellipsoids which are an extension of the ellipsoid model, but with fractional powers. The advantage of using this parametric model is that it can evolve into a variety of shapes. Deformable models are an improvement over the superellipsoid model, in that, modeling capabilities are enhanced by providing for deformations like tapering, bending, etc. In this thesis we will focus on such deformable models as our hybrid model.

A goal of this research is to automatically recover accurate volumetric models of 3D objects for computer-aided design (CAD), object recognition, and for robotic applications. A physics-based approach is used to infer 3D shape models from a collection of sensor measurements. The basic idea of this approach is to fit a 3D volumetric model to a collection of range datapoints. Superellisoids are used to provide us with an approximate initial model. Deformable models are then used to refine the initial superellipsoid shape descriptors to model a variety of shapes and to provide a more accurate description of the scene, e.g., for purposes of object recognition. For this reason, the initial superellipsoid model should evolve to a more elaborate deformable model. This can be achieved by an active process where an observer interacts with the environment by moving about and collecting more information and inferring a dynamic representation of the world [**36**]. In the model fitting problem, usually a single viewpoint of the camera is used. Even with the active process, only superellipsoid models have been fit to data. The principal contribution of this thesis is to investigate the use of deformable models in a dynamic modeling process which is achieved using an active exploration strategy [**36**]. To our knowledge, this is the first time deformable models have been used in such a context. A number of issues arise in using these sort of models, e.g., the stability and uniqueness of the resulting model descriptors.

Initially, a superellipsoid model is fit to a data set from one view only. The next logical step would be to incrementally recover volumetric deformable models by exploring the object from several different viewpoints. Here we investigate the use of Whaite and Ferrie's autonomous explorer [**36**] as a means of acquiring the necessary data. However, such a dynamic strategy raises a number of issues. The deformable model has to be initialized at every instance of the data being added to the existing data-sets. Stability, normalization, and convergence of the deformable model parameters are important under this exploration strategy are vital. Various trials of the exploration process are performed on a object starting from a different initial viewpoint each time. We show that at each stage the stability of the model parameters is maintained as the model is initialized with the addition of new datapoints. The results indicate that the parameters of our deformable model representation are stable over various iterations of the experiments and that the model is unique with respect to sampling and viewpoint. This is essential, for example, in parametric object recognition where model parameters are used to index into a database of stored object prototypes [**1**].

Our approach to deformable model representation is based on the work of Pentland and Sclaroff There are two classes of deformable models - those based on parametric solid modeling primitives using superquadrics [**24**], and those based on mesh-like surface models, such as employed by Terzopoulos *et al.* [**33**] These approaches deal with model recovery from uniformly sampled 3D data, or data sets from a single viewpoint. Whaite & Ferrie [**36**], however, have argued that such approaches do not sufficiently constrain the model, resulting in many possible model representations that would fit the data accurately as shown in Figure 2.5. Our work takes this fact into account, so that the final model is incrementally recovered over time by exploring the object from several viewpoints. In doing so, we are able to maintain the stability of the model, viewpoint independence and other factors which are inherent in such a dynamic modeling environment.

The incremental model recovery has been successfully implemented and the strategies explored in this thesis have been integrated in a dynamic model building process (active exploration strategy). Specifically, this research makes the following contributions:

- This thesis investigates the use of the deformable model representation [**24**] in the context of the active exploration process, which is used to acquire data, fuse measurements and construct approximate shape models from different viewpoints. This leads to a dynamic incremental model recovery approach in which the deformable models evolve gradually.

- A number of experiments have been performed to infer unique model descriptions and to test the stability of the model parameters. We show that, on adding new data from the current data acquisition step, the model can be correctly initialized and that the resulting parameters are stable. Invariance to sampling and viewpoint is demonstrated - model parameters are shown to converge to acceptable limits inspite of variations in sampling and viewpoint trajectory.
- During the volumetric model fit to the range data, special care has to be exercised in the attachment of the data points to the solid model. Correspondence routines have been developed for fitting a volumetric model to uniformly sampled, partial, and 3D data sets. Such types of correspondence issues has been addressed by Pentland [**24**]. The routines developed in this research are general enough to fit 3D datasets that are either partial or uniformly sampled.
- The model fitting strategies developed in this thesis have been integrated with an active exploration process developed in our laboratory. The process is completely bottom-up and does not depend on any a priori knowledge of the environment.

2. Approach

The hybrid representation used in this work is a physically-based, parametric model which is based in part on Terzopoulos's work with deformable models [**33**] and on Pentland and Sclaroff's work with modal analysis [**24**]. The notion of employing physical constraints in shape modeling has been suggested by many authors. However, the seminal paper by Terzopoulos *et al.* [**34**], who obtained 3D solid models by fitting simulated rubbery sheets and tubes, has focused attention on modeling methods which rely on physics and mathematics to simulate the dynamics of real objects. In this modeling framework, objects are modeled as deformable elastic models which are constrained by forces applied to the objects. These forces deform the objects and simultaneously satisfy the constraints. The above shape modeling techniques make use of finite element methods (FEM) since they can accurately describe common vision tasks like fitting, interpolation, or correspondence problems [**24**]. This thesis encompasses various aspects of volumetric modeling techniques and demonstrates the utility of these models for shape interpolation and as general representations for 3D shape.

Superellipsoids are used as primitives for shape representation. This has been demonstrated by various authors [**31, 9**] . While useful for general shape classification, superellipsoid models often fall short when it comes to differentiating between two similar objects - a richer model is required. Hence we are interested in developing a method for determining the parameters of a richer model such as a deformable using a superellipsoid model as a basis. These kinds of shape abstractions are essential for efficient matching in object databases for a higher level process like object recognition. Pentland [**24**] has illustrated that object recognition can proceed by comparing the recovered model parameters. There are a variety of representations possible to take us from the very low-level measurements to a higher level description [**33**]. Low level descriptors are much easier to recover but object recognition, which requires matching in object databases, is quite complex because of the numerous representations possible. On the other hand, a higher level descriptive model has a lot of parameters associated with it making parameter recovery an involved process. However object recognition is simplified since our deformable model representation is unique. So the notion of superquadrics, for example can be expanded to include non-rigid transformations such as shearing, tapering, bending and pinching, leading to a general class of deformable superquadrics [**24**] providing a much richer representation of 3D shape. This class can model a large variety of geometrical and natural shapes at the level of parts (Figures 2.1 and 2.2). Here, we show the various shapes generated by varying the shape exponents and the deformation modes of a superquadric. In this thesis we will use Pentland and Sclaroff's model analysis techniques to refine the initial approximations (superellipsoids) into a richer description of the scene.

In the presence of a uniformly sampled 3D data set, the methodology outlined by Pentland and Sclaroff [**24**] would generally lead to a stable and unique set of model parameters. Range sensors, on the other hand, provide only 2 & 1/2 D data, i.e. distance to the object's surface for some particular vantage point. Whaite and Ferrie have shown [**36**] that in the absence of additional constraints the fitting problem is underdetermined in this context. Their solution to this problem was to use model parameter uncertainity as feedback to an automatic gaze-planning algorithm. The methodology, which they refer to as the autonomous explorer, is to sequentially collect data from different vantage points until model parameter variance falls below a prescribed threshold. A question posed by this research is whether the same approach can suffice for more elaborate deformable models. Our results show that this is basically the case.

The exploration strategy (Figure 2.3) has been tied in with the deformable model recovery in this research. Low-level vision tasks such as data acquisition, visual reconstruction lead up to higher level tasks such as segmentation to partition a shape into principal components that correspond to the parts of each object. Then part boundaries are described by parametric forms (i.e. models) such as superquadrics. These model parameters are then used as a starting point from which deformable parameters are obtained providing us with a richer model. Experimental validation for the stability of the model is provided based on the parameter values obtained by the autonomous explorer at each stage of data acquisition.

3. Organization of the thesis

The thesis is organized as follows. Chapter 2 describes the shape modeling techniques in the context of deformable models. The bottom-up vision strategy used in this thesis is explained. The notion of superquadrics and the need for *autonomous exploration* is explained in greater detail.

Chapter 3 describes the motivation for using FEM (Finite Element Method) and describes the FEM representation and modal analysis. The methodology employed by Pentland [**24**] for recovering volumetric deformable models from range data is discussed. This provides the basic theoritical background for this thesis.

Chapter 4 demonstrates the use of Pentland's method of deformable model recovery under the autonomous exploration strategy proposed by Whaite & Ferrie. This is the major contribution in this thesis. The algorithm used for the deformable model recovery is explained in detail. Special emphasis is placed on establishing the correspondence between the data points and nodes on the model. Several examples of the deformable model fitting are presented. The methodology for performing the experiments is illustrated in detail. The sensor control and data acquistion method used in the laboratory are discussed.

Chapter 5 shows a number of examples of recovering finite-element based deformable models from range data. Experiments are done with uniformly sampled, partial data and 3D range data. Autonomous exploration results are presented and the stability of the model parameters is demonstrated over multiple trials of the exploration process.

Chapter 6 Conclusions are drawn and research directions for future work are suggested.

CHAPTER 2

Shape Modeling Concepts

1. Introduction

In this research, we consider shape modeling techniques in the context of an active exploration strategy using deformable volumetric shape models. What is required is a mechanism by which an observer can dynamically acquire more information about the object in order to produce a richer description. We will make use of the strategy of *autonomous exploration* [**36**] as the basis for developing such a mechanism. A brief description of this strategy is presented later in this chapter. Focus is also on models which can be used to fit complex 3D shapes which can capture both gross shape features and local shape details. The theory of superquadrics on which the deformable models is based, is also presented in order to make the presentation self-contained.

2. Research Background

Volumetric primitives are essential to generate compact object-oriented descriptions and to define global part structure. 3D representations such as generalized cylinders and cones [**18, 12, 8**], polyhedral models, and 3D symmetric axis transforms [**17**] have been commonly used by model-based vision systems. Parametric models such as generalized cylinders and their derivatives are appropriate as volumetric primitives because they give compact constrained estimates of the overall shape. A generalized cylinder representation was first recommended in 1971 by Binford as a convenient representation for 3D surfaces for the purposes of vision [**5**], but generalized cylinders are overly restrictive in that they can accurately represent only perfectly axial symmetric shapes. Due to the asymmetry inherent in many synthetic and natural shapes, generalized cylinders can result in a loss of crucial information. Marr and Nishihara [**16**] were among the first to propose a hierarchical representation of objects in terms of parts. Their work used generalized cylinders to describe each part, thereby limiting the scope of the representation to objects adequately describeable as collections of generalized cylinders.

Motivated by the generalized cylinder idea and the need to go beyond geometry and exploit computational physics in the modeling process, Terzopoulos, Witkin, and Kass [**34**] proposed a 3D shape model similar to generalized cylinders, but enhanced with deformation parameters and constraints which are expressed as forces applied to the model. The applied forces deform the models such that they satisfy the available constraints over time. Model recovery requires setting constraints on multiple parameters often requiring human interaction. The fitting function for these distributed models is complicated and numerical differentiation is required. The result, being a number of equations which are not closed-form and thus are difficult to solve.

The generalized cylinder representation requires the specification of an axis, generally a space curve, and the cross-section function. Pentland [**19, 20**] proposed the use of a simpler part model with scalar parameters: the superquadric ellipsoid with parameterized deformations [**2**] like tapering, bending, etc. Pentland's proposal has spawned a flurry of efforts to reconstruct superellipsoids with global geometric deformations from 3D data, and these have been met with some success [**7, 31**]. Superquadric models possess descriptive adequacy [**27**] and they are also quite stable with respect to scale, noise, and configuration [**30, 31**].

In [**24**], Pentland followed a physically based approach by proposing another method for fitting deformable part models based on superellipsoids. Inspired by modal analysis (a technique for analyzing the vibration of linear mechanical systems under peroidic forcing conditions [**4**]), he applied polynomial deformation "modes" to superquadrics which has the metaphor of pushing, pinching, and pulling a lump of clay into a shape.

2.1. Basic Modeling Primitive - Superquadrics To trade some flexibility in shape description for a simpler and more efficient recovery of models, we make use of the family of superquadrics [**2**] as the basic volumetric primitive. This family of parametric shapes was first discovered by Danish designer Peit Hein [**11**] as an extension of basic quadric surfaces and solids. Superquadrics have been proposed for use as primitives for shape representation in computer graphics [**2**] and computer vision [**19, 30**].

Advantages of these volumetric primitives [**2**] are:

- Superquadric models are implicitly constrained by the class of the shapes that they can model and do not require any setting of external constraints from case to case.
- An implicit advantage for ease of model recovery is that the function that has to be minimized is differentiable everywhere. This is due to the fact that a superquadric function is defined with a single analytic function.
- Boolean operations can be defined on superquadric primitives and variation of shapes can be achieved with two shape parameters or with global parametric deformations. Superquadrics with parametric deformations encompass a large variety of natural shapes yet are simple enough to solve for their parameters.
- Due to their built-in symmetry, superquadric models predict the shape of the occluded parts by assuming global symmetry.

Superquadrics are a generalization of the equation of an ellipsoid ($(x/a_1)^2 + (y/a_2)^2 + (z/a_3)^2 = 1$), but rather than using only powers of two, a superquadric allows fractional powers, e.g.,

$$((x/a_1)^{2/\epsilon_1} + (y/a_2)^{2/\epsilon_1})^{\epsilon_1/\epsilon_2} + (z/a_3)^{2/\epsilon_2} = 1.$$

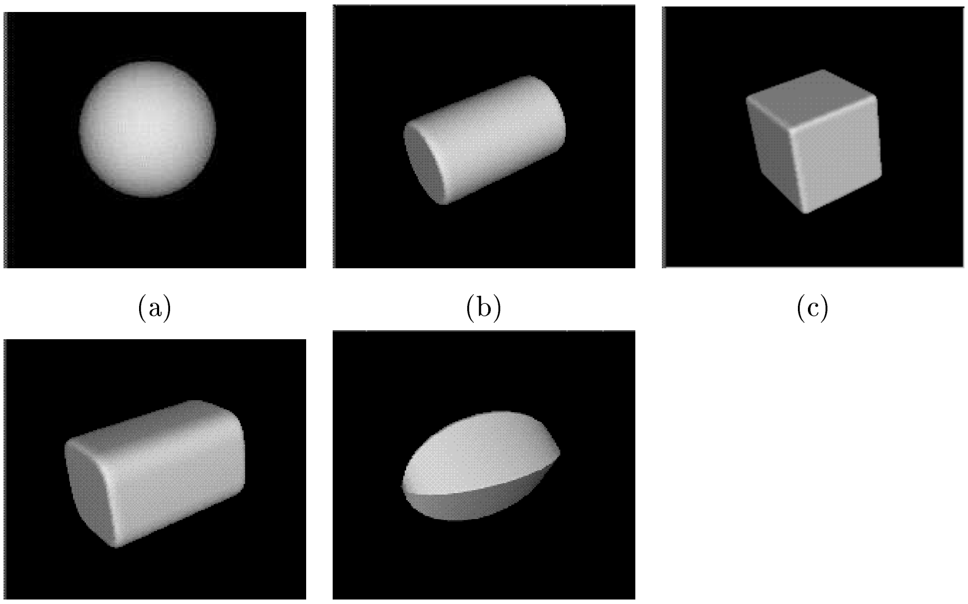
The variables x and y are grouped together to allow separate control of the $x - y$ profile and the $z - (xy)$ profile via the shape parameters ϵ_1 and ϵ_2.

The parametric representation of a superquadric surface is defined by the following 3D vector:

$$\mathbf{x}(\eta, \omega) = \begin{pmatrix} \mathbf{a}_1\mathbf{cos}(\eta)^{\epsilon_1}\mathbf{cos}(\omega)^{\epsilon_2} \\ \mathbf{a}_2\mathbf{cos}(\eta)^{\epsilon_1}\mathbf{sin}(\omega)^{\epsilon_2} \\ \mathbf{a}_3\mathbf{sin}(\eta)^{\epsilon_1} \end{pmatrix} \tag{2.1}$$

where $-\pi/2 \leq \eta \leq \pi/2$ and $-\pi \leq \omega < \pi$. The vector $\mathbf{x}$ originates in the coordinate center and sweeps out a closed surface in space when the two independant parameters, angles η and ω, change in the given intervals, ω is the angle between the x-axis and the projection of vector $\mathbf{x}$ in the x-y plane, while η is the angle between $\mathbf{x}$ and its projection in the x-y plane. Parameters η and ω correspond to latitude and longitude angles of the vector $\mathbf{x}$ expressed in spherical coordinates. Parameters $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ define the superquadric size in $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ coordinates respectively, ϵ_1 is the squareness parameter in the latitude plane and ϵ_2 is the squareness parameter in the longitude plane.

The shapes generated by various values of ϵ_1 and ϵ_2 are a superset of the standard constructive solid geometry (CSG) modeling primitives. Superquadrics can model a large set of standard building blocks, such as spheres, cylinders, parallelepipeds and shapes in



(a) (b) (c) (d) (e)

FIGURE 2.1. Shapes generated by varying the shape exponents of a Superquaric. (a) sphere ($\epsilon_1 = 1.0$, $\epsilon_2 = 1.0$) (b) cylinder ($\epsilon_1 = 0.1$, $\epsilon_2 = 1.0$) (c) cube ($\epsilon_1 = 0.1$, $\epsilon_2 = 0.1$) (d) parallelepiped ($\epsilon_1 = 0.1$, $\epsilon_2 = 0.5$) (e) bevel-shape ($\epsilon_1 = 1.0$, $\epsilon_2 = 2.0$).

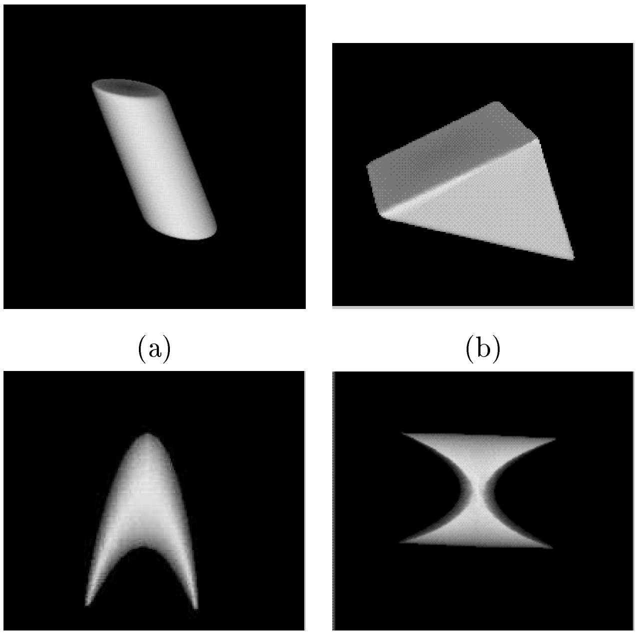
between indicated in Figure 2.1. When both ϵ_1 and ϵ_2 are 1, the surface vector defines an ellipsoid or, if $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are all equal, a sphere. When ϵ_1 is $\ll 1$ and $\epsilon_2 = 1$, the superquadic surface is shaped like a cylinder. Parallelepipeds are produced when both ϵ_1 and $\epsilon 2$ are $\ll 1$. Flat beveled shapes are produced when either ϵ_1 or $\epsilon_2 = 2$, and pinched shapes are produced when either ϵ_1 or $\epsilon_2 > 2$. Modeling capabilities of superquadrics can be enhanced by deforming them in different ways, including shearing, tapering, bending, and pinching as shown in Figure 2.2.

Based on the implicit equation of a superquadric surface, the *inside-outside function* [**31**] can be defined by the following function:

$$f(x,a) = ((x/a_1)^{2/\epsilon_1} + (y/a_2)^{2/\epsilon_1})^{\epsilon_1/\epsilon_2} + (z/a_3)^{2/\epsilon_2}$$

where if
$f(x,a) = 1$ then x is on the surface,
$f(x,a) > 1$ then x is outside the surface,
$f(x,a) < 1$ then x is inside the surface.

2.2. Exploration strategy Object models are constructed through a process of *autonomous exploration* [section 2.3]. Articulated volumetric descriptions are built up through



(a) (b) (c) (d)

FIGURE 2.2. Deformation modalities of a Superquaric. (a) a sheared cylinder (b) a tapered cube (c) a bent ellipsoid (d) a pinched sphere.

a sequence of exploratory probes. Figure 2.3 shows a block diagram of the resulting implementation. It corresponds to a classical model of bottom-up vision in which sensor data is transformed into various levels of representation through successive stages of processing [**9**].

These levels include:

(a) Sensor Control and Data Acquisition: Data acquisition is through a laser rangefinding system mounted on the end-effector of an inverted PUMA robot, the details of which will be provided in the Chapter 4.

(b) Visual Reconstruction: To facilitate the estimation of the transformation parameters relating different viewpoints and to provide the necessary structure from which to perform shape analysis, a visual reconstruction procedure is used to turn the discrete sampled data from the rangefinder into a piecewise-smooth (C^2) representations of the surfaces in the scene [**10**]. Data acquired from different vantage points can then be fused by determining the correspondence between features in adjacent views [**32**]. A temporal extension to the reconstruction procedure is used to determine the transformation parameters on a local basis. An important feature of this algorithm is its ability to deal with non-rigid motions.

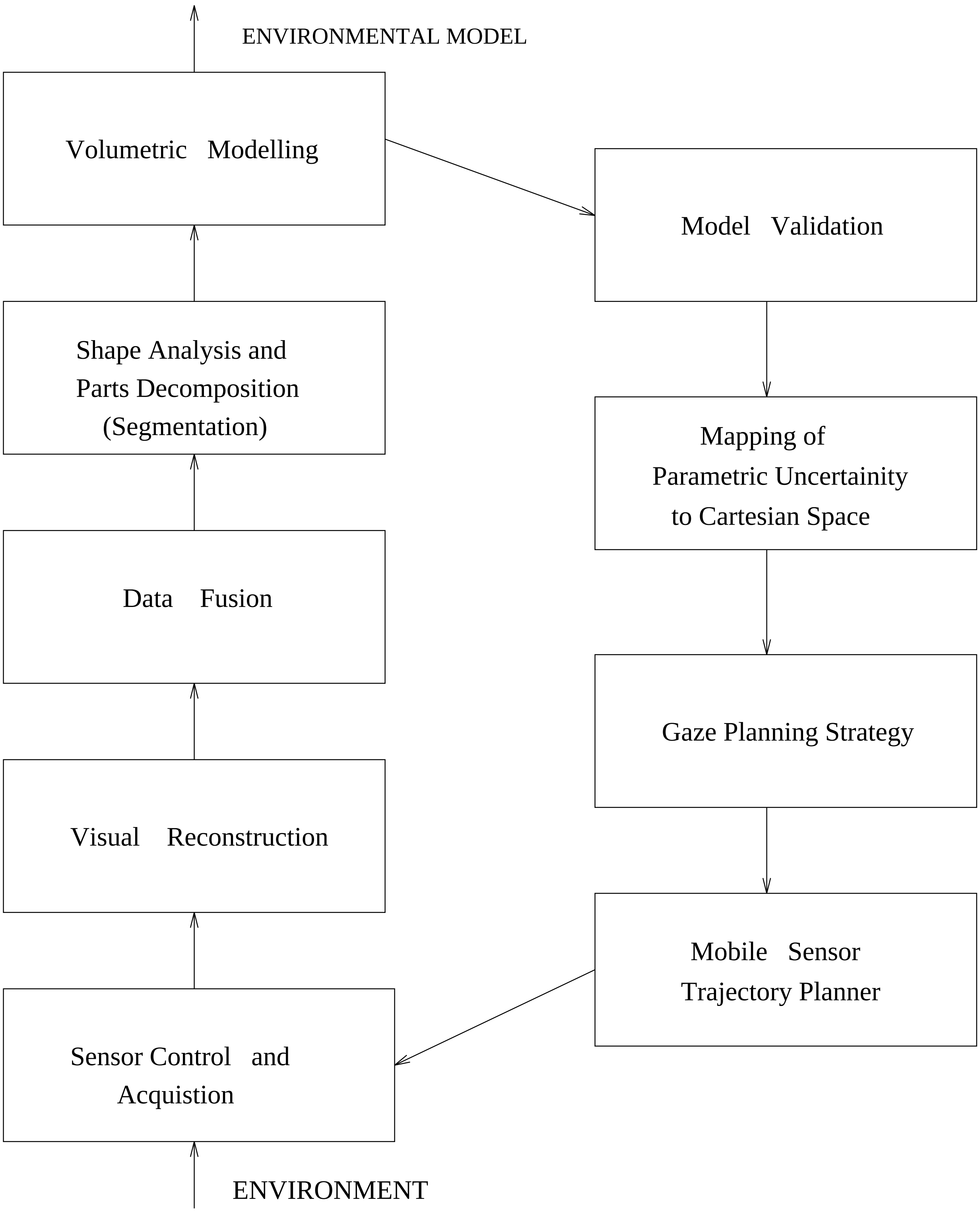
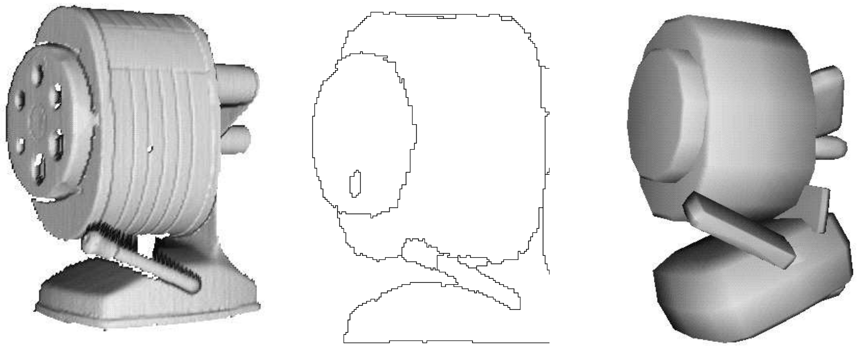


FIGURE 2.3. Process flow in the autonomous exploration strategy. The left hand corresponds to a classical bottom-up vision strategy. The right hand side corresponds to the feedback based on parametric uncertainity to close the loop. [from Whaite & Ferrie [**36**]]



(a) (b) (c)

FIGURE 2.4. Original Model of a pencil sharpener, Segmented parts corresponding to the object and Final Fitted Model using the superquadric representation. (a) shows a pencil sharpener (b) shows the segmented parts (c) shows the final fitted model using a superquadric representation. [from Lejuene [**14**]]

(c) Shape Analysis and Parts Decomposition: At the next higher level of abstraction, reconstructed surface information from multiple viewpoints is used to determine surface boundaries corresponding to the parts of an object. Objects in a scene are represented as conjunctions of convex solids. Boundaries between parts of objects thus correspond to concave discontinuities and/or negative local minima in the principal curvatures of the surface [**9**]. These features are made explicit as a by-product of the reconstruction procedure. However, the task of interpolating such features into part boundaries is non-trivial and the subject of much work in the literature, e.g. the work of Kimia *et al* [**13**]. The strategy employed by Lejuene and Ferrie [**14**] makes use of a special case of the more general model described by Kimia and has been adapted for surfaces. The algorithm developed in the strategy [**14**] is quite successful at partitioning range images into regions corresponding to parts for both man-made objects and natural forms as is evident from Figure 2.4.

(d) Volumetric Modelling: Finally, at the highest level of abstraction, surface regions defined by part boundaries are described by parametric forms (i.e. models) such as superquadrics [**31**]. In addition to serving as a basis for the characterization of uncertainity, these descriptors provide additional cues for maintaining correspondence at the level of parts, for describing general shape properties, and for recognition.

(e) Feedback: Just as feedback serves to reduce uncertainity in a conventional control system, the strategy uses feedback to minimize the parametric model uncertainity. Gaze planning is based on this uncertainity and is used to guide the sensor to the next viewpoint section in order to make the representation clearer.

2.3. The need for autonomous exploration A superquadric fit to the range data determines the model parameters. But the data might not sufficiently constrain the superquadric model, and as a result, numerous representations are possible which would fit the data accurately. This can be seen in the example in Figure 2.5 which shows the fitting process applied to points sampled from the surface of a noisy hemisphere. The models are fit to the data to within the same error of fit [**35**], yet the models look quite different. The problem is that the data acquired does not sufficiently constrain the model. This should not be surprising given that only part of the surface is visible in a given view. There should exist a way in which an observer could explore around the object to gather more data and come up with a better representation.

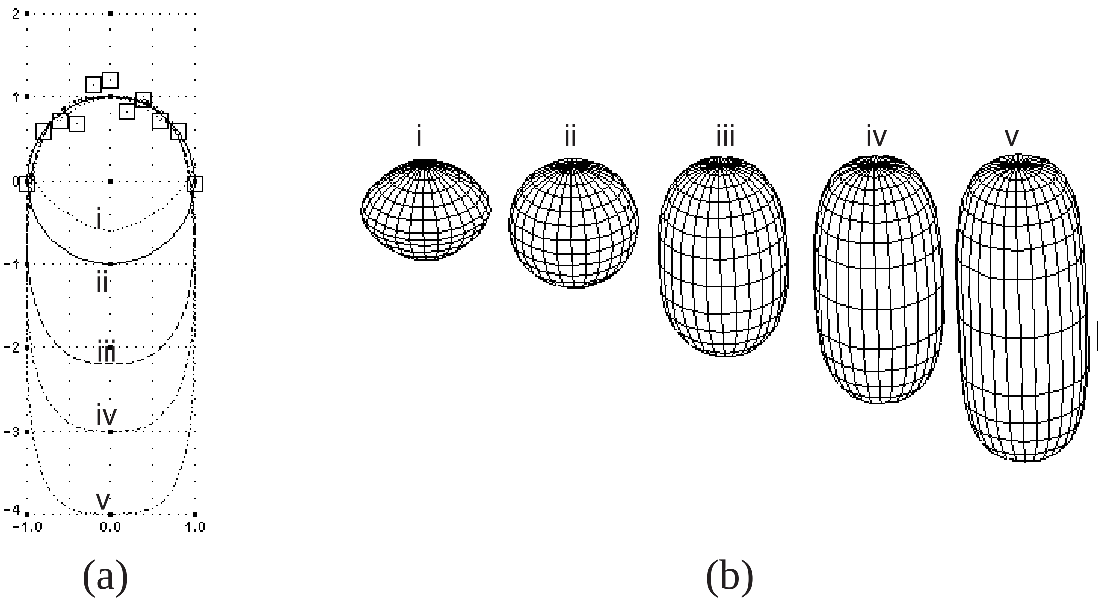


FIGURE 2.5. Model uncertainty. (a) Data points sampled from the surface of a noisy hemisphere. (b) Fits to the data using superellipsoid models. Each of the 5 models shown fits the data to within the same tolerance. [from Whaite & Ferrie [**36**]]

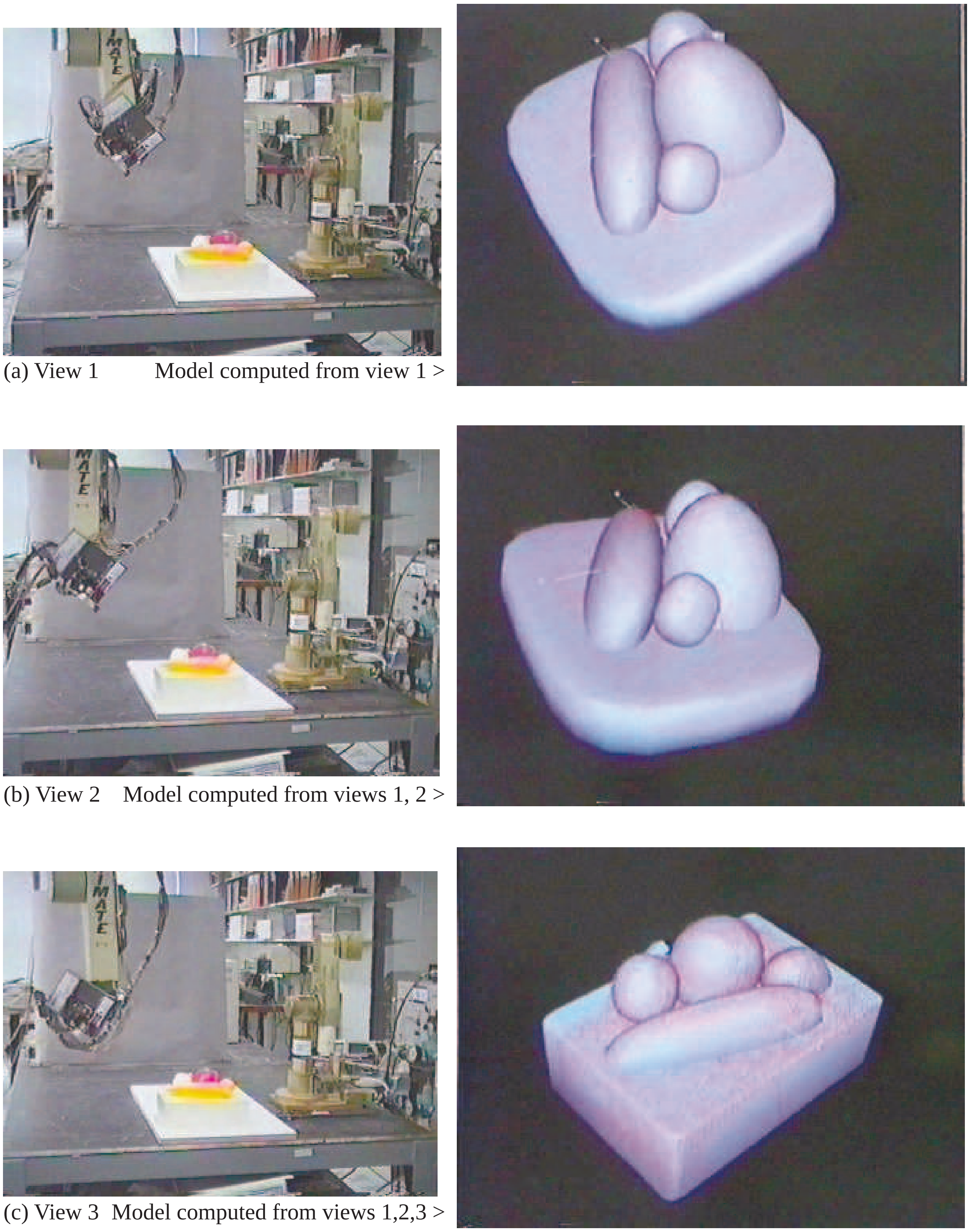
Whaite and Ferrie [**36**] have proposed such an autonomous explorer and have come up with representations which are acceptable. The system is entirely bottom-up and does not depend on any a priori knowledge of the environment. The development of the theory begins with the description of a system where the next location of a sensor is determined by a set of control parameters. The interaction between the sensor and its environment is modeled by the linear combination of a set of basis functions. A maximum likelihood estimate of unknown model parameters from a set of noisy measurements is found and the parameter covariances are used to calculate a measure of the model uncertainity. This leads to a gaze-planning strategy that uses the model uncertainity as a basis for selecting viewpoints. A representation, called an uncertainity surface, is computed from the model parameter covariance matrix. For a given model the gaze-planning strategy amounts to a gradient descent on the uncertainity surface. The next viewpoint is chosen to correspond to the direction on the view sphere which has the largest uncertainity gradient relative to the current position [**36**].

A working model of this explorer, Figure 2.6a shows a scene consisting of a wooden block on top of which is an assortment of 4 fruits. The system starts off by coarsely sampling the workspace until the block and fruit come into view. It then positions the rangefinder such that the block and fruit are at the center of its field of view. This defines the initial position shown in Figure 2.6a. The corresponding model determined from this initial viewpoint is shown immediately to the right. All 5 parts comprising the "object" are correctly localized, but there are significant errors in the positions and shapes of each part. This is about the best that can be expected given only a single view of the scene without additional constraint information.

Next, a second viewpoint is computed based on the uncertainty surface of the part closest to the rangefinder. This new viewpoint is shown in Figure 2.6b along with the model computed from the previous and current views. The additional data serves to further constrain the shapes of the fruit, but provides little additional constraint on the wooden block. Using the same part of the fruit as in the previous iteration, an uncertainty surface is computed and a third viewpoint determined (Figure 2.6c). This viewpoint brings the scanner low on the horizon from which more of the wooden block is visible. The resulting model, shown to the right, incorporates data from all three viewpoints and now correctly represents the shape of each part. It has been shown that the overall process is stable, producing near identical results when the initial positions of objects in the scene are perturbed [**36**].

The resulting model looks satisfactory but superellipsoids have limited expressiveness. The superquadrics do provide us with a good approximate model which can then be used for computing a more refined description using deformable superquadric models.

This exploration strategy provides us with a framework that can be used to determine a more refined deformable model that we seek. The idea will be to use the autonomous explorer to determine an initial set of model parameters at each acquisition step. Superellipsoids can be used for this purpose since they can be recovered reliably. They are subsequently used as the starting point from which a set of deformation parameters can be determined, leading to a more refined second order model. However the task of recovering



(a) View 1 Model computed from view 1 >

(b) View 2 Model computed from views 1, 2 >

(c) View 3 Model computed from views 1,2,3 >

FIGURE 2.6. Autonomous exploration sequence. (a) The initial viewpoint and corresponding scene model. (b) Second viewpoint determined from the uncertainties in the initial model. Scene model computed from fusion of data in the first and second views. (c) Final viewpoint and composite scene model computed from all 3 views. [from Whaite & Ferrie [**36**]]

these model parameters raises a number of concerns which are at the heart of the thesis - namely how the parameters are estimated, how the model is initialized, how the correspondence between data points and the model is determined, as well as the fitting procedure itself.

CHAPTER 3

Theoritical background

1. Introduction

The main emphasis of this chapter will be to characterize the accurate geometry of objects in a scene by inferring deformable models (as an improvement over the superellipsodial model) from 3D range data. Volumetric descriptions of this nature are useful because they can serve as a compact description of a scene while still retaining the information required by traditional higher level processes such as object recognition, path planning, or control of grasp.

Recently, researchers have addressed the problems of recovering volumetric part descriptions using deformable models. Two classes of such deformable models exist - those based on parametric solid modeling primitives [**19**], and those based on finite element based mesh-like surface models [**34**]. The description of shape by using orthogonal parameters produces a unique, compact description that is well-suited for recognition and database search but suffers from the lack of degrees of freedom to account for finer surface details. The deformable mesh approach, on the other hand is good at shape description, but produces descriptions which are neither unique nor compact, and consequently cannot easily be used for higher levels of processing such as object recognition. Both the approaches are slow, requiring hundreds of iterations in the case of a physically-based mesh formulation and dozens of iterations in the case of a parametric formulation.

These problems can be resolved to some degree by adopting an approach based on the finite element method (FEM) and parametric solid modeling using implicit functions [**29**]. The approach provides both the convenience of parametric modeling and the expressiveness of the physically based mesh formulation. In addition, it can provide much greater accuracy at physical simulation. The optimal parametrization is obtained from the eigenvectors of the object's finite element method (FEM) model. These eigenvectors are referred to as the 3D object's free vibration or deformation modes and together form a frequency-ordered orthonormal basis set for shape [**4**]. In this representation, objects are described using the metaphor of modeling clay, that is, shape is thought of as the result of pushing, pinching, and pulling on a lump of elastic material [**20, 19**]. The parametrization is unique, and is obtained by a multiscale orthonormal linear transform (similar to the Fourier transform) that maps the object's point-by-point motion in Cartesian coordinates into a coordinate system based on the object's intrinsic deformation modes [**23**]. The approach adopted in this thesis follows from the pioneering work done by Pentland & Sclaroff [**24**].

2. FEM Representation

The FEM is a standard engineering tool for simulating the dynamic behavior of an object. The motivation for using FEM in vision is that one is often concerned with estimating changes in position, orientation and shape which can be accurately described by FEM [**24**]. The dynamic behavior of the FEM can be used to solve fitting, interpolation, or correspondence problems by allowing the specification of forces that are a function of sensor measurements.

2.1. Background In the FEM the continuous variation of displacements throughout an object can be replaced by displacements at a finite number of nodal points as a function of impinging forces [**4**]. This can be expressed as:

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{D}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} = \mathbf{f} \tag{3.1}$$

where $\mathbf{u}$ is a $3n$ x 1 vector of the (x, y, z) displacements of the n nodal points relative to the objects' center of mass, $\mathbf{M}$, $\mathbf{D}$ and $\mathbf{K}$ are $3n$ by $3n$ matrices describing the mass, damping, and material stiffness between each point within the body, and $\mathbf{f}$ is a $3n$ x 1 vector describing the (x, y, z) components of the forces acting on the nodes.

Perhaps the major drawback of the FEM is its large computational expense, because the matrices $\mathbf{M}$, $\mathbf{D}$, and $\mathbf{K}$ are quite large: for instance, an object whose geometry is defined by 100 points produces 300 x 300 matrices, corresponding to the 300 unknown coordinates of the 100 nodal points, (x_i, y_i, z_i). In most situations, therefore, hundreds or thousands of very large matrix multiplications are required for each second of simulated time. Furthermore, for 3D models the computation scales as $O(n^3)$ as the number of points n defining the object geometry increases [**21**].

Another important problem with physically based methods is that all the degrees of freedom are coupled together. Consequently, closed-form solutions are impossible and solutions to the inverse problems encountered in vision are very difficult [**24**].

The approach followed by Pentland [**26**] is to diagonalize Equation 3.1 , so that the degrees of freedom will become uncoupled and thus give closed-form solutions. This is accomplished by *Modal Analysis* [**4**] .

2.2. Modal Analysis The modal method [**4, 24**] is adopted for the dynamic analysis of an object in order to achieve linear scaling in object complexity. In the modal approach an object's behavior is described in terms of its natural *strain* or *vibration* modes, rather than in terms of point-by-point displacements. These modes form the basis set over which the object dynamics is calculated. This basis set diagonalizes the mass, stiffness, and damping matrices so that the object dynamics may be described by linear superposition of the independent vibration modes. The modal method is equivalent to the FEM in expressiveness and accuracy, but has the additional virtue that it separates non-rigid object behavior into independent modes of deformation, each of which may be separately analyzed and (often) solved in closed form [**25**]. This in turn can lead to a much more efficient and stable computational scheme. The approach in this section is developed along the lines of [**4, 24**].

An object's strain modes may be found by simultaneously diagonalizing $\mathbf{M}$, $\mathbf{D}$, and $\mathbf{K}$. Because these matrices are normally positive definite symmetric, and $\mathbf{D}$ is a linear function of $\mathbf{M}$ and $\mathbf{K}$, Equation 3.1 can be transformed into $3n$ independent differential equations by use of the *whitening transform* [**4**], which is the solution to the following eigenvalue problem:

$$\lambda\phi = \mathbf{M}^{-1}\mathbf{K}\phi \tag{3.2}$$

where λ and ϕ are the eigenvalues and eigenvectors of $\mathbf{M}^{-1}\mathbf{K}$.

Using the transformation $\mathbf{u} = \phi\bar{\mathbf{u}}$ we can then re-write Equation 3.1 as follows:

$$\phi^T\mathbf{M}\phi\ddot{\bar{\mathbf{u}}} + \phi^T\mathbf{D}\phi\dot{\bar{\mathbf{u}}} + \phi^T\mathbf{K}\phi\bar{\mathbf{u}} = \phi^T\mathbf{f} \quad . \tag{3.3}$$

In this equation $\phi^T\mathbf{M}\phi$, $\phi^T\mathbf{D}\phi$, and $\phi^T\mathbf{K}\phi$ are diagonal matrices, so that if we let $\bar{\mathbf{M}} = \phi^T\mathbf{M}\phi$, $\bar{\mathbf{D}} = \phi^T\mathbf{D}\phi$, $\bar{\mathbf{K}} = \phi^T\mathbf{K}\phi$, and $\bar{\mathbf{f}} = \phi^T\mathbf{f}$ then we can write Equation 3.3 as $3n$ independent equations:

$$\bar{M}_i\ddot{\bar{u}}_i + \bar{D}_i\dot{\bar{u}}_i + \bar{K}_i\bar{u}_i = \bar{f}_i \quad , \tag{3.4}$$

where $\bar{M}_i$ is the i^{th} diagonal element of $\bar{\mathbf{M}}$, and so forth.

What Equation 3.4 describes is the time course of one of the object's *strain* or *vibration modes*. The constant $\bar{M}_i$ is the generalized mass of mode i, that is, the inertia of the i^{th} vibration mode. Similarly, $\bar{D}_i$, and $\bar{K}_i$ describe the damping and spring stiffness associated with mode i, and $\bar{f}_i$ is the amount of force coupled with this vibration mode. The i^{th} row of ϕ describes the *deformation* the object experiences as a consequence of the force $\bar{f}_i$, and the eigenvalue λ_i is proportional to the natural resonance frequency of that vibration mode.

To obtain an accurate simulation of the dynamics of an object, one simply uses linear superposition of these modes to determine how the object responds to a given force [**26**]. Either explicit, implicit, or (in simple situations) closed form solution techniques may be used to calculate how each mode varies with time.

There are several advantages of using this modal representation:

- It's efficiency which is typically one or two orders of magnitude more efficient and scales as $O(n)$ rather than the $O(n^3)$ scaling of the standard finite element method [**21**].
- A key aspect of the modal representation is that it provides a natural hierarchy of scale, so that we can smoothly vary the level of detail by adding in or discarding high-frequency modes.
- By establishing a correspondence (mapping) between the number of modes and the sensor measurements available, the shape recovery problem can be kept overconstrained and also relatively insensitive to noise. A shape description that is unique with respect to sampling and viewpoint can be calculated if the number of modes is less than the number of sensor measurements. In the modal representation, the high frequency modes that allow nodal points to move relative to one another have been discarded; and hence the representation is relatively insensitive to sampling and viewpoint. This makes the modal approach well-suited for object recognition and other spatial database tasks.
- The modal representation decouples the degrees of freedom within the non-rigid dynamical system of Equation 3.1, but it does not by itself reduce the total number of degrees of freedom. Once decoupled, however, we can separately analyze the various modes in order to determine which ones are required in order to obtain an accurate description of an object's non-rigid dynamic behavior [**20**].

The most important observation is that modes associated with high resonance frequencies (large eigenvalues) normally have little effect on object shape [**20**], the reason being:

- Typically, higher frequencies have small amplitudes, since the displacement amplitude for each mode is *inversely* proportional to the *square* of the mode's resonance frequency.
- Damping is proportional to the mode's resonance frequency. Thus higher frequency vibrations dissipate quickly.
- Frequencies are excited in proportion to the frequency content of the input force. The force generated by simple collisions typically have a roughly Gaussian time course, so that an individual resonance frequency f receives energy proportional to $e^{-f^2/\sigma}$. Thus low frequencies typically receive more energy than high frequencies.

Because of the above factors, one can conclude that the high-frequency modes have very little amplitude, and negligible effect. So most multi-body interactions can be adequately modeled by use of only rigid-body, linear, and quadratic strain modes [**25**].

Although discarding high-frequency modes generally has little effect on accuracy, it can have a profound effect on computational efficiency. The most obvious effect of discarding modes is to reduce the number of equations, with the result that only $O(n)$ operations and $O(n)$ storage locations are required (where n is the number of nodes defining the object geometry), as opposed to the $O(n^3)$ operations and $O(n^2)$ storage locations required when using standard methods. Indeed, when n is large the major cost of a low-frequency modal approach is collecting the point forces into modal forces by computing $\bar{\mathbf{f}} = \phi^T \mathbf{f}$, and then translating the modal displacements into point displacements by computing $\mathbf{u} = \phi\bar{\mathbf{u}}$ [**21**].

Another important benefit of using only low-order modes to describe object deformations is that they change very slowly as a function of object shape. Consequently the same modes ϕ may be used for a *range* of different - but similar - undeformed shapes without incurring substantial error. This allows the modes to be precomputed, avoiding the expense of solving for ϕ at run time [**24**].

3. Recovering Volumetric Models from the sensor measurements

The laser range finder gives us the sensor measurements (in the global coordinate system) that originate from the surface of the object being scanned. Mathematically, this can be represented by:

$$\mathbf{X}^w = [x_1^w, y_1^w, z_1^w, ..., x_m^w, y_m^w, z_m^w]^T. \tag{3.5}$$

The general approach to estimate an initial model estimate based on moments of inertia for the sampled range data. This provides the overall object position, orientation, and size. The next step is to establish the correspondence (mapping) between the sensor measurement points and the particular nodes on our deformable model, i.e., to determine a mapping between $\mathbf{X}$, points on the undeformed surface, and $\mathbf{X}^w$, the sensor measurements that specify the points' target position after deformation. This correspondence is one of the most problematic aspects in physically based modeling paradigm.

3.1. Recovery of Parametric Deformations Deformable superquadrics can be recovered using the same technique as for the recovery of nondeformed superquadrics. Some additional parameters describing deformations must be recovered as well that will provide the transformation from superellipsoids to deformable models. Any shape deformation can be recovered provided an inverse transformation is available such that x, y, z components of the nondeformed superellipsoid can be expressed in terms of X, Y, Z components of the deformed superellipsoid and the necessary deformation parameters [**31**].

A shape deformation is a function D which modifies the global coordinates of points in space

$$\mathbf{X} = \mathbf{D}(\mathbf{x}) \tag{3.6}$$

where $\mathbf{x}$ are the points on the undeformed solid and $\mathbf{X}$ are the corresponding points after deformation. Both $\mathbf{x}$ and $\mathbf{X}$ are expressed in the object-centered coordinate system.

The modal representation of shape can be combined with analytic shape primitives (e.g. superellipsoids) by first describing each mode by an appropriate polynomial function, and then using global deformation techniques [**3**] to warp the shape primitives into the appropriate overall form. The modes have been classified by the complexity of the associated deformation, e.g., as 0^{th} order (rigid body) modes, 1^{st} order (linear deformation) modes, 2^{nd} order (quadratic deformation) modes as illustrated in the Figure 3.1 [**25**].

The polynomial deformation mappings that correspond to each of the modes are determined by a linear regression of a polynomial with m terms in appropriate powers of x, y, and z, against the n triples of x, y and z that compose ϕ_i, a $3n$ x 1 vector containing the elements of the i^{th} row of ϕ,

$$\alpha = (\beta^T \beta)^{-1} \beta^T \phi_i \qquad , \tag{3.7}$$

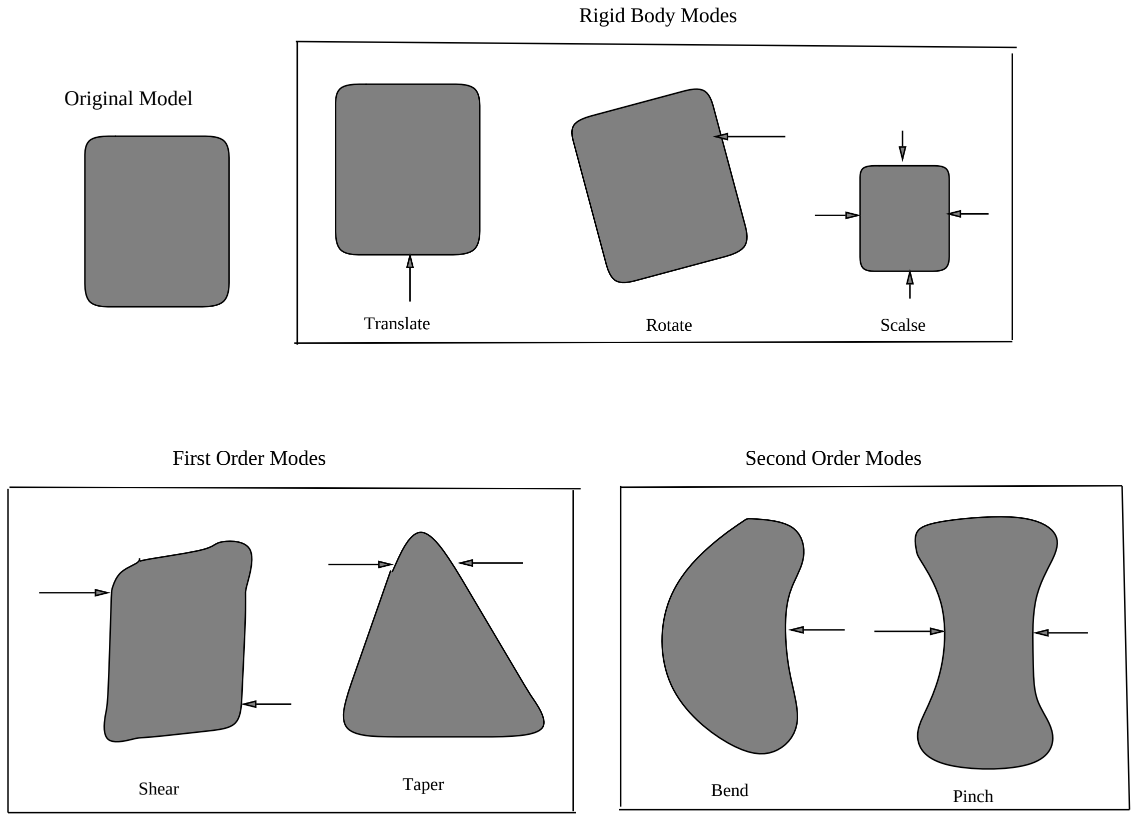


FIGURE 3.1. The lowest-frequency vibration modes of a object

where the i^{th} row of ϕ describes the deformation the object will experience as a consequence of the modal displacement $\mathbf{u}$, α is an m x 1 matrix of the coefficients of the desired deformation polynomial, β is an $3n$ x m matrix whose first column contains the elements of $\mathbf{u} = (x_1, y_1, z_1, x_2, y_2, z_2, ...)$, and whose remaining columns consist of the modified versions of $\mathbf{u}$ where the x, y, and/or z components have been raised to the various powers, e.g.,

$$\beta = \begin{pmatrix} x_1 & x_1^2 & x_1 & x_1 & \dots \\ y_1 & y_1 & y_1^2 & y_1 & \dots \\ z_1 & z_1 & z_1 & z_1^2 & \dots \\ x_2 & x_2^2 & x_2 & x_2 & \dots \\ y_2 & y_2 & y_2^2 & y_2 & \dots \\ z_2 & z_2 & z_2 & z_2^2 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \quad . \tag{3.8}$$

The polynomial deformations are combined into a 3 by 3 matrix of polynomials, **D**, referred to as the modal deformation matrix [**25**]. This matrix transforms the pre-deformation point positions $\mathbf{X} = [\mathbf{x}, \mathbf{y}, \mathbf{z}]^T$ into deformed positions

$$\mathbf{X}^* = \begin{pmatrix} d_{00} & d_{01} & d_{02} \\ d_{10} & d_{11} & d_{12} \\ d_{20} & d_{21} & d_{22} \end{pmatrix} \mathbf{X} \quad . \tag{3.9}$$

For further discussion refer to the Appendix 1. The question of which polynomial powers are appropriate for a particular column of ϕ can normally be decided by inspection (noting that the order of the deformation is related to the associated eigenvalue). The result is a polynomial model of unit amplitude deformation associated with mode i. The effect of the i^{th} mode on the overall object shape is then captured by simply scaling the i^{th} polynomial deformation according to the mode's amplitude $\bar{u}_i$ [**25**]. Once the deformed superquadrics have been recovered, the next step is to establish the correspondence between the data points and the nodes on our deformable model which is discussed in the next section 3.2.

3.2. Establishing Correspondence by Attachment Points The problem of correspondence is the matching of data points to corresponding attachment points for fitting a model to point data. To establish the correspondence between sensor measurements and the nodes of the undeformed model, the sensor measurements are projected onto the elliptical coordinate system [section 2.2] by a projection function, P(x), thereby determining the latitude ω and longitude η of each data point.

If $(\bar{x}, \bar{y}, \bar{z})$ is the coordinate of the sensor measurement in the elliptical coordinate system, we can find ω by:

$$\frac{\bar{y}}{\bar{x}} = \frac{cos\eta \; sin\omega}{cos\eta \; cos\omega} = tan \; \omega \tag{3.10}$$

Thus,

$$\omega = atan(\frac{\bar{y}}{\bar{x}}). \tag{3.11}$$

Given ω, η is determined by either

$$\eta = atan(\frac{\bar{z} \; sin\omega}{\bar{x}}) \tag{3.12}$$

or

$$\eta = atan(\frac{\bar{z} \; sin\omega}{\bar{y}}), \tag{3.13}$$

depending on whether $\bar{x}$ or $\bar{y}$ is larger.

This projection implicitly defines a correspondence between the undeformed surface points and their desired positions after deformation. When the number of data points is large it is more efficient to project the data onto a predetermined grid of undeformed surface points.

3.3. Fitting Procedure The fitting can proceed to recover the deformation modes once point correspondences have been established. The fitting methodology follows from [**29**]. The task is to deform the original undeformed point positions to their deformed positions. The fitting results in the recovery of the deformed implicit function [section 1] which best fits the data points.

The fitting stage begins with the formulation of the ϕ matrix which defines a coordinate transformation that diagonalizes the object's finite element equations. The effect of each of the deformation parameters u in the modal deformation matrix $\mathbf{D_u}$ on the position of the undeformed points $\mathbf{X}$ is calculated to obtain a m x 3n matrix ϕ whose i^{th} column ϕ_i is

$$\phi_i = (\frac{\delta x_1}{\delta u_i}, \frac{\delta y_1}{\delta u_i}, \frac{\delta z_1}{\delta u_i}, ... \frac{\delta x_n}{\delta u_i}, \frac{\delta y_n}{\delta u_i}, \frac{\delta z_n}{\delta u_i})^T \tag{3.14}$$

ϕ is determined by applying a small amount of each deformation u_i and measuring the resulting change in the coordinates of each point.

The matrix ϕ is the Jacobian of $\mathbf{D_u}$ at each point in $\mathbf{X}$ and its columns are approximately orthogonal. The deformation parameters (modes) are recovered using weighted least squares. The weighted least squares is solved using Singular Value Decomposition [**28**] which can be used to solve a possibly rank-reduced weighted least squares problem. The weights used should be inversely proportional to the measurement variance. Fitting is achieved by using the ϕ matrix in a modified Newton-Raphson iteration to obtain the minimum RMS error estimate of the deformation parameters u as:

$$\mathbf{u^{k+1}} = \phi^{-\mathbf{1}}(\bar{\mathbf{X}} - \mathbf{X^k}) + \mathbf{u^k}. \tag{3.15}$$

where $\mathbf{X^k}$ is the projection of the data points on the surface defined by the deformations $\mathbf{u^k}$ at iteration k , and $\mathbf{u^0} = 0$, $\mathbf{X^0} = \mathbf{X}$. A few iterations are often sufficient to obtain a satisfactory estimate of the deformation parameters $\mathbf{u}$. The ϕ matrix is usually not a square matrix, so the psuedoinverse ϕ^{-1} is used, where

$$\phi^{-1} = (\phi * \phi^T)^{-1}\phi^T \tag{3.16}$$

$\phi * \phi^T$ is only a square m x m matrix, so the calculation is numerically inexpensive.

The fitting algorithm proceeds by first establishing the correspondence between the model and the data points. Then the model is fit using the fitting algorithm (number of iterations can be controlled). The ϕ matrix is computed so as to linearize the model according to the calculated modes and loop until the modal parameters do not change significantly.

The fitting results in deformable models which fit the datapoints quite reliably. The problem is that unless the data has been uniformly sampled from all around the object, the model may not fit the data correctly. The methods described in this chapter suffer from this drawback. There should exist a way in which data points from all around the object can be collected so as to constrain the resulting deformable model to fit the data accurately. This can be achieved by making use of the autonomous exploration process which was introduced in Chapter 2. The next chapter will take these factors into account to incrementally recover deformable models in a dynamic fashion.

CHAPTER 4

Deformable model recovery in a dynamic strategy

1. Introduction

In this chapter, we integrate the ideas presented in the previous chapters. Usually, the methodolgy for fitting [**20, 33**] considers fitting models to data points which have been obtained from a single viewpoint of a camera (a laser-range finder, in our case). This research will incorporate data from other views of the object in a dynamic fashion in order to infer a volumetric deformable model. Pentland's work with deformable models is combined with an existing autonomous exploration strategy put forward by Whaite & Ferrie. The result is that now we can dynamically recover deformable models.

2. Methodology

The approach adopted in this thesis follows from the pioneering work done by Pentland & Sclaroff [**24**]. An initial approximate superellipsoid model is first recovered using a method based on [**31**]. An iterative fit to this model is achieved using a gradient descent procedure [**9**]. Correspondence is then established between the datapoints and the model [**24**]. Correspondence routines are used by Pentland & Sclaroff to establish which datapoints to attach to the model during fitting. We have developed routines which are an an improvement over the correspondence routines to account for the different kinds of datasets, for example, separate routines to deal with partial data, uniformly sampled sparse data, dense range data, etc. Fitting is then performed [**29**] to deform the initial model into a deformable one.

The major contribution in this research is that all the above mentioned steps will be performed within the framework of an active exploration strategy. An important consideration is the the stability & uniqueness of the deformable model parameters. Viewpoint independence, sampling, or sensor noise should not affect the resulting 3D description. We should be able to come up with stable & unique representations even when viewing the model from several different viewpoints.

Numerous experiments were performed to establish the stability of the underlying superellipsoid parameters and the deformation parameters. The experimental set-up consisted of the laser-rangefinder and the object on the rotary-table [Figure 4.9]. The exploration process was started using an initial view (defined by a latitude, longitude and the radius of the view sphere). Ten paths of the exploration were performed with each of the views being computed according to the uncertainity surface of the object closest to the rangefinder [**35**]. Five trials of exploration were performed by starting the exploration from a different initial viewpoint at each trial. The models were initialized after ten steps of the exploration and five more steps of the exploration were performed subsequently.

Stability of the model parameters during the model recovery is a major concern. A parameter is considered to be stable if its value with gaussian noise added falls within 10% of its value before any noise was considered. The assumption is valid for all the superellipsoidal and deformable parameters as a change of 10% does not cause considerable change to the model. Figure 4.1 will illustrate parameter variation of models by 10%. It is evident that the model shows negligible change visually.

2.1. Initial Fit to an Ellipsoid Model The model recovery process using an iterative fitting procedure will converge to an acceptable solution provided it has a good initial estimate of the rotation and translation parameters of the volumetric primitive. Only a rough estimate of the extent parameters is required [**9**]. The shape parameters are not critical and can be initialized so that the superquadric starts as an ellipsoid ($\epsilon_1 = 1.0$, $\epsilon_2 = 1.0$). The general methodology in this section follows from [**31**].

The position of the initial estimate for the superquadric is set to the center of gravity of all **N** range points:

$$p_x = \bar{x};\ p_y = \bar{y};\ p_z = \bar{z}. \tag{4.1}$$

An important point to be considered here is that since only the front side of the object is visible to us, we must employ a symmetry assumption to determine the z coordinate. After the center of gravity is found, we compute the orientation of the object centered coordinate

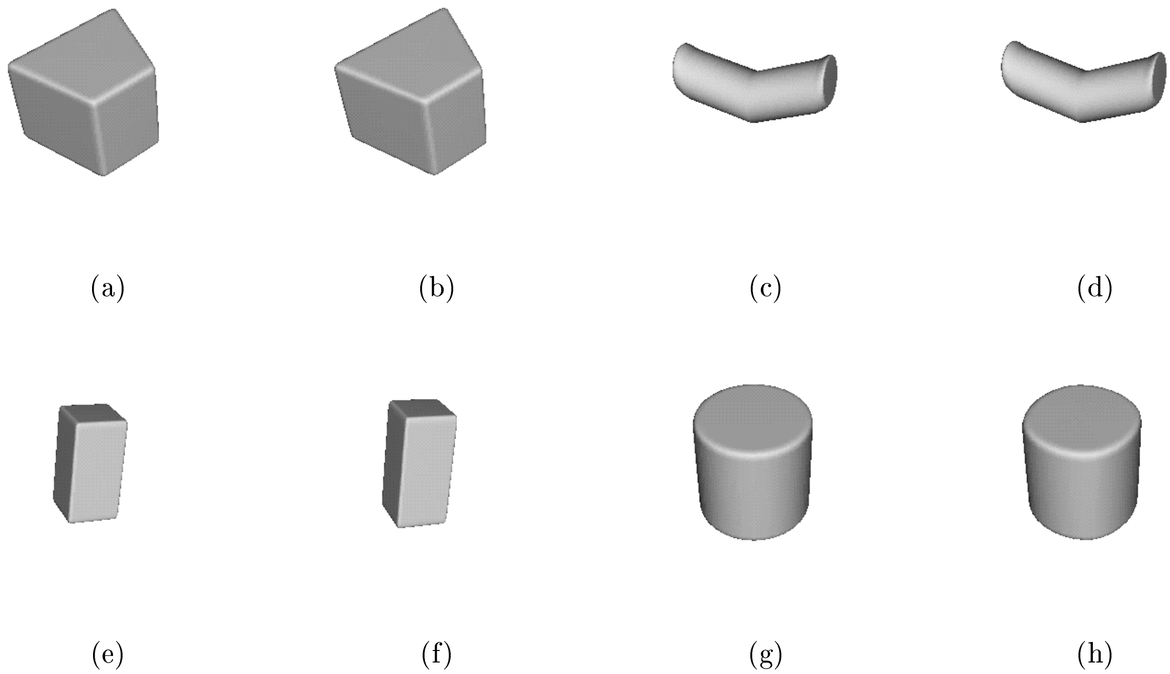


FIGURE 4.1. Superellipsoid and deformable parameter variation. (a) shows tapered cube. (b) shows the same cube with tapering (txy) increased by 10%. (c) shows bent cuboid. (d) shows the same cuboid with the bending (byz) increased by 10%. (e) shows a cuboid with rotation, translation, size ($rz = 20$, $tx = 20$, $ay = 50$). (f) shows the same cube with the parameters varied by 10% (g) shows a cylinder with shape exponents ($\epsilon_1 = 0.1$, $\epsilon_2 = 1.0$). (h) shows the same cylinder with shape exponents increased by 10%.

system. This is assumed to be the same as the axes of inertia. So we first compute the matrix of central moments or the inertia matrix. The inertia matrix for sample range data consists of the moments of inertia and the cross products of inertia:

$$\mathbf{I} = \begin{pmatrix} (y_i-\bar{y})^2+(z_i-\bar{z})^2 & -(y_i-\bar{y})(x_i-\bar{x}) & -(z_i-\bar{z})(x_i-\bar{x}) \\ -(x_i-\bar{x})(y_i-\bar{y} & (x_i-\bar{x})^2+(z_i-\bar{z})^2 & -(z_i-\bar{z})(y_i-\bar{y}) \\ -(x_i-\bar{x})(z_i-\bar{z}) & -(y_i-\bar{y})(z_i-\bar{z}) & (x_i-\bar{x})^2+(y_i-\bar{y})^2 \end{pmatrix} \tag{4.2}$$

Central moments are the moments with respect to the center of gravity $(\bar{x}, \bar{y}, \bar{z})$. We want to find a rotation matrix $\mathbf{R}$ which makes the matrix of moments $\mathbf{I}$ diagonal. Rotation matrix $\mathbf{R}$ that orients the object centered coordinate system along the axis of minimum and maximum inertia can be assembled out of the eigenvectors of matrix $\mathbf{I}$. Eigenvector e_1 with the smallest eigenvalue λ_1 corresponds to the minimum-inertia line and the eigenvector e_3 with the largest eigenvalue to the maximum inertia line. The minimum-inertia line is also known as the principal axis. The Jacobi method is used for computing the matrix of eigenvectors which consists of a sequence of orthogonal similarity transformations designed to annihilate one of the off-diagonal matrix elements. This method is foolproof for all real symmetric matrices which is the case with matrix $\mathbf{I}$.

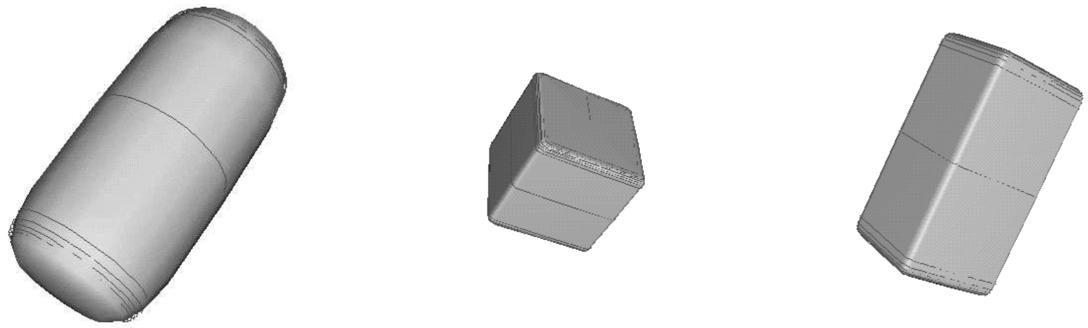
Now the object centered coordinate system is oriented so that the new axis z lies along the longest side for elongated objects (axis of least inertia) and along the shortest for flat objects (axis of largest inertia), based on the assumption that bending and tapering deformations normally affect objects along their longest side. From the elements of the rotation matrix $\mathbf{R}$, we compute the equivalent Euler angles. The size of the initial etimate is simply the distance between the outermost range points along each coordinate axis of the new object centered coordinate system.

The result is an initial ellipsoid estimate with three approximate estimates of position, orientation, and size. The method works very well with static range data points spanning the surface of the object in 3D. In such cases, usually the initial estimates are very close to the actual parameter values. In the course of this thesis, the same method is used to give an initial model estimate for range data points in a active exploration process [section 6]. Data points from different viewpoints are continually being added to the initial dataset, thereby refining the initial estimate each time. So, with the addition of new datapoints, the initial superellipsoidal model is continually changing. The method used in this section is quite capable of handling such events.

2.2. Iterative fit to a Superquadric Model To fit a superquadric model to datapoints, Bajscy [**31**] and Ferrie *et al.* [**9**] have used the Levenberg-Marquardt gradient descent procedure [**28**], an iterative least squares minimization technique that works as follows. A metric, $D(\mathbf{x}, \mathbf{a})$, is defined that measures the distance of each data point, $\mathbf{x}$, from the superellipsoid surface described by the parameters $\mathbf{a}$. From an initial guess [section 2.1] the parameters are incrementally changed in steepest descent fashion to minimize the squared sum

$$\chi^2(\mathbf{a}) = \sum_{i=1}^{n} \frac{\mathbf{D}^2(\mathbf{x}_i, \mathbf{a})}{\sigma_i^2} \tag{4.3}$$

of the metric over all the data points, $\mathbf{x}_i$. Ideally the data points are weighted by their measurement error, σ_i, so that those with less error have more importance. In our case the individual measurement errors are unknown so we set $\sigma_i = 1$ and estimate the total error once a fit has been found. The procedure repeats with the new parameters until there is negligible improvement in the squared error.



(a) (b) (c)

FIGURE 4.2. Uniformly sampled 3D data points and the intial estimated model (a) for a cylinder (b) for a cube (c) for a cuboid.

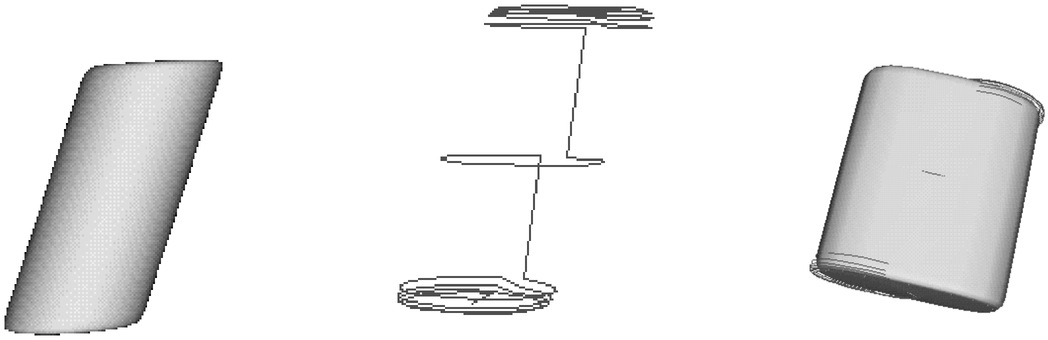
The result of the iterative fitting procedure is that for a superquadric in an arbitrary position 11 parameters are found: three estimates for the position, orientation, and size which define the elliptical coordinate system and two shape parameters ϵ_1 and ϵ_2 . This superquadric will serve as the initial guess of the undeformed implicit surface.

2.3. Initial shape estimation results In this section the algorithm developed for the sections 2.1 and 2.2 will be tested on simple geometric shapes such as cubes, cylinders and cuboids. The data used was simulated by a method discussed in section 4.3.

The algorithms give very good results for simple geometric shapes, shown in Figure 4.2. The problem arises when estimating a model which has non-rigid transformations such as shearing, tapering, bending, etc as is illustrated from the next example 4.3.

It is evident from the Figure 4.3 that though the initial model takes on the general shape of the data, it fails to get the sheared appearance of the cylinder. Evidently, the next step would be to recover deformation modes to apply to the initial estimated model to characterize this non-rigid transformation (shearing).

2.4. Correspondence Two types of correspondence algorithms were developed. Out of the cloud of data-points (given by the laser range-finder) the routines map the sample data points to the closest points on the deformable ellipsoid (i.e., only the closest data point to a node on the deformable model is selected). For two or three views of an object and with a collection of dense data points all around the object, this correspondence gives very good results. But for partial data (e.g. when only the front-facing information of an object is available), a distance threshold is used so that the the data points affect only the



(a) (b) (c)

FIGURE 4.3. A sheared cylinder, data points sampled from the figure and the intial estimated model (a) A cylinder sheared along the x-axis (b) data points sampled from the cylinder (c) Initial estimated model.

model points at overlapping regions and the rest of the model does not change to maintain global symmetry. This is achieved by weighting the data points in accordance to a distance threshold (i.e., if a data point is quite far away, then very minimal weight will be assigned to it). When a corresponding point isn't found for a particular surface (model) point, then it is assumed that the surface point doesn't move much (i.e., it matches itself) and is assigned a variance. The main drawback facing the algorithm is to deal with very sparse data. A correspondence may be found which may cause sideways stretching of the model causing the nodal points to move about on the object's surface, which may cause the model to deform badly out of shape. This is taken care of by modifying the algorithm leading to a second correspondence algorithm.

The second correspondence routine developed is similar to the above in that a single data point among the cloud of data points is attached to the closest node of the deformable model, but this time the data point selected is along the normal to the surface of the model. This method is used to account for the partial data cases.

Other methods of correspondence were tried as well in which the data points were projected along the normal to the surface followed by gaussian smoothing to average out the contribution of the points. The main drawback with this averaging method was that sharp, well-defined edges of objects like cubes, cuboids, etc were smoothed out during fitting due to the averaging of the data points.

Figure 4.4 illustrates the different cases of correspondence discussed in this section.

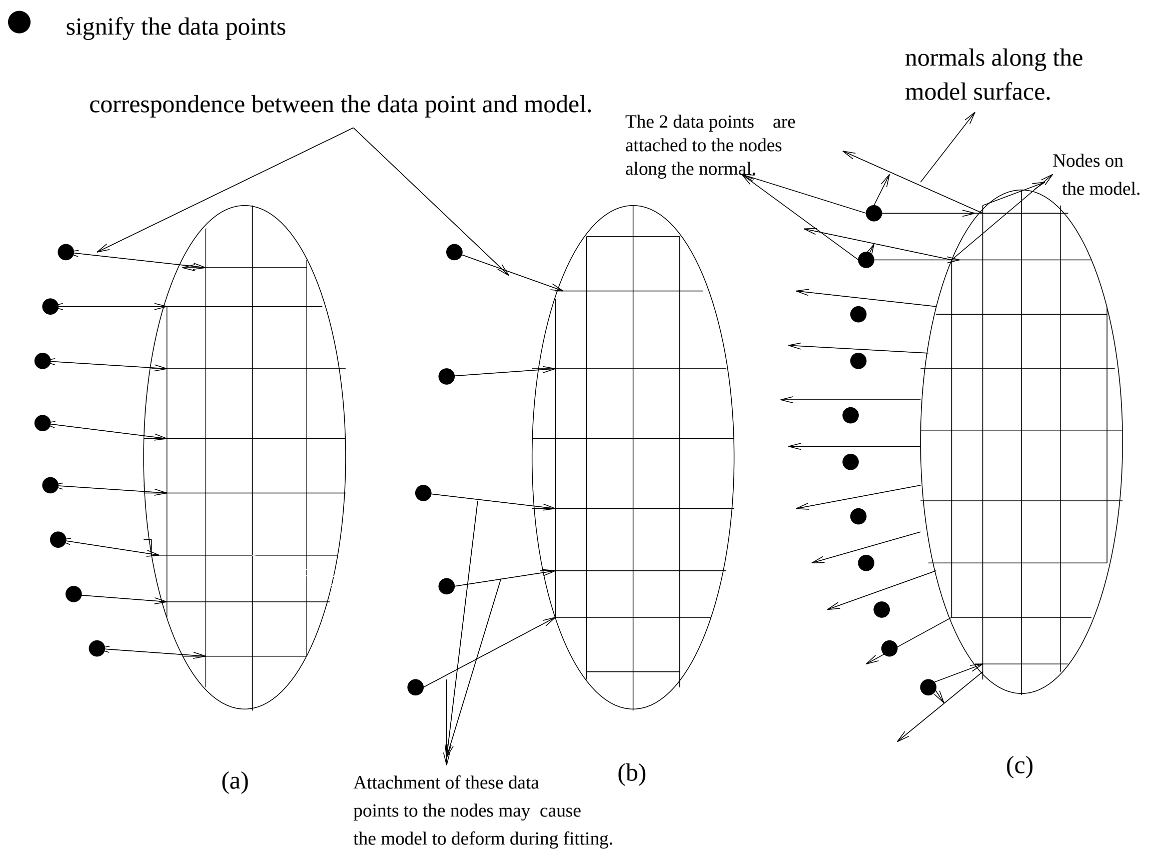


FIGURE 4.4. A graphical representation of the various types of correspondence routines (a) shows the case when a data point point is mapped to the nearest node on the model. (b) shows the case when correspondence between a sparse data point and the model may cause the model to deform out of shape during the fitting process. (c) shows the case when the the correspondence between the data point and the model was selected along the normal to the model.

3. Experimental set-up

This section will present a general overview of the experimental set-up used for gathering the range data in our laboratory. The general principles of the laser range finding system used, the limitations and the short-comings of the equipment used are also discussed.

3.1. Sensor Control and Data Acquisition Data acquisition is through a laser range-finding system mounted on the end-effector of an inverted PUMA robot as shown in Figure 4.6.

The range finder used for the experiments is the result of a collaboration between McGill University and the NRCC (National Research Council of Canada). It is a 2-axis scanner capable of scanning a complete surface from a fixed view point. The basic principle is explained in Figure 4.7.

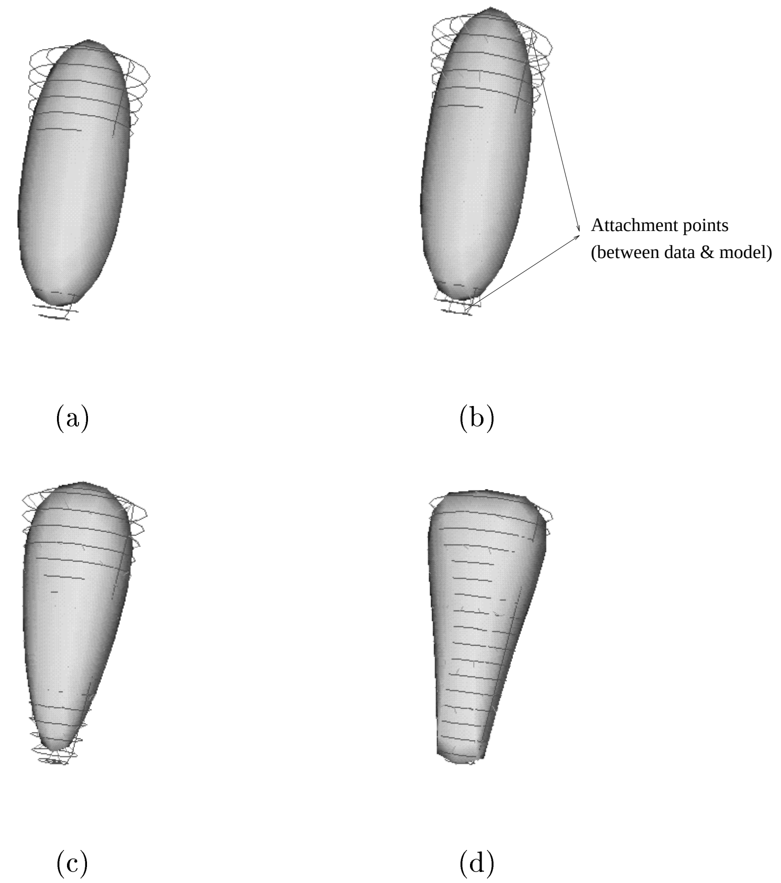


FIGURE 4.5. Initial model, Attachment points and Fitted model of an artificially generated tapered cylinder. (a) shows the range image (dotted lines) and the initial model fitting this range image data. (b) shows the attachment points between the data points and the model. (c) shows the model after 1 iterations. (d) shows the fitted model after 5 iterations.

The range of this scanner is of one meter with a field of view of 40 degrees in x and 28 degrees in y. Each of these directions (x and y) can be divided by a maximum of 256 positions. The precison is in the order of 1 mm at 1 meter of distance and improves in a non-linear fashion as the distance is reduced. A good technical description of the construction and the principle of this scanner can be found in [**6**].

It is possible to obtain calibrated data (real triplets x, y and z in mm for each point) in terms of the camera coordinate frame. For this purpose, a calibration procedure has to be applied carefully. The result of the calibration is a set of look-up tables and few simple geometric equations relating the indexes i and j of an image to their physical location (x and y in mm in the camera coordinate frame).

FIGURE 4.6. Mobile laser range-finding system used for data acquisition.

In addition, because of the angle between the outgoing and incoming rays, the scanner is sometimes unable to read a point position because the beam is blocked on its way back by an obstacle. This produces a shadow or a *hole* in the information about the surface.

3.2. Precision Stages A set of precision stages controlled by stepper motors is used to precisely move an object in front of the range finder. This allows for four degrees of freedom (2 rotations and 2 translations). This set-up is shown in Figure 4.8.

The theoretical precision obtained is about 79 steps per mm in displacement in x and in y, 100 steps per degree for the rotation about the z-axis and 0.56 step per degree for the rotation about the x-axis. However, the real precision is a bit lower if one takes into account the mechanical play in the gears (backlash).

In this set-up, multiple views of an object are obtained by keeping the camera fixed and by moving the object with the stages in front of the camera.

3.3. Acquisition of multiple-view data A robot (Puma 560) Figure 4.9 attached to the ceiling has also been used to move the camera around a rotary table to collect many views of an object.

In this chapter, we have integrated the Pentland's work with the deformable model recovery within the framework of an active exploration process [**35**]. Until now, most of the research has dealt with recovering models from single view datapoints. We have demonstrated a method by which deformable models can be incrementally recovered in a dynamic fashion. This chapter integrates the various methodologies and strategies proposed in this thesis. The sensor control and data acquisition process is discussed. Results of this reserach are presented in greater detail in the next chapter.

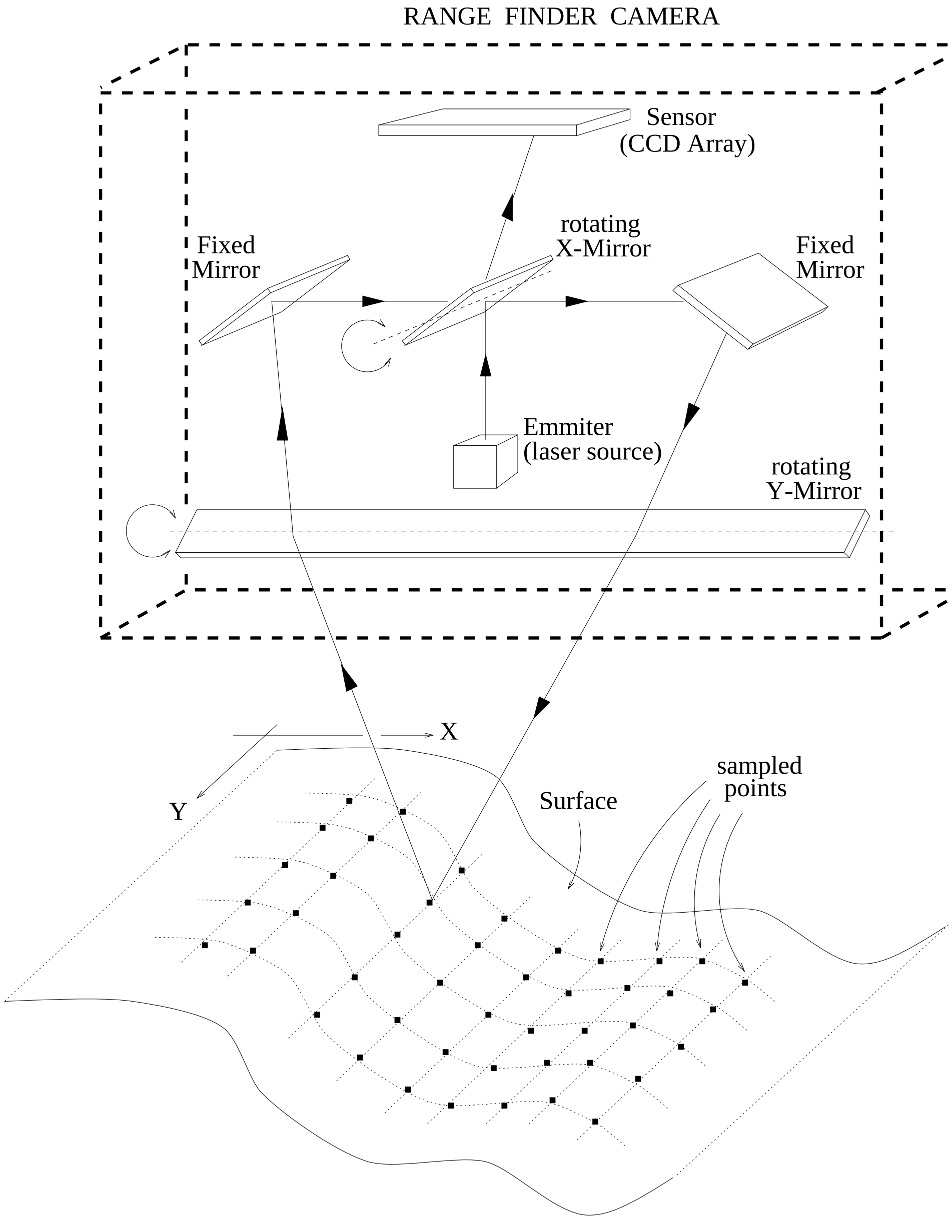


FIGURE 4.7. Schematic of the range finder used. A triangulation based technique is used to localize the hitting point of a laser beam after it has been reflected and diffused through a series of mirrors (y-mirror, fixed mirror and a relective x-mirror). The *detector* is in fact a linear CCD array which gives a reading function of the distance of the surface. By changing the angles of the x-mirror and y-mirror, it is possible to scan a 2 dimensional patch of the surface. [from [**32**]]

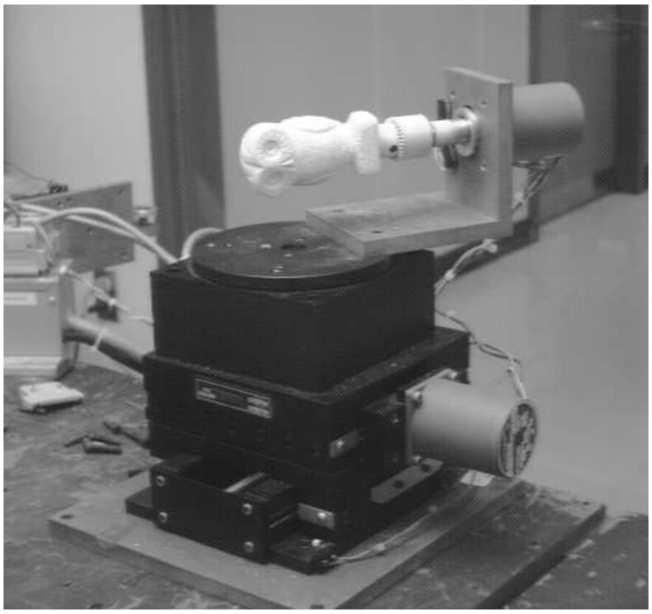


FIGURE 4.8. Precision stages used in the data acquisition process. These stages were used to move the objects in front of the range finder. The combination of stages allows for four degrees of freedom (2 displacements and 2 rotations).

FIGURE 4.9. Robot (Puma 560) hanging from the ceiling and used to move the range finder around the rotary table to gather multiple-view data.

CHAPTER 5

Results and Discussion

1. Introduction

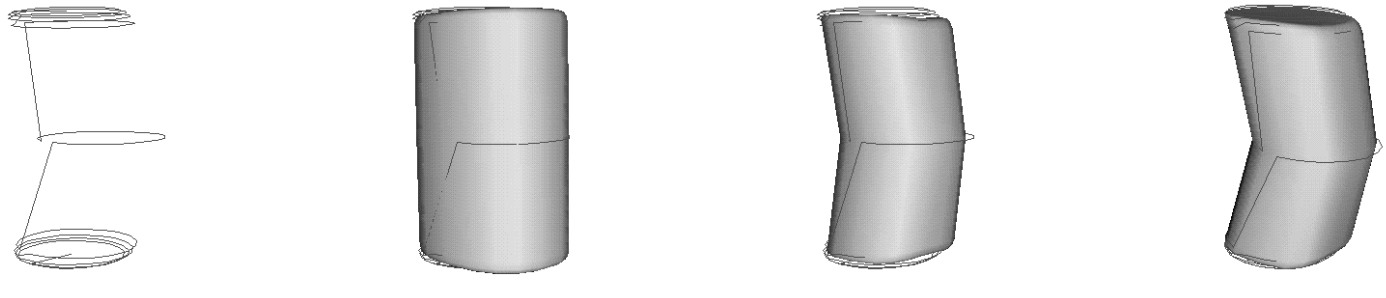
In this chapter, many different examples of recovering 3D volumetric, finite-element based models from range data by the methods proposed in this thesis are presented. For this purpose, examples with synthetic images, partial data range images and 3D range data will be presented. The need for additional views will be made evident from the model fitting to partial range data which will lead to experiments with complete 3D data sets. Autonomous exploration will be performed on a real object and issues related to model uniqueness and stability of parameters will be addressed.

2. Model fitting to artificially generated objects

The volumetric model recovery process will be first demonstrated on artificially generated objects. The data used was simulated by deriving a sampling equation (i.e., an η and ω generator which looks at the derivative of the superellipsoid equation along the isoparametric curve, and steps along the curve at constant intervals) which generates sample points in parameter space that are uniform in 3D.

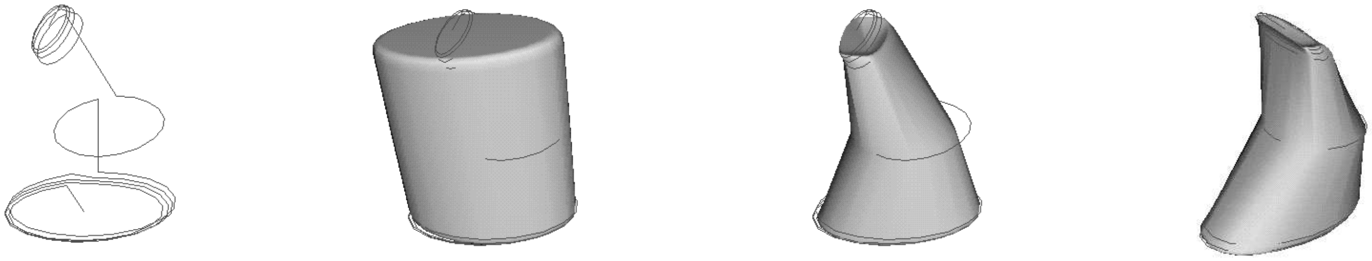
Figure 5.1 shows the fitting to a cylinder bent along the yz plane. The initial superellipsoid fit to the range data overfits the data as shown in Figure 5.1b. Correspondence is then established between a grid of surface points and the closest data points. In one iteration, the cylinder appears slightly bent and becomes more so after 5 iterations as shown in Figure 5.1c and Figure 5.1d. Table 5.1 shows the parameters during the fitting process. The parameters of interest are the bending parameters along the x axis (yx plane) and z axis (yz plane) which contribute to the bending.

Similarly, from Table 5.2, the bending parameters along z axis (yz plane) contribute to the bending. It is evident from the Figure 5.2 that more bending has been established. 10%



(a) (b) (c) (d)

FIGURE 5.1. Various stages of fitting of an artificially generated cylinder bent along the yz plane. (a) shows the range image (dotted lines) of a bent cylinder. (b) shows the initial model fit to this range image data. (c) shows the model after 1 iteration. (d) shows the model after 5 iterations.



(a) (b) (c) (d)

FIGURE 5.2. Various stages of fitting of an artificially generated cylinder tapered and bent along the yz plane. (a) shows the range image (dotted lines) of a bent cylinder. (b) shows the initial model fitted to this range image data. (c) shows the model after 1 iteration. (d) shows the model after 5 iterations.

Gaussian noise has been added to the range data and the perturbed parameters are shown in Table 5.4. The model parameters are within acceptable limits (within 10%), showing the stability of the fitting procedure. Similar examples of fitting are demonstrated in Figure 5.3 and the resulting model parameters in Table 5.3.

3. Experiments with single-part object partial range data

In this section, results of the models fitted to partial data are shown. The correspondence algorithm used for finding the attachment points was discussed in section 3.2. Figure

Parameter	Initial parameter extimates	Parameters after 1 iteration	Parameters after 5 iterations
TX, TY, TZ	0 -26.1 0	0 -31.3 0	0 -46.9 0
AX, AY, AZ	103.2 133 194.8	89.2 106.9 193.6	94.5 102.2 199.2
TBP (xy plane)	0 0 0	14.3 0 -4.1	13.2 0 5.3
TBP (xz plane)	0 0 0	0 0 5.4	0 0 -3.9
TBP (yx plane)	0 0 0	0 -24.7 3.7	0 -27.5 0.9
TBP (yz plane)	0 0 0	0 33.8 6.2	0 58.5 9.1
TBP (zx plane)	0 0 0	0 0 11.5	0 0 10
TBP (zy plane)	0 0 0	7.2 0 8.6	4.1 0 2.4

Above are the parameters of different model types over iterations of the fitting process. TBP refers to tapering, bending and pinc respectively. TX, TY, TZ refers to the translation along the x, y, z axis respectively. AX, AY, AZ refers to the x, y, z sizes respectiv Sampling = 0.39.

TABLE 5.1. Table showing the parameters of different model types over iterations of the fitting process.

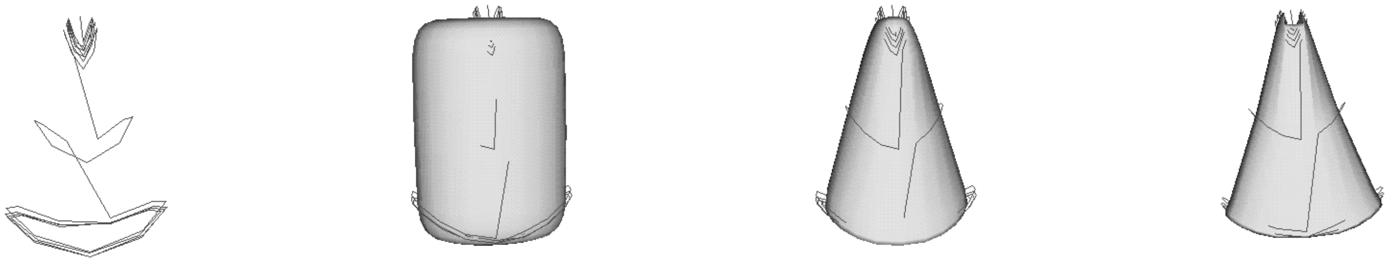
5.4 illustrates a pyramid model fitting which has been scanned from the top. It is evident from the figure that the initial superquadric model gets the general shape of the pyramid. During the deformable model fitting in the later iterations, the correspondence between the data points and the model is minimal. So the deformable model is similar to the inital superellipsoid model visually. Table 5.5 shows the tapering parameters of the pyramid. Tapering is considered as it is the parameter responsible for getting the shape of the pyramid. Even with addition of noise, the significant parameter (tapering) does not change much. This is evident from Table 5.5 as the parameter variation is within (± 10%).

Figure 5.5 shows a model fitted to the range data of a pear scanned from the top. The correspondence between the initial fitted model and the data points has to be established carefully so that the model does not deform badly out of shape during the subsequent model fitting process. This is necessary since we have no prior knowledge about the hidden surface of the object (i.e., in our case, the other side of the pear). Correspondence in this case is established only with the data points on the top of the model. The model points on the underside of the model are not attached to the data points, otherwise the model would deform to a thin plate during subsequent iterations. Table 5.6 shows the variation of the parameters of the pear during fitting. The deformable parameters do not vary significantly to affect the model. The pear is modeled by the pose and size parameters as shown in the table. Addition of noise varies these parameters within acceptable levels (± 10% parameter variation) that shows the stability of the model.

Cylinder bent and tapered along yz plane

Parameter	Initial parameter extimates	Parameters after 1 iteration	Parameters after 5 iterations
TX, TY, TZ	0 3.7 4.1	0 -11.5 8.8	0 -55.6 6.5
AX, AY, AZ	97.1 168.1 196.9	87.6 152.4 197.2	90.1 106.2 200.4
SX, SY, SZ	0 0 0	-0.4 0 0	-1.4 0 0
TBP (xy plane)	0 0 0	-3.7 0 -2.4	-4.9 0 3.7
TBP (xz plane)	0 0 0	8.2 0 4.6	12.5 0 -1.6
TBP (yx plane)	0 0 0	0 -6.3 -28.4	0 1.9 7.6
TBP (yz plane)	0 0 0	17.4 22.9 -1.8	70.1 66.4 -1.7
TBP (zx plane)	0 0 0	0 -7.3 6.3	0 -7.8 8.1
TBP (zy plane)	0 0 0	2.7 -5.3 4.5	4.8 -4.9 5.2

TABLE 5.2. Parameters for a Cylinder bent and tapered along yz plane.



(a) (b) (c) (d)

FIGURE 5.3. Various stages of fitting of an artificially generated cylinder tapered along yz plane and bent along the zx plane. (a) shows the range image (dotted lines) of a bent cylinder. (b) shows the initial model fitted to this range image data. (c) shows the model after 1 iteration. (d) shows the model after 5 iterations.

Figure 5.6 shows some examples of model fitting to some simple shapes. The Figures 5.4, 5.5, and 5.6 show that the model fitting is accurate over the portions of the surface covered by data points, but it begins to deteriorate at the boundary of the data near the "equator" because of the underside of the model, which remains too spherical due to the lack of the data points. Clearly, if more data is available from different vantage points, we can come up with a better model. Section 5 will show how the model improves as more views of the object are taken.

Cylinder bent (yz plane) and tapered (zx plane)

Parameter	Initial parameter extimates	Parameters after 1 iteration	Parameters after 5 iterations
TX, TY, TZ	0 0 7.6	0 0.8 -20.9	0 1.2 -36.7
AX, AY, AZ	94 158.3 214.4	83.9 129.8 204.2	97.6 99.6 205.5
SX, SY, SZ	0 0 0	0.2 0 0	-0.2 0 0
TBP (xy plane)	0 0 0	0 0 0.1	1.7 0 4.6
TBP (xz plane)	0 0 0	9.9 0 7.6	4.8 0 -3.5
TBP (yx plane)	0 0 0	0 -1.4 -23.4	0 -2.7 0.8
TBP (yz plane)	0 0 0	16.9 -0.7 7.6	73.4 -0.9 1.4
TBP (zx plane)	0 0 0	0 43.4 21.9	0 67.6 17.6
TBP (zy plane)	0 0 0	0.02 4.9 4	-0.6 7.9 3.5

TABLE 5.3. Parameters for a Cylinder bent (yz plane) and tapered (zx plane).

(a) (b) (c) (d)

FIGURE 5.4. Partial range data of a pyramid scanned from the top view and the various stages of fitting. (a) shows the range image (dotted lines) of the pyramid. (b) shows the initial model fit to the range image data. (c) shows the model after 1 iteration. (d) shows the model after 5 iterations.

4. Experiments with partial data of complex multi-part objects

This section shows the results of experiments with range images of articulated complex scenes which could consist of many parts. The segmentation [**14**] of the image is performed to break the image into its component parts and the volumetric 3D models of each part is recovered. The stability of the model is observed over a few iterations.

Bending parameters of the bent cylinder along yz plane.

B (yz plane)	0	33.8	58.5
B (yx plane)	0	-24.7	-27.5

With 10% addition of gaussian noise.

B (yz plane)	0.23	34.3	62.3
B (yx plane)	0.14	-26.5	-28.5

Bent and tapered cylinder along yz plane.

B (yz plane)	0	22.9	66.4
T (yz plane)	0	17.4	70.1

With 10% addition of gaussian noise.

B (yz plane)	0.5	23.6	71.4
T (yz plane)	0.43	15.4	75.6

Bent (zx plane) and tapered (yz plane) cylinder.

B (zx plane)	0	43.4	67.6
T (yz plane)	0	16.9	73.4

With 10% addition of gaussian noise.

B (yz plane)	0.03	38.6	68.7
T (yz plane)	0.4	17.4	70.2

T and B refer to tapering and bending respectively.

TABLE 5.4. Bending and tapering parameters & parameter variation with the addition of noise.

Figure 5.7 shows the fitting to the range data of a mushroom set. The inital model (a superquadric) is quite accurate at getting the 3D shape of the mushrooms. So, during the deformable model fitting, the the data points correspond to the model quite closely, thereby attachment is minimal. The result is that the deformable models look similar to the initial superquadric estimate.

The Figure 5.8 depicts the fitting to a group of fruits. The data sets in Figures 5.7 and 5.8 are both partial. The fitted models capture the structure of the data quite well in the

Tapering parameters of the pyramid along y axis (zy-plane and zy-plane).

T (zy plane)	0	19.6	24.3
T (xy plane)	0	22.9	25.4

With 10% addition of gaussian noise, the parameters are

T (zy plane)	0.23	22.34	24.9
T (xy plane)	0	19.9	27.3

The tapering parameters for the pyramid. The 3 values refer to the parameters of the initial model, after 1 iteration and after 5 iterations.

TABLE 5.5. Variation of the tapering parameters of the pyramid.

Translation, rotation and size parameters for the pear

TX, TY, TZ	5.1 9.1 109.16	4.65 9.6 113.45	3.3 5.4 112.8
RX, RY, RZ	-32.8 23.45 -132.9	-32.5 24.9 -143.7	-31.6 24.5 -149.6
AX, AY, AZ	63.6 57.8 121.0	64.3 59.7 124.65	62.3 63.4 126.5

With 10% addition of noise, the parameters are

TX, TY, TZ	4.76 8.3 99.8	5.14 8.7 120.3	3.1 4.9 120.8
RX, RY, RZ	-30.7 22.9 -127.4	-34.8 25.1 -165.2	-30 26.7 -154.7
AX, AY, AZ	67.8 56.6 120.7	69.8 63.4 120.6	60.6 62.7 136.5

Tapering, bending and pinching parameters of the pear

TBP (xy plane)	0 0.76 0.89	-1.3 -2.54 -2.49	-0.5 0.63 0.92
TBP (xz plane)	0.43 0.6 1.34	0 0.6 1.86	0 0.1 -2.3
TBP (yx plane)	0.3 0.7 1.53	-0.75 -1.56 0	-0.8 1.9 4.2
TBP (yz plane)	0.78 1.2 2.54	0 4.7 0.65	-0.5 0.97 0.9
TBP (zx plane)	0.8 1.0 1.49	0.2 3.2 1.54	0 1.2 0.2
TBP (zy plane)	0 0.2 0.69	-3.4 -0.9 0	2.43 0.67 1.6

The superellipsoid and deformable parameters for the pear. The 3 values refer to the parameters of the initial model, after 1 iteration and after 5 iterations.

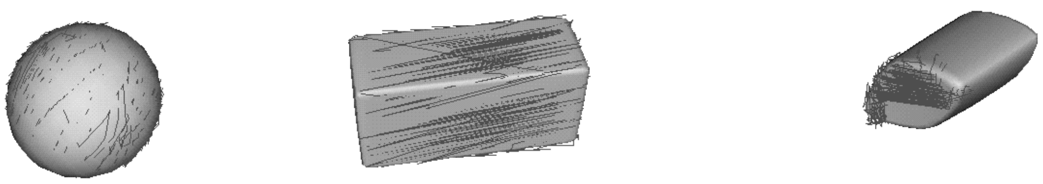
TABLE 5.6. Variation of the superellipsoid and deformable parameters for the pear.

case of the mushrooms. The algorithm breaks down for the data set in Figure 5.8. It is difficult to get the shape of the bananas since data-points corresponding to the banana are just for a single view and the correspondence routine has to be careful not to attach the data-points to the underside of the banana. This is taken care off by accumulating data from around the object as shown in Figure 5.14.

Some more examples are presented for partial multi-part range data-sets. The results are interesting because the model tries to come up with a description of shape even though it might be occluded by other parts of the same object as shown in Figure 5.9. The

(a) (b) (c) (d)

FIGURE 5.5. Partial range data of a pear scanned from the top view, the initial fitted model and various stages of the fitting process. (a) shows the range image (dotted lines) of the pear. (b) shows the initial model fit to the range image data. (c) shows the model after 1 iteration. (d) shows the model after 5 iterations.

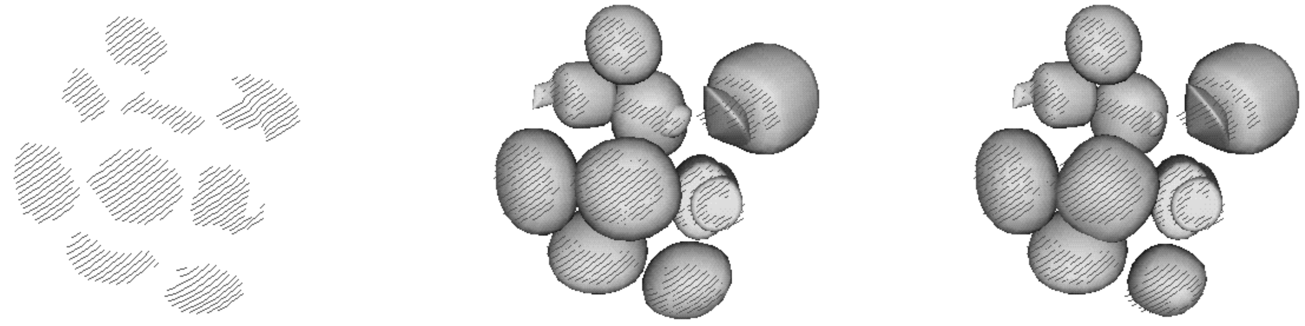


(a) (b) (c)

FIGURE 5.6. Partial range data of a sphere, cube and cylinder and the initial fitted model to these data. (a) shows the partial data of a sphere and the initial fitted model. (b) shows the partial data of a cube and the initial fitted model. (c) shows the partial data of a cylinder and the initial fitted model.

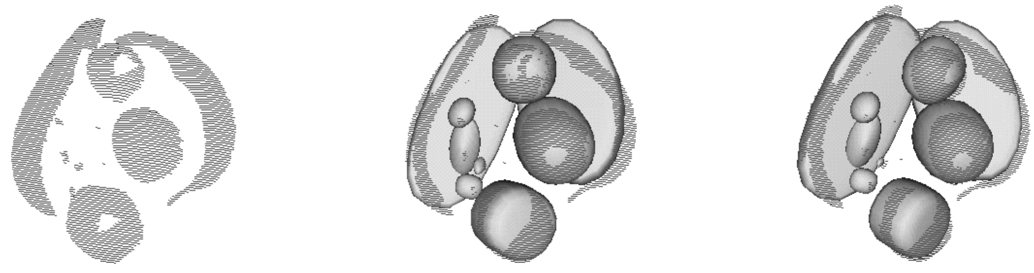
segmentation plays a big role in breaking up the object into its component parts since the number of data-points contibuting to a part might vary considerably, for example, data-points in the fingers and the torso of the doll in the Figure 5.10 are significantly different.

Figures 5.9 and 5.10 show the recovery of the volumetric 3D models to range data. The models obtained are quite accurate over the data points and capture the 3D structure of the object involved even though the range image is from a single view only. The results presented in this section intuitively suggest the need for additional views so that the complete 3D structure of an object can be captured by the fitted model.



(a) (b) (c)

FIGURE 5.7. 3D range data of a group of mushrooms, the initial superquadric and final deformable model . (a) shows the range data of the mushrooms. (b) shows the initial model fitted to the data. (c) shows the model after 5 iterations.

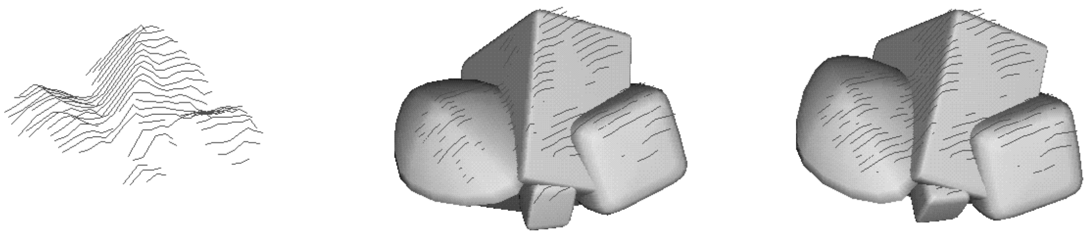


(a) (b) (c)

FIGURE 5.8. 3D range data of a group of an assortment of fruits , the initial superquadric and final deformable model . (a) shows the range data of the fruits. (b) shows the initial fitted model to the data. (c) shows the model after 5 iterations.

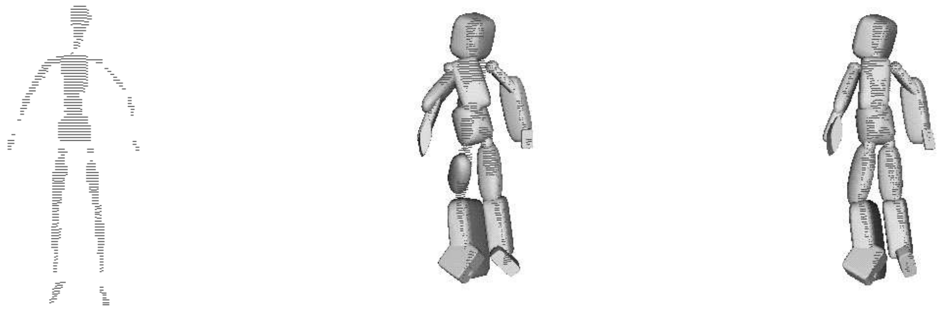
5. Results with 3D data sets

In this section, results of models fitted to the multi-view data are shown. Experiments are done with views at angles of 90, 180 and 270 degrees in addition to the initial view of the object. It is shown how the model improves over additional views. Mutiple view data is produced by a three-step procedure, called multiple view integration [**38**]. In the first step - *data acquisition* - range data from different views are collected as viewer-centered data descriptions specified in each camera coordinate system. In the second step - *view*



(a) (b) (c)

FIGURE 5.9. Range image of a pile of rocks, the initial superquadric and final deformable model. (a) shows the range data of the rocks. (b) shows the initial fitted model to this data. (c) shows the model after 5 iterations of the fitting procedure.



(a) (b) (c)

FIGURE 5.10. Range image of a doll, the initial fitted model and final fitted model after a few iterations. (a) shows the range data of the doll. (b) shows the initial fitted model to this data. (c) shows the model after 5 iterations of the fitting procedure.

registration - a transformation between a camera coordinate system and the world coordinate system is determined. In the last step - *view integration* - range data in each camera coordinate system are transformed in the world coordinate system and the redundant data seen in more than one view are removed. Data acquisition of the multiple views of an object is accomplished with a laser scanner which scans objects supported by a turntable as shown in the Figure 4.9. View integration is performed by using the view transformation, surface normals and residuals of the normal computation. The integrated data are expressed as a sequence of points in R^3.

Cuboid

Parameter	Initial parameter extimates	Parameters after 1 iteration	Parameters after 5 iterations
TX, TY, TZ	-11.7 -0.6 5.7	-10.6 -2.8 1.8	-7.7 -3.2 3.1
RX, RY, RZ	48.5 19.2 161.6	47.8 19.5 161.7	44.7 18.5 162.9
AX, AY, AZ	45.5 21.9 16.8	43.5 21.2 16.7	40.8 19.5 16.2
SX, SY, SZ	0 0 0	-0.1 0.3 0.1	-0.3 1.2 0.4
TBP (xy plane)	0 0 0	0.5 1.7 1.5	2.2 3.3 3.8
TBP (xz plane)	0 0 0	-0.4 0.5 0.8	0.3 0.4 2.8
TBP (yx plane)	0 0 0	-0.5 -4.3 0.3	-1.3 -2.8 4.0
TBP (yz plane)	0 0 0	-0.1 -0.4 -1.8	0.04 -0.6 -1.8
TBP (zx plane)	0 0 0	0.4 -0.24 -1.2	1.2 -2.7 1.4
TBP (zy plane)	0 0 0	0.7 0.8 -1.3	0.1 0.55 -0.6

TABLE 5.7. Table showing the variation of the superellipsoid and deformable parameters for the cuboid.

Addition of 10% noise

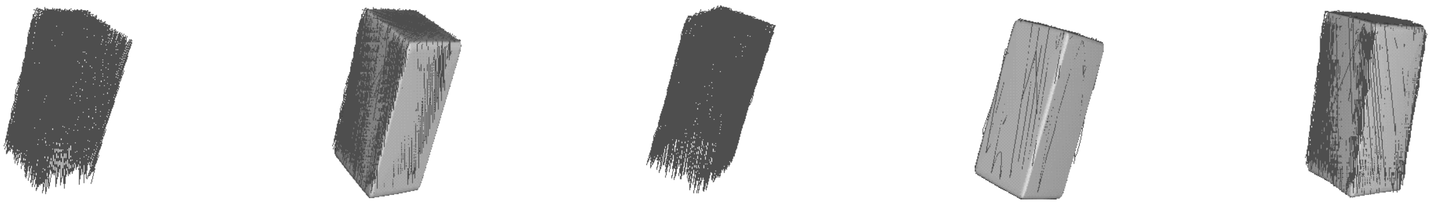
TX, TY, TZ	-10.9 -0.5 4.9	-10.9 -3.34 2.2	-6.9 -3.3 2.85
RX, RY, RZ	51.34 18.76 172.3	48.9 21.34 154.7	44.9 23.4 172.6
AX, AY, AZ	45.8 22.3 18	41.7 19.6 14.2	37.6 19.7 15.4

TABLE 5.8. Variation of the cuboid parameters with the addition of noise.

Figure 5.11a shows the range data of a cuboid. The initial model fit in Figure 5.11b fits the data points quite well. The figures that follow show how the fitted model improves as views are added. Figure 5.11d shows a model fitted to data points integrated from views taken of the cuboid at 0 degrees and 90 degrees. Figure 5.12b show the same model fitted to three views of the cuboid (from 0, 90, 180 degrees). Figure 5.12d shows the model fitted to 4 views (from 0, 90, 180, and 270 degrees). Table 5.7 shows that the variation of the deformable parameters is almost negligible. The shape of the cuboid is easily established by the size, shape, position and orientation parameters. The variation of these parameters over the fitting process does not change significantly (within $\pm$ 10% variation) even with the addition of noise to the range data as is evident from the Table 5.8. Thus the model is stable during the fitting procedure.

(a) (b) (c) (d)

FIGURE 5.11. Range data of a cuboid from views at 0 and 90 degrees and the initial fitted model. (a) Range data of a cuboid at the initial position of the object on the rotary table. (b) Initial fitted model to the data. (c) Range data of the cuboid integrated from views at 0 and 90 degrees to the object. (d) Initial fitted model to this data.



(a) (b) (c) (d) (e)

FIGURE 5.12. 3D range data of a cuboid from views at 180 and 270 degrees to the first view and the initial model fitted to the composite range data from the views. (a) 3D range data of a cuboid from a view at an angle of 180 degrees to the first view. The resultant view adds the range data obtained from all the 3 views (0, 90 and 180 degrees). (b) The initial fitted model to the above data. (c) 3D range data of a cuboid from a view at an angle of 270 degrees to the first view. The resultant view adds the range data obtained from all the 4 views (0, 90, 180 and 270 degrees) to give an almost 3D description of the object. (d) Initial fitted model to this data. (e) Model after a few iterations.

Figures 5.13 and 5.14 depict the fitting process applied to non-convex objects which are overfitted by the initial fitting algorithm. The fitting algorithm is quite successful at getting the bend and taper characterstics of the data and is also quite stable during the fitting process. Table 5.9 & 5 verify that the deformable parameters are stable after few iterations of the fitting process since the parameter variation is within (± 10%). Figure 5.13 is achieved by tapering along the x axis (yx and zx planes) whereas the bend characterstics of Figure 5.14 are achieved by bending along the z axis (along the yz-plane). Tables 5.10 and

Tapered cuboid

Parameter	Initial parameter extimates	Parameters after 1 iteration	Parameters after 5 iterations
TX, TY, TZ	-0.5 0.4 1.5	0.35 0.8 0.6	1.5 2.5 -0.9
RX, RY, RZ	-41.9 23.05 -149.5	-41.8 24.11 -149.6	-40.3 25.2 -149.8
AX, AY, AZ	36.6 19.3 19.4	36.1 16.7 17.02	36.2 15.9 16.6
SX, SY, SZ	0 0 0	-0.18 0.5 -0.04	0.12 0.6 0.08
TBP (xy plane)	0 0 0	0.08 0.6 -0.4	-0.09 1.6 -1.9
TBP (xz plane)	0 0 0	-0.4 1.5 -0.8	-0.3 2.9 -0.6
TBP (yx plane)	0 0 0	4.8 1.01 1.8	5.9 2.5 2.6
TBP (yz plane)	0 0 0	-0.11 -0.15 -1.6	0.6 0.48 -0.04
TBP (zx plane)	0 0 0	5.5 -0.5 0.5	6.5 -1.6 1.12
TBP (zy plane)	0 0 0	-0.08 -0.7 -1.3	-0.7 -1.01 -0.19

TABLE 5.9. Variation of the superellipsoid and deformable parameters for a tapered cuboid.

Addition of 10% gaussian noise

RX, RY, RZ	-42.3 24.21 -160.3	-42.3 22.9 -153.7	-39.8 27.3 -163.4
AX, AY, AZ	38 17.8 19.6	34.7 17.6 18.2	36.9 17 14.35
T (yx plane)	0.12	5.2	6.13
T (zx plane)	0	5.9	7.2

TABLE 5.10. Variation of the significant parameters for a tapered cuboid with the addition of noise.

5.12 illustrate the variation of the significant parameters with noise. Again, the parameter variation falls within acceptable levels (± 10% parameter variation), indicating the stability of the model.

The results presented in this section illustrate the fact that the model gets the shape of the object with the accumulation of data points from all around the object incrementally (Figure 5.13 and 5.14). What we would like to investigate next is if a similar improvement over the superellipsoid can be achieved by a deformable solid model using an active process. The active strategy we propose to use is the *autonomous exploration* process [**36**].

6. Autonomous Exploration results

This section will illustrate various experiments that were performed in our laboratory on a working autonomous explorer [**36**]. In this section, the autonomous explorer will be

Bent cuboid

Parameter	Initial parameter extimates	Parameters after 1 iteration	Parameters after 5 iterations
TX, TY, TZ	-2.4 -2.8 2.6	-3.5 -0.25 0.05	-3 1.5 -1.18
RX, RY, RZ	6.9 -65.8 -127.8	-0.3 -68.5 -134.4	4.2 -68.9 -135.6
AX, AY, AZ	44.6 16.5 7.37	42.9 14.17 7.26	44.9 16.9 7.33
SX, SY, SZ	0 0 0	-0.4 -0.09 -0.1	-0.9 -0.07 0.06
TBP (xy plane)	0 0 0	-3.4 0.3 -2.6	-4.6 1.116 -6.6
TBP (xz plane)	0 0 0	0.3 -0.37 0.5	0.9 -0.08 0.07
TBP (yx plane)	0 0 0	-0.08 -3.6 -1.4	-0.5 -4.2 -8
TBP (yz plane)	0 0 0	1.2 5.28 0.57	2.06 7.12 -0.06
TBP (zx plane)	0 0 0	0.12 0.18 0.7	0.04 1.16 0.44
TBP (zy plane)	0 0 0	-0.02 0.35 0.39	-0.27 0.5 1.27

TABLE 5.11. Table shows the variation of the superellipsoid and deformable parameters for a bent cuboid.

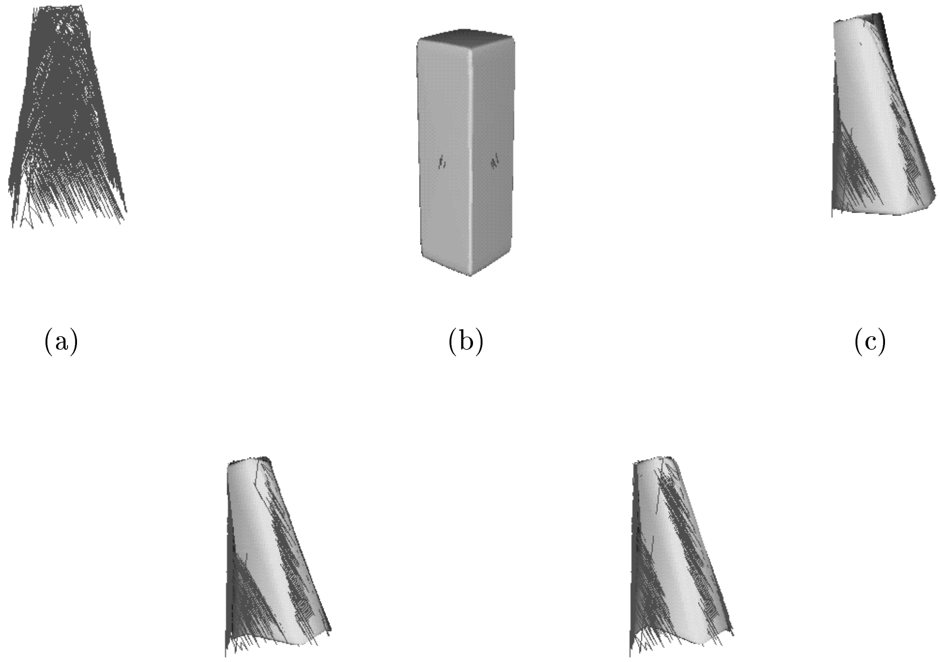
Addition of 10% gaussian noise

RX, RY, RZ	7.2 -67.1 -120.4	-0.2 -69.4 -138	5.1 -73.4 -147.8
B (yz plane)	0.2	5.2	6.6

TABLE 5.12. Variation of the significant parameters for a bent cuboid with the addition of noise.

used on a real object. The object of the experiment is to infer a description of the scene by collecting data around the object. Stability of the model parameters and model uniqueness are addressed by observing the parameters over exploration paths starting from a different vantage viewpoint each time.

Figure 5.15 shows the superellipsoid parameters of a pear under a exploration sequence. The position, extent and shape parameters over the various trials converge to acceptable limits. The orientation parameters seem to be out of phase for some of the trials. This is due to the degenerate symmetries of the rotational angles. Rotational symmetries (0 and 180 degrees) were taken care off by canonicalization of the model parameters. This was achieved by making the z axis of the object the longest axis. The deformable model parameters of the pear under the same trials of the exploration are shown in Figure 5.16 and the parameters all coverge to acceptable limits (within 10% parameter variation). The parameters are presented in the Table 5.13 & 5.14. Analysis of the parameters show that the deformable parameters do not contribute much to the shape of the pear, thereby the

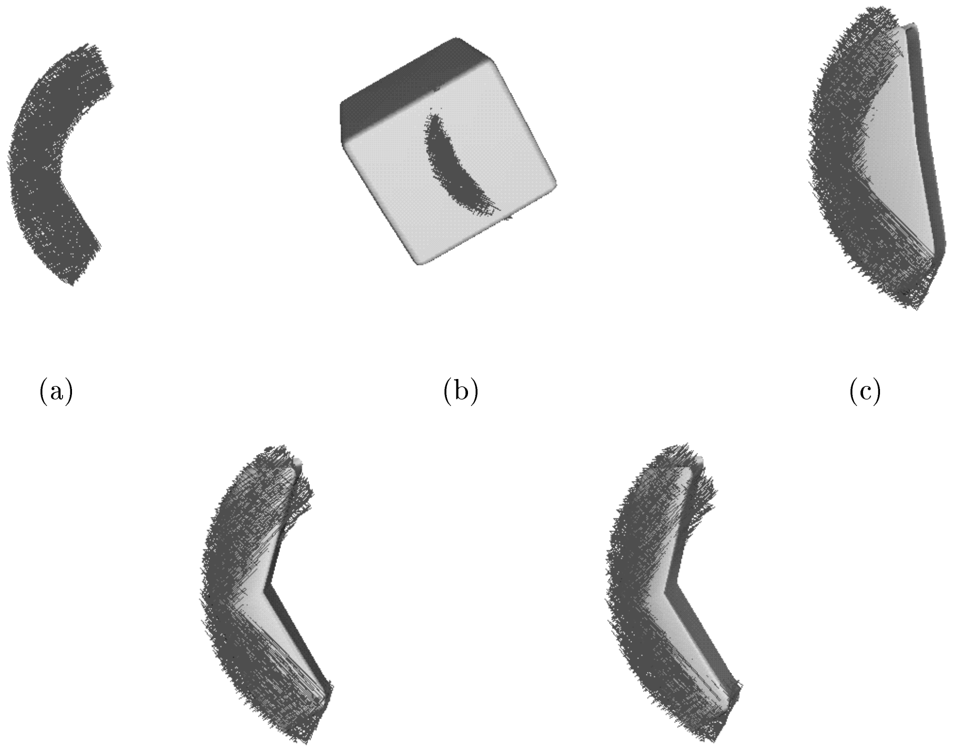


(a) (b) (c) (d) (e)

FIGURE 5.13. 3D range data of a tapered cuboid, the initial fitted model and the various stages of the model to these data. (a) shows the range data of the tapered cuboid. (b) shows the initial fitted model to the data using ϵ_1 and $\epsilon_2 = 0.1$. (c) shows the model after 1 iterations. (d) shows the model after 5 iterations. (e) shows the model after 10 iterations.

shape is determined by the basic superellipsoid parameters (translation, orientation, size and shape exponents). These parameters are shown in Tables 5.13 & 5.14 over the 5 trials from a different initial viewpoint each time. The parameters are within the acceptable limits for Trials 1,3 & 4. For the other 2 trials, the pose and shape parameters seem to be incorrect. The models are shown in Figures 5.17 and 5.18. It can be observed that the deformable models capture the shape of the pear and are a definite improvement over the superellipsoid models.

Figure 5.19 and 5.20 show the variation of the superellipsodial and the deformable parameters of a tapered cuboid under our exploration scheme. Five different trials have been performed on the object with each consecutive trial starting from a different intial position. The parameters have been plotted against the exploration steps for the 5 trials.



(a) (b) (c)

(d) (e)

FIGURE 5.14. 3D range data of a bent cuboid, the initial fitted model and the various stages of the model to these data. (a) shows the range data of the bent object. (b) shows the initial fitted model to the data using ϵ_1 and $\epsilon_2 = 0.1$. (c) shows the model after 1 iterations. (d) shows the model after 5 iterations. (e) shows the model after 10 iterations.

Exploration has been performed in such a way that a coverage of data-points from around the object is obtained. The deformable & superellipsoidal parameters are shown in Tables 5.15 & 5.16. The tapering parameters are considered, since we are dealing with a tapered object in the scene. The tapering occurs along the x axis (along the zx & yx planes) and along the z axis (along the yz plane). These parameters are necessary to deform the inital superellipsoid model to the final tapered model. A saw-tooth waveform is observed in the tapering & bending parameters in the Figure 5.20. On each trial, one of the parameter reaches its maximum value. The reason for this is that for one of the tapering parameters to reach a certain value, the other parameter has to be compensated to account for the tapering so as to not deform the object out of shape.

Variation of parameters over trials (1,2,3) of exploration

	1st Trial	2nd trial	3rd trial
TX, TY, TZ	-14.3 -1.47 -4.5	-14.2 -2.09 3.7	-14.25 -3.8 0.3
RX, RY, RZ	29.3 -3.47 1.8	13.4 6.7 -2.08	-34.25 -2.3 -178.8
AX, AY, AZ	77 29.4 29.9	72.9 31.9 31.9	75.7 32 29.7
SX, SY, SZ	0.34 0.21 0.25	-0.14 -0.31 -0.09	-0.2 -0.02 0.15
TBP (xy plane)	-0.03 -0.27 0.4	-1.8 0.42 -0.78	-0.5 0.12 0.04
TBP (xz plane)	0.65 -0.37 -0.13	-1.34 -0.35 -0.23	-0.33 0.71 -0.37
TBP (yx plane)	-1.87 -0.84 0.06	-0.7 0.06 -0.2	0.2 0.8 0.3
TBP (yz plane)	0.86 -0.65 0.01	0.55 0.07 0.1	0.26 0.51 -0.04
TBP (zx plane)	-0.95 0.21 -0.27	-1.58 0.21 -0.2	0.65 -0.4 -0.34
TBP (zy plane)	0.11 0.01 0.6	-1.28 0.22 -0.7	0.24 -0.35 0.39
E1, E2	0.9 1.66	1.16 1.53	0.9 1.69

TABLE 5.13. Superellipsoid and deformable parameter variation of a pear over 3 trials (1,2,3) of the exploration process.

The superellipsodial parameters (Figure 5.19) fit an approximate model over the exploration steps. The reason for the overfitted initial model could be attributed to the fact that the model (tapered cuboid) cannot be fitted reasonably using a superellipsoid. But this model provides us with a good initial guess. The variation of parameters are within acceptable limits ($\pm$ 10% parameter variation). There is a considerable improvement over the superellipsoid model as shown in Figure (5.21 and 5.22).

Similar experiments are performed on a bent object. Figure 5.23 and 5.24 illustrate the parameters over the exploration sequence. The superellipsoids, again provide us with a good initial model. Bending is achieved along the x axis (along the yx plane). The deformable parameters are shown in the table5.17. The parameters remain within acceptable limits ($\pm$ 10% parameter variation) over the different trials. Some of the tapering and pinching parameters also aid in deforming the initial superellipsoid model to the final deformable model. Again, considerable improvement over the initial model is achieved as seen in Figures (5.25 and 5.26).

Variation of parameters over trials (4,5) of exploration

	4th trial	5th trial
TX, TY, TZ	-24.2 -0.24 1.4	-27.4 -1.13 -0.6
RX, RY, RZ	69.9 -2.2 0.52	103.19 12.36 177.15
AX, AY, AZ	83.3 35.04 34.2	84.9 33.18 33.5
SX, SY, SZ	0.57 -0.02 0.15	0.1 0.12 -0.04
TBP (xy plane)	0.15 0.8 0.5	-0.02 0.04 0.07
TBP (xz plane)	-0.09 -0.07 0.26	-0.28 0.18 -0.02
TBP (yx plane)	0.27 -0.34 0.32	0.35 -0.01 0.06
TBP (yz plane)	1 -0.53 0.33	0.78 -0.26 0.22
TBP (zx plane)	-0.77 0.7 -0.54	0.2 -0.58 -0.3
TBP (zy plane)	-0.68 0.86 -0.01	-0.18 -0.63 -0.28
E1, E2	1.14 1.58	1.24 1.42

Table 5.14. Superellipsoid and deformable parameter variation of a pear over the 4th and 5th trials of the exploration process.

Variation of parameters over 3 trials (1,2,3) of exploration

	1st trial	2nd trial	3rd trial
TX, TY, TZ	-47.8 25.8 323.6	-41.07 25.06 316.5	-47.68 25.78 323.04
RX, RY, RZ	-2.05 -54.6 -165.9	0.12 -61.14 -165.64	-3.6 -54.3 -168.75
AX, AY, AZ	45.84 20.56 20.26	46.35 22.67 22.06	45.69 20.4 19.8
SX, SY, SZ	0.06 -0.03 -0.07	0.65 0.97 0.93	-0.07 0.06 0.21
TBP (xy plane)	0.32 -5.5 1.15	2.57 -14.43 9.58	1.07 -5.7 0
TBP (xz plane)	-0.15 -6.02 0.5	-2.96 2.65 0.33	0.27 -4.94 0.9
TBP (yx plane)	-3.85 0.02 0.93	2.04 0.72 1.96	-4.4 0.03 0.78
TBP (yz plane)	0.3 0.3 1.01	-8.56 -0.4 3.4	0.02 -0.09 1.36
TBP (zx plane)	-4.74 -0.7 1.18	-0.36 7.3 -1.15	-4.9 0.48 1.2
TBP (zy plane)	-0.17 -0.11 2.12	-1.5 4.7 1.11	-0.27 0.5 2
E1, E2	0.45 0.7	0.1 0.6	0.43 0.64

Table 5.15. Variation of the superellipsoid and deformable parameters of a tapered object over 3 (1,2,3) trials of the exploration process.

Variation of parameters over the 4th and 5th trials of exploration

	4th trial	5th trial
TX, TY, TZ	-47.8 25.8 323.7	-47.56 26 323.8
RX, RY, RZ	-1.95 -54.5 -165.8	-1.9 -54.5 -165.8
AX, AY, AZ	45.76 20.6 20.3	45.6 20.58 20.3
SX, SY, SZ	0.06 -0.03 -0.04	0.09 -0.03 0.01
TBP (xy plane)	0.33 -5.44 1.15	0.5 -5.2 1.14
TBP (xz plane)	-0.08 -6 0.54	0.08 -5.46 0.5
TBP (yx plane)	-3.8 -0.01 0.88	-3.8 -0.01 0.8
TBP (yz plane)	0.33 0.27 0.97	0.27 0.44 1.26
TBP (zx plane)	-4.84 -0.66 1.4	-4.8 -0.6 1.6
TBP (zy plane)	-0.13 -0.15 2.18	-0.15 -0.06 2.14
E1, E2	0.45 0.7	0.45 0.72

TABLE 5.16. Variation of the superellipsoid and deformable parameters of a tapered object over 4th and 5th trials of the exploration process.

Variation of parameters over 3 (1,2,3) trials of exploration

	1st trial	2nd trial	3rd trial
TX, TY, TZ	-39.98 26.8 312.3	-39.03 27.18 313.24	-40.18 27.16 312.21
RX, RY, RZ	79.44 11.76 54.08	79.12 11.71 54.73	79.43 11.5 54.5
AX, AY, AZ	51.84 19.72 9.94	52.5 19.07 10.15	51.9 19.82 9.99
SX, SY, SZ	0.23 0.06 -0.09	0.09 0.27 -0.26	0.1 0.1 0.09
TBP (xy plane)	14.45 1.16 17.44	13.25 -0.33 16.06	14.74 1.35 17.64
TBP (xz plane)	-0.34 0.87 -0.5	-0.21 -0.02 -1.26	-0.11 1.17 -0.48
TBP (yx plane)	-1.56 13.98 7.09	-1.56 14.57 7.65	-1.28 13.69 6.97
TBP (yz plane)	0.04 1.07 -1.32	0.27 -0.03 -1.3	-0.14 1.14 -1.25
TBP (zx plane)	-0.68 0.07 0.9	-0.85 0.05 1.06	-0.84 0.65 1.45
TBP (zy plane)	-0.36 0.51 -0.56	-0.12 0.24 0.07	-0.33 0.9 0
E1, E2	0.1 1.38	0.1 1.39	0.1 1.38

TABLE 5.17. Table showing the variation of the superellipsoid and deformable parameters of a bent object over 3 (1,2,3) trials of the exploration process

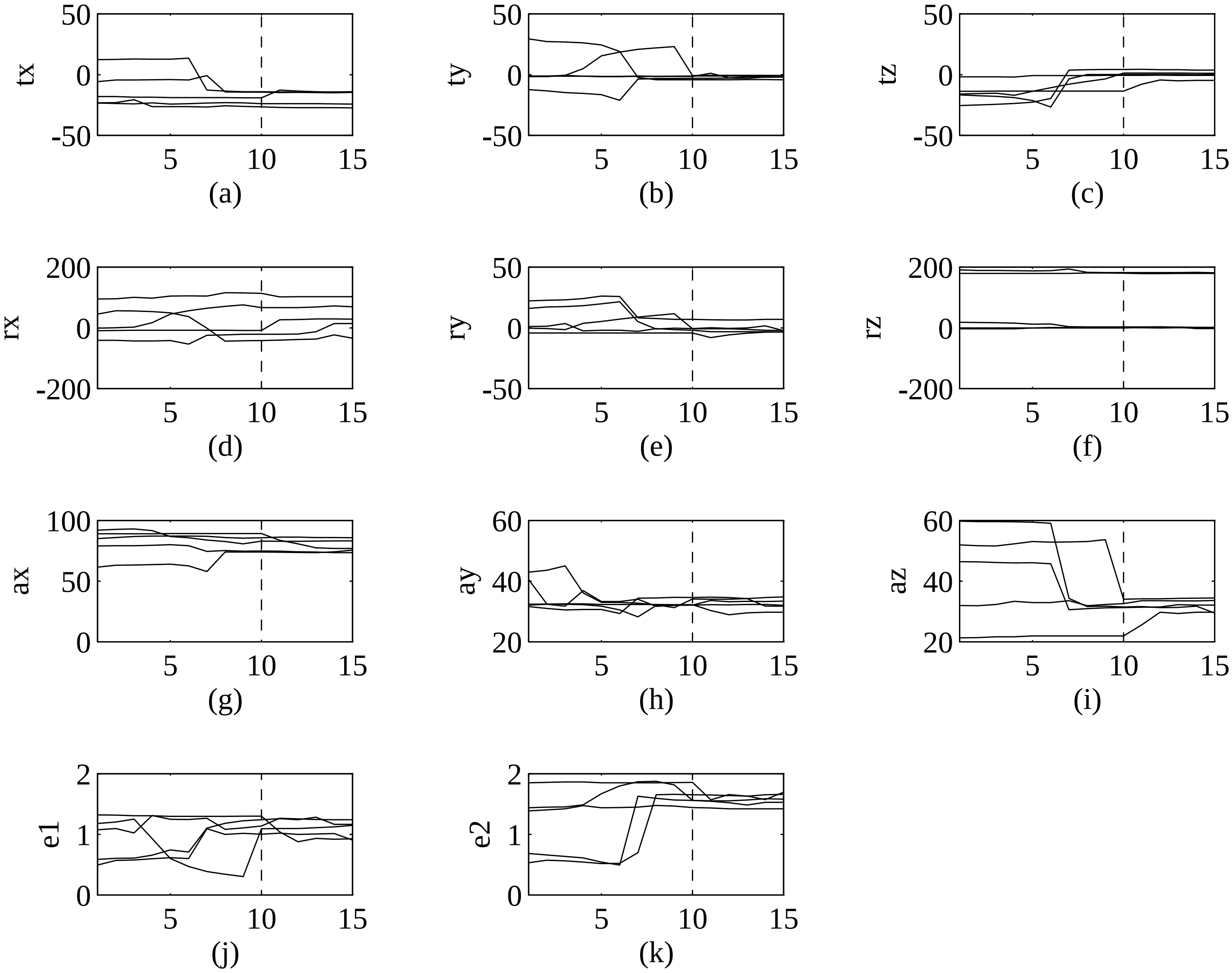


FIGURE 5.15. Superellipsoid parameters of a pear under a exploration sequence. (a), (b) and (c) show the variation of the translation parameters (tx, ty, tz) over the 15 exploration steps and 5 trials (d), (e) and (f) show the variation of the rotation parameters (rx, ry, rz). (g), (h) and (i) show the variation of the extent parameters (ax, ay, az). (j) and (k) show the variation of the shape exponents (e1 and e2).

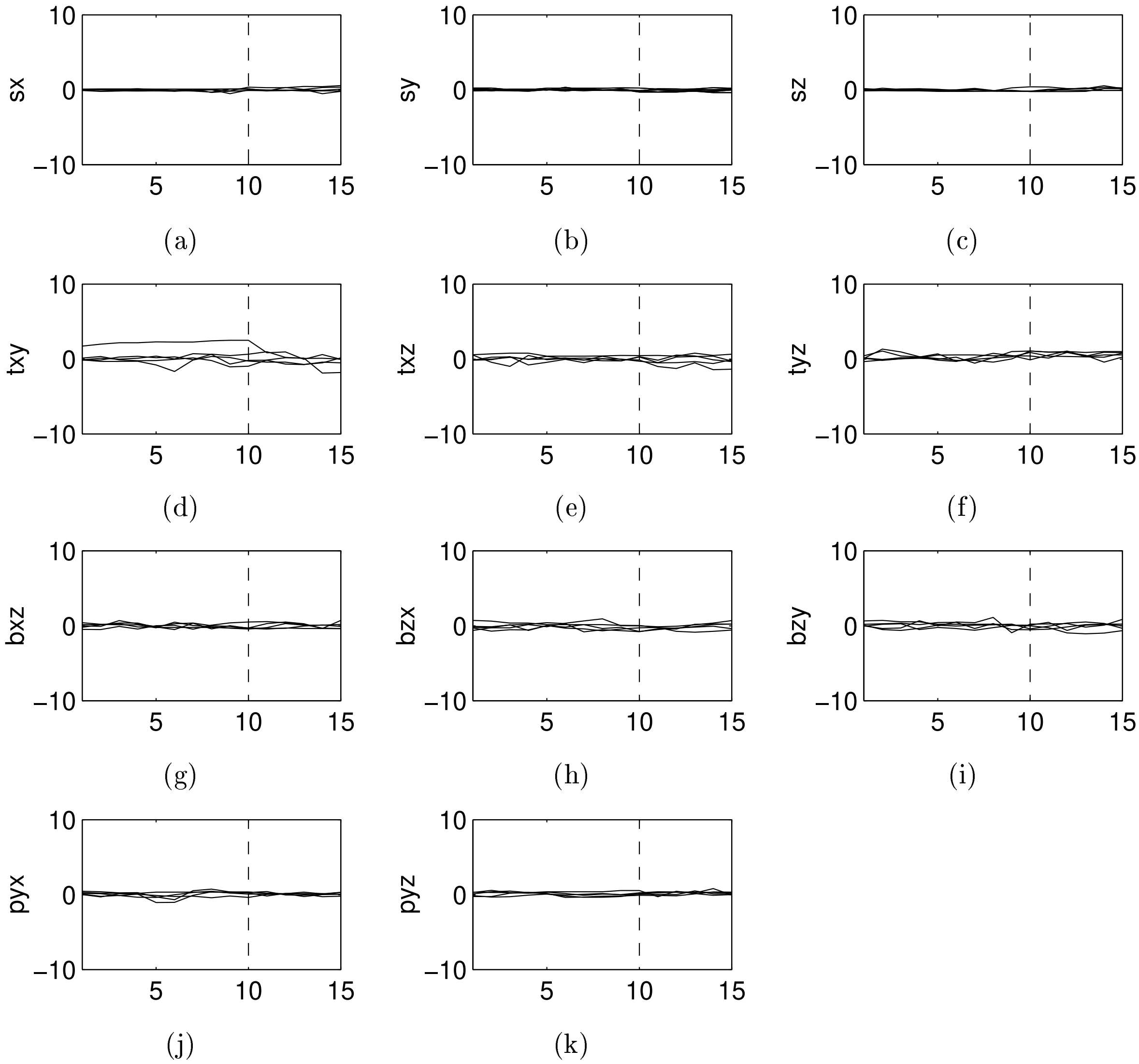


FIGURE 5.16. Deformable parameters of a pear under a exploration sequence. (a), (b) and (c) show the variation of the shearing parameters (sx, sy, sz) over the 15 exploration steps and 5 trials. (d), (e), (f) show the variation of the tapering parameters (txy, txz, tyz). (g), (h) and (i) show the variation of the bending parameters (bxz, bzx, bzy). (j) and (k) show the variation of the pinching parameters (pyx, pyz).

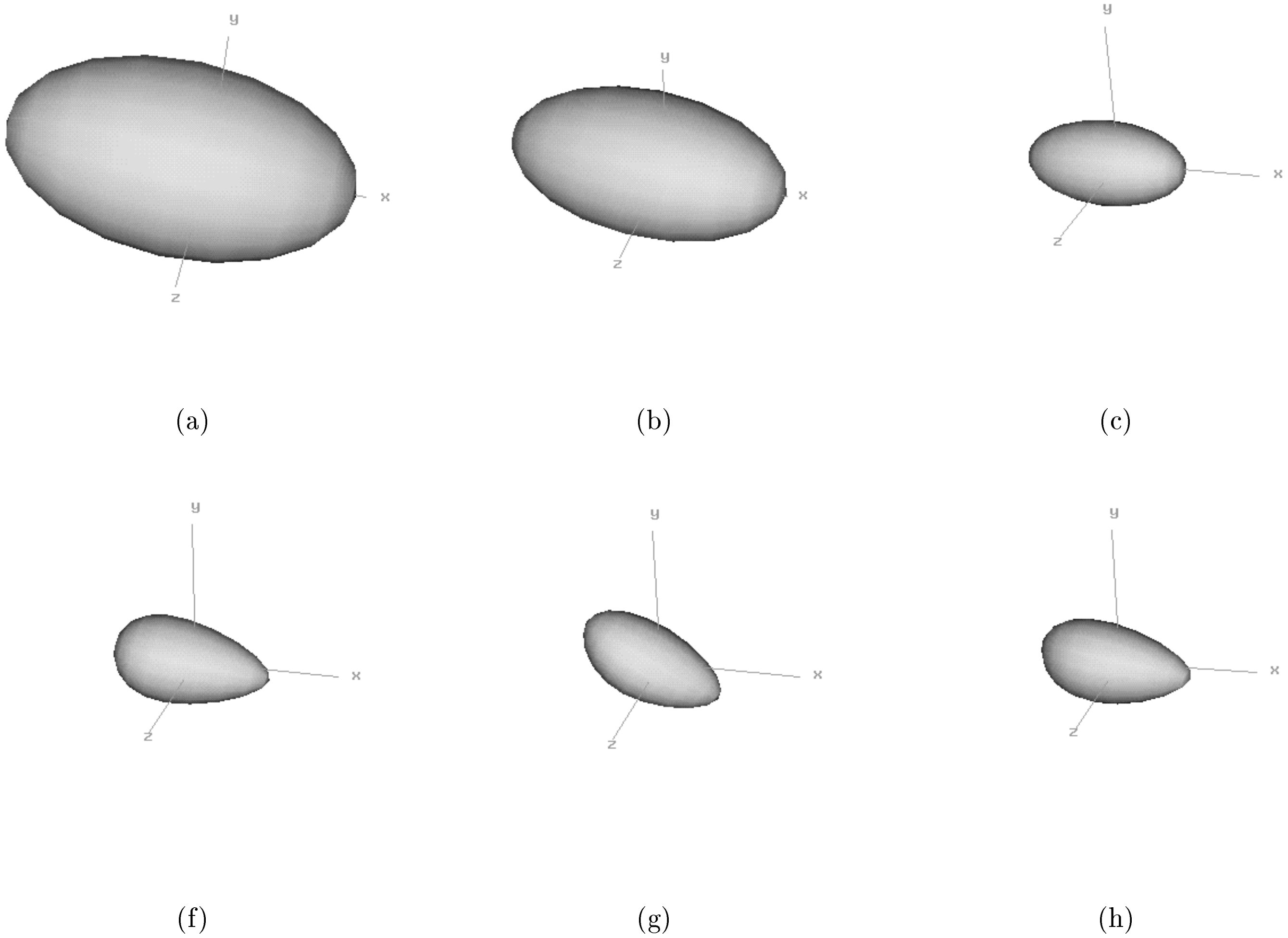
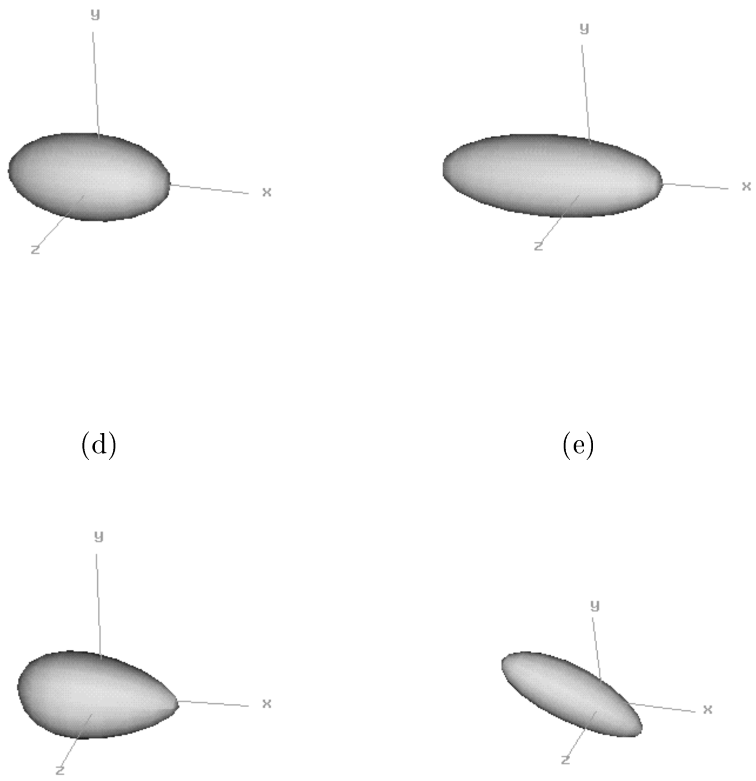


FIGURE 5.17. Superellipsoid and deformable models of a pear after 20 exploration and 5 trials (1st, 2nd and 3rd trials are shown.) (a) and (f) shows the superellipsoid and deformable respectively from an initial view position of (-90, 0) - 1st trial . (b) and (g) shows the superellipsoid and deformable respectively from an initial view position of (90, 0) - 2nd trial . (c) and (h) shows the superellipsoid and deformable respectively from an initial view position of (0, -90) - 3rd trial .



(d) (e)

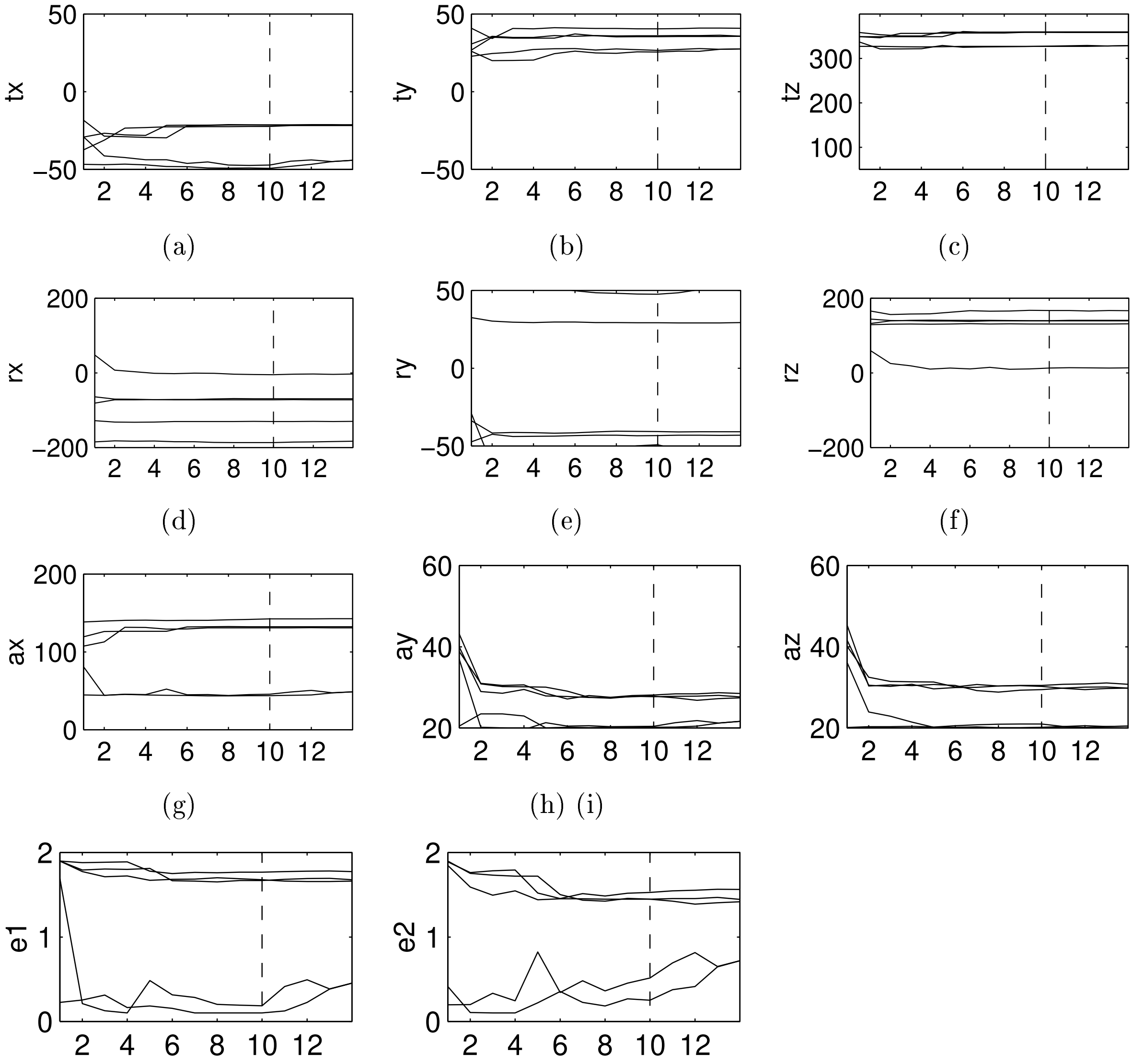
(i) (j)

FIGURE 5.18. Superellipsoid and deformable models of a pear after 20 exploration and 5 trials (the 4th and 5th trials are shown.) (d) and (i) shows the superellipsoid and deformable respectively from an initial view position of (80, 0) - 4th trial . (e) and (j) shows the superellipsoid and deformable respectively from an initial view position of (0, -75) - 5th trial .

Variation of parameters over 4th & 5th trials of exploration

	4th trial	5th trial
TX, TY, TZ	-40.06 27.2 312.2	-39.98 26.85 312.3
RX, RY, RZ	79.44 11.72 54.07	79.44 11.76 54.08
AX, AY, AZ	51.83 19.72 10	51.83 19.72 9.94
SX, SY, SZ	0.23 0.05 -0.09	0.23 0.06 -0.09
TBP (xy plane)	14.45 1.17 17.44	14.45 1.16 17.44
TBP (xz plane)	-0.34 0.87 -0.5	-0.34 0.87 -0.49
TBP (yx plane)	-1.56 13.96 7.1	-1.56 13.98 7.1
TBP (yz plane)	0.03 1.09 -1.32	0.04 1.07 -1.32
TBP (zx plane)	-0.68 0.53 1.29	-0.68 0.07 0.89
TBP (zy plane)	-0.36 0.97 -0.15	-0.36 0.51 -0.56
E1, E2	0.1 1.38	0.1 1.38

TABLE 5.18. Table showing the variation of the superellipsoid and deformable parameters of a bent object over 4th & 5th trials of the exploration process



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

FIGURE 5.19. Superellipsoid parameters of a tapered cuboid under a exploration sequence. (a), (b) and (c) show the variation of the translation parameters (tx, ty, tz) over the 15 exploration steps and 5 trials (d), (e) and (f) show the variation of the rotation parameters (rx, ry, rz). (g), (h) and (i) show the variation of the extent parameters (ax, ay, az). (j) and (k) show the variation of the shape exponents ($e1$ and $e2$).

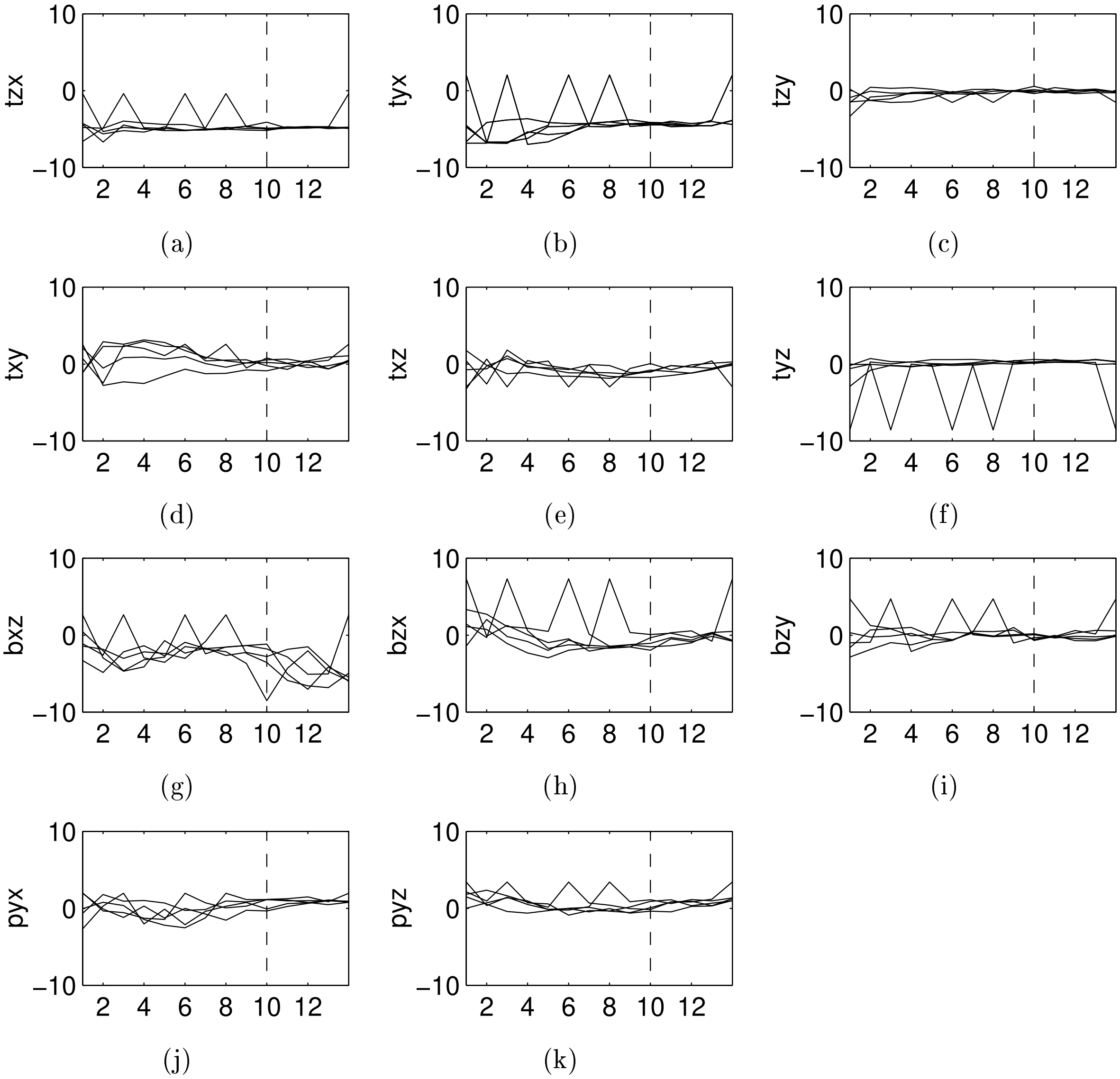


FIGURE 5.20. Deformable parameters of a tapered cuboid under a exploration sequence. (a), (b), (c), (d), (e) and (f) show the variation of the tapering parameters (tzx, tyx, tzy, txy, txz, tyz) over the 15 exploration steps and 5 trials. (g), (h) and (i) show the variation of the bending parameters (bxz, bzx, bzy). (j) and (k) show the variation of the pinching parameters (pyx, pyz).

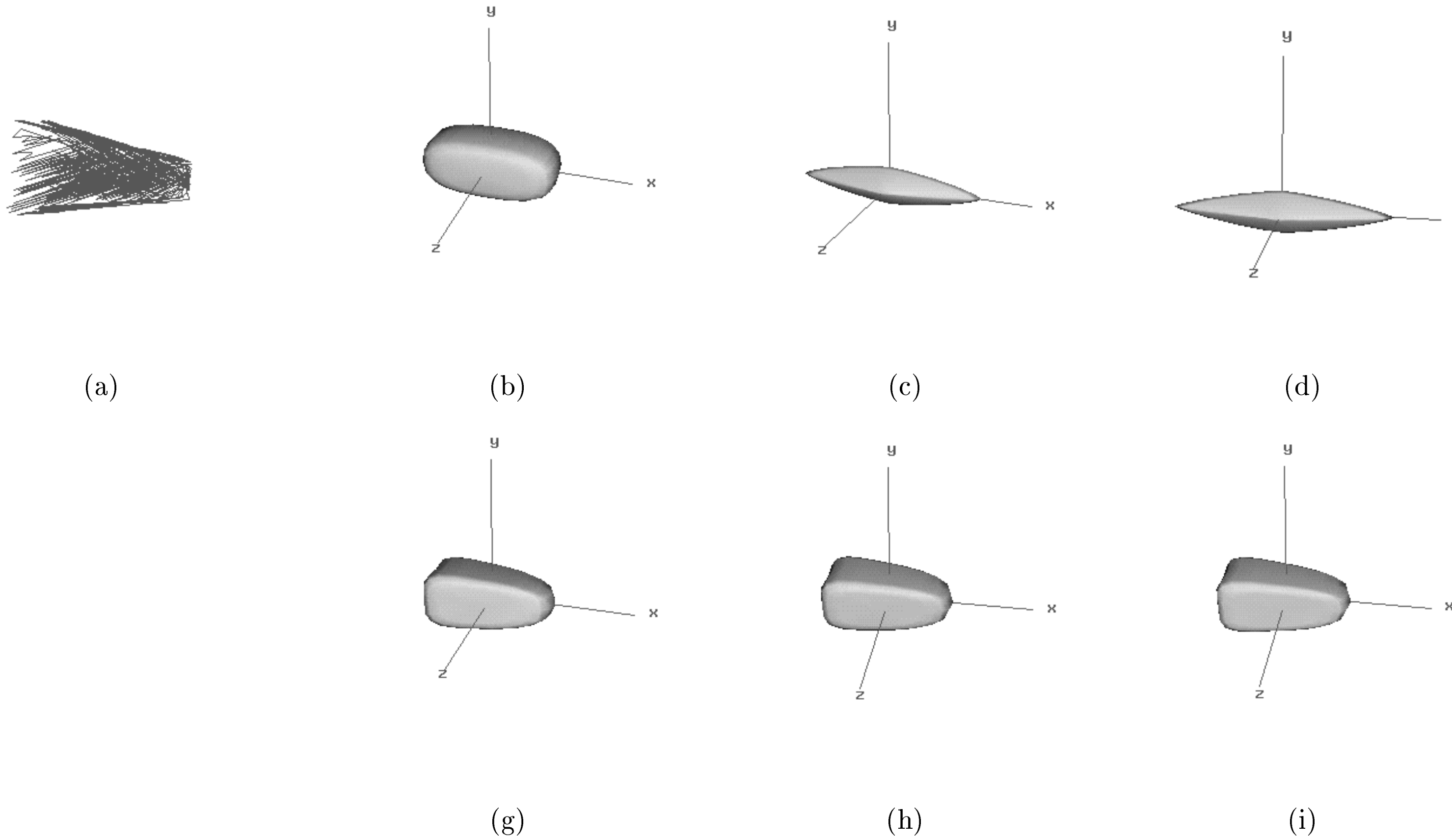
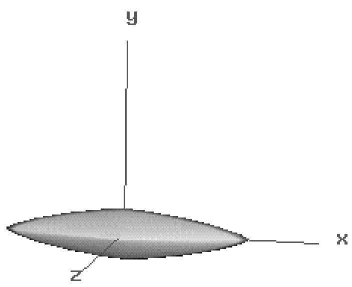
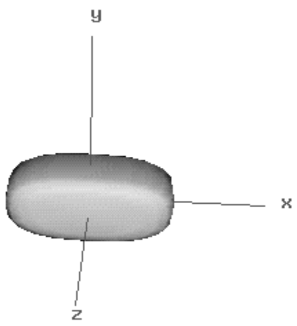


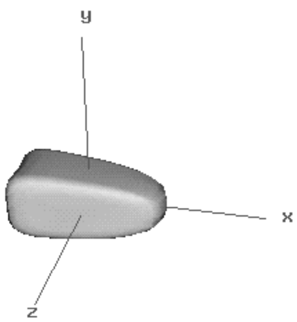
FIGURE 5.21. Superellipsoid and deformable models of a tapered cuboid after 15 exploration paths and 5 trials (1st, 2nd and 3rd trials are shown.) (a) shows the range data of the original model (tapered cuboid). (b) and (g) shows the superellipsoid and deformable respectively from an initial view position of (0, 45) - 1st trial . (c) and (h) shows the superellipsoid and deformable respectively from an initial view position of (-45, 90) - 2nd trial . (d) and (i) shows the superellipsoid and deformable respectively from an initial view position of (45, 180) - 3rd trial .



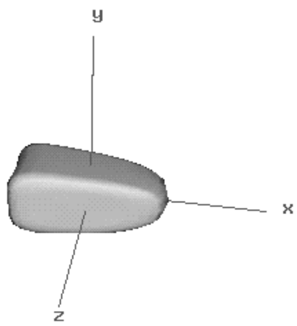
(e)



(f)



(j)



(k)

FIGURE 5.22. Superellipsoid and deformable models of a tapered cuboid after 15 exploration paths and 5 trials (4th and 5th trials are shown.) (e) and (j) shows the superellipsoid and deformable respectively from an initial view position of (45, -90) - 4th trial . (f) and (k) shows the superellipsoid and deformable respectively from an initial view position of (0,135) - 5th trial .

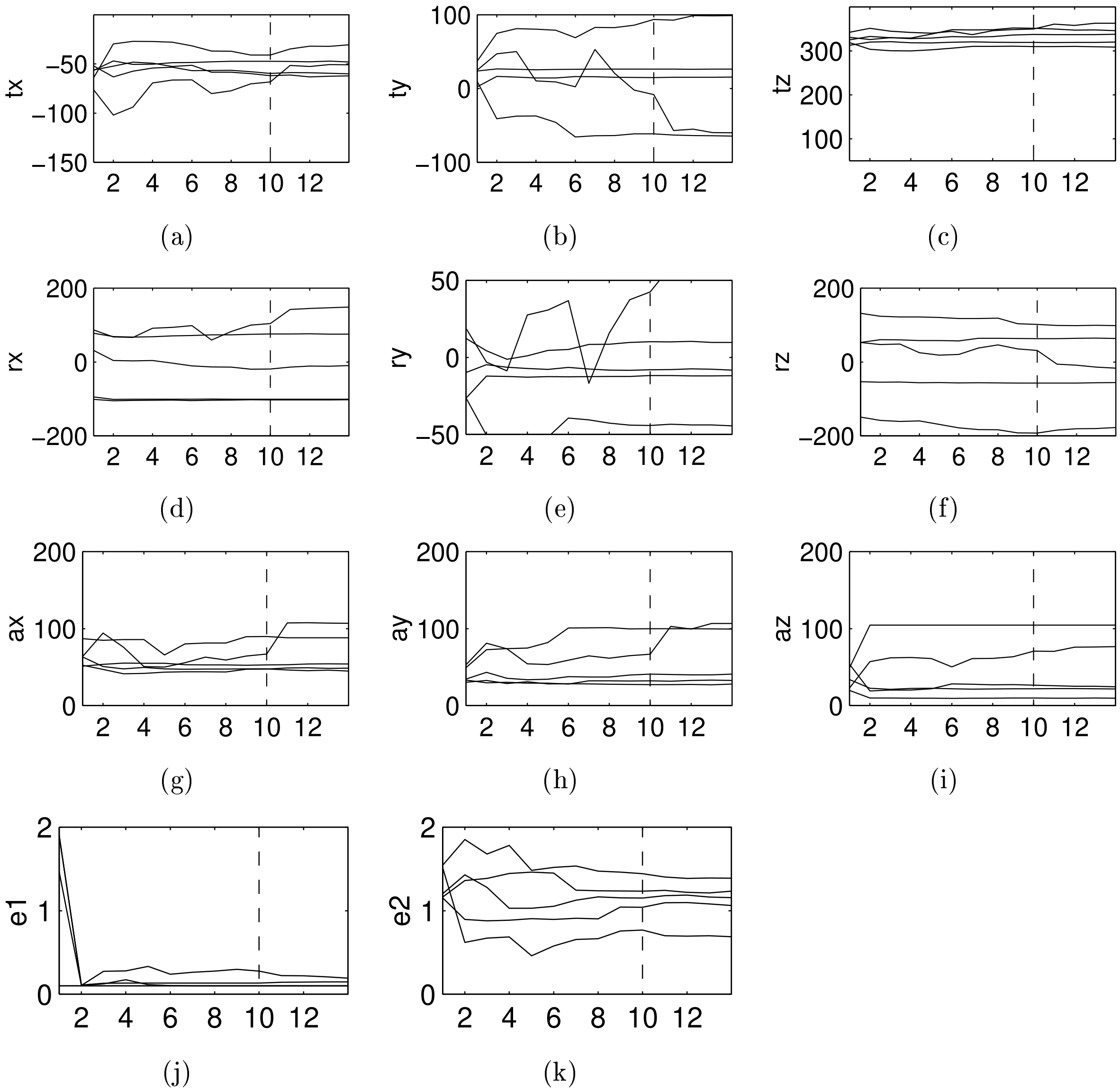


FIGURE 5.23. Superellipsoid parameters of a bent object under a exploration sequence. (a), (b) and (c) show the variation of the translation parameters (tx, ty, tz) over the 15 exploration steps and 5 trials (d), (e) and (f) show the variation of the rotation parameters (rx, ry, rz). (g), (h) and (i) show the variation of the extent parameters (ax, ay, az). (j) and (k) show the variation of the shape exponents ($e1$ and $e2$).

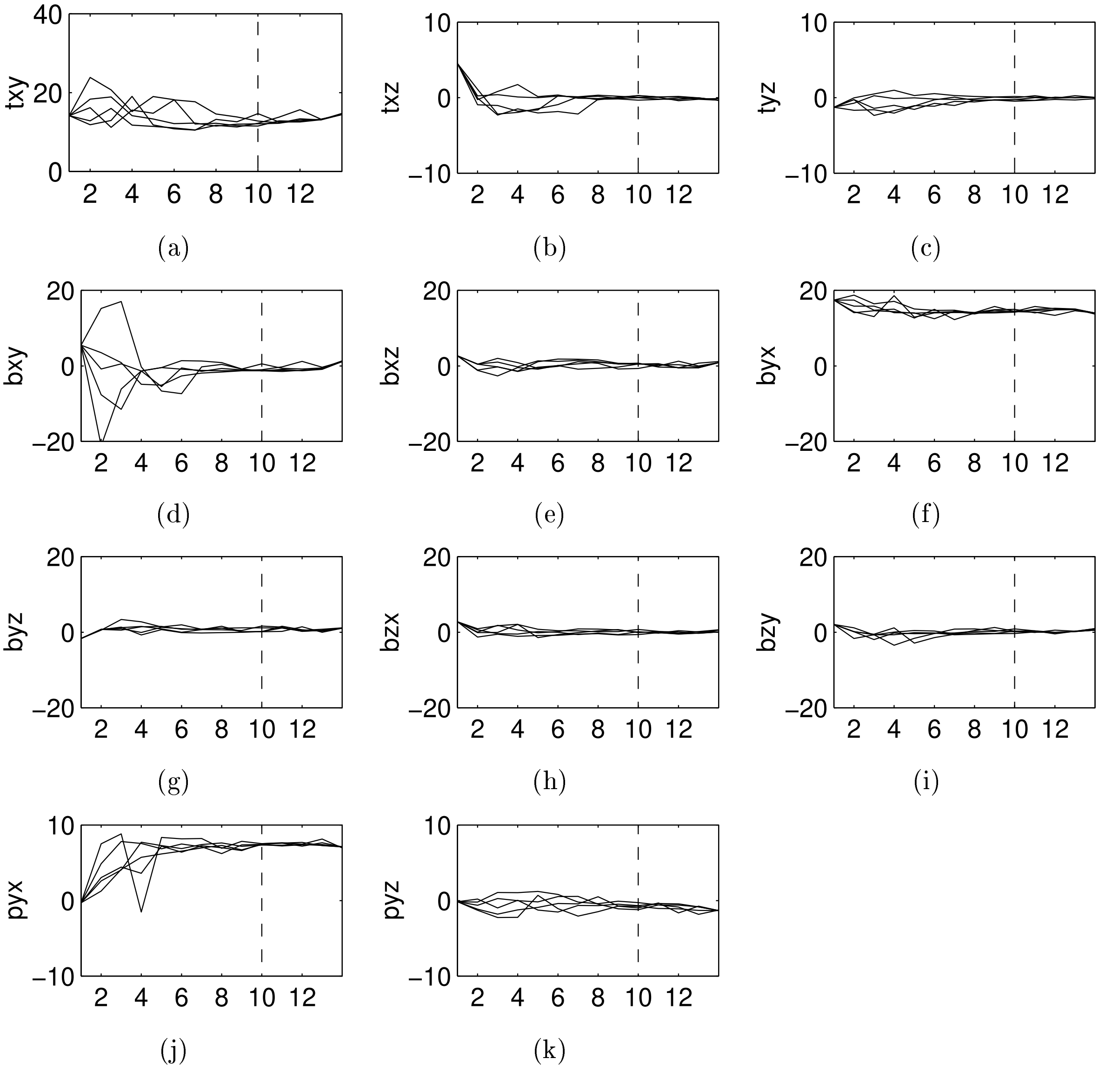


FIGURE 5.24. Deformable parameters of a bent object under a exploration sequence. (a), (b) and (c) show the variation of the tapering parameters (txy, txz, tyz) over the 15 exploration steps and 5 trials. (d), (e), (f), (g), (h) and (i) show the variation of the bending parameters (bxy, bxz, byx, byz, bzx, bzy). (j) and (k) show the variation of the pinching parameters (pyx, pyz).

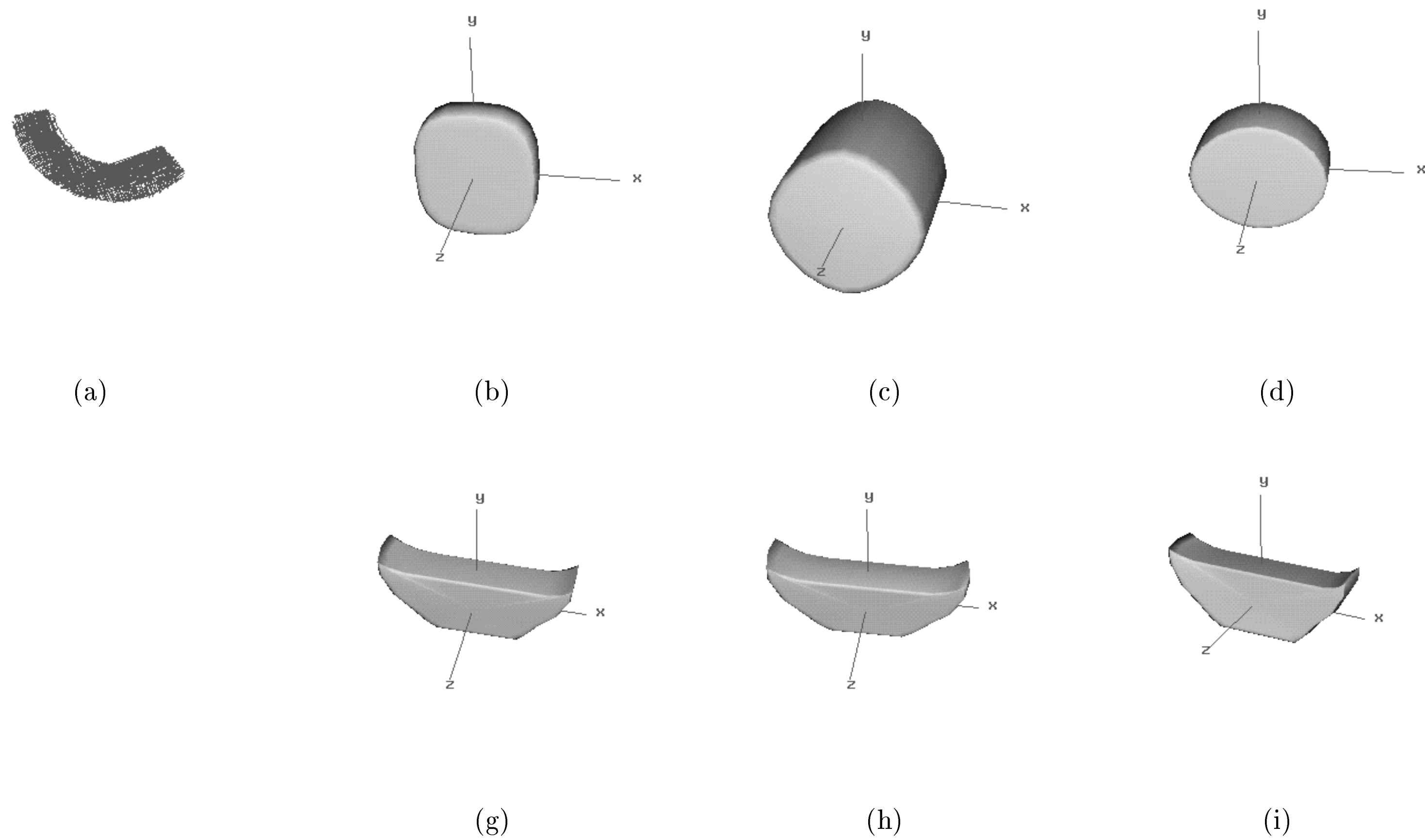


FIGURE 5.25. Superellipsoid and deformable models of a bent cuboid after 15 exploration paths and 5 trials (1st, 2nd and 3rd trials are shown) (a) shows the range data of the original model (bent cuboid). (b) and (g) shows the superellipsoid and deformable respectively from an initial view position of (0, 45) - 1st trial . (c) and (h) shows the superellipsoid and deformable respectively from an initial view position of (-45, 90) - 2nd trial . (d) and (i) shows the superellipsoid and deformable respectively from an initial view position of (45, 180) - 3rd trial .

Results are shown with varying levels of range data (partial, complete, etc..) to infer volumetric models. Issues of stability of model parameters, normalization of parameters, model uniqueness are all touched upon. Incremental recovery of the volumetric models is achieved by using an active strategy, namely, autonomous exploration to infer a much more complete and richer description of the scene.

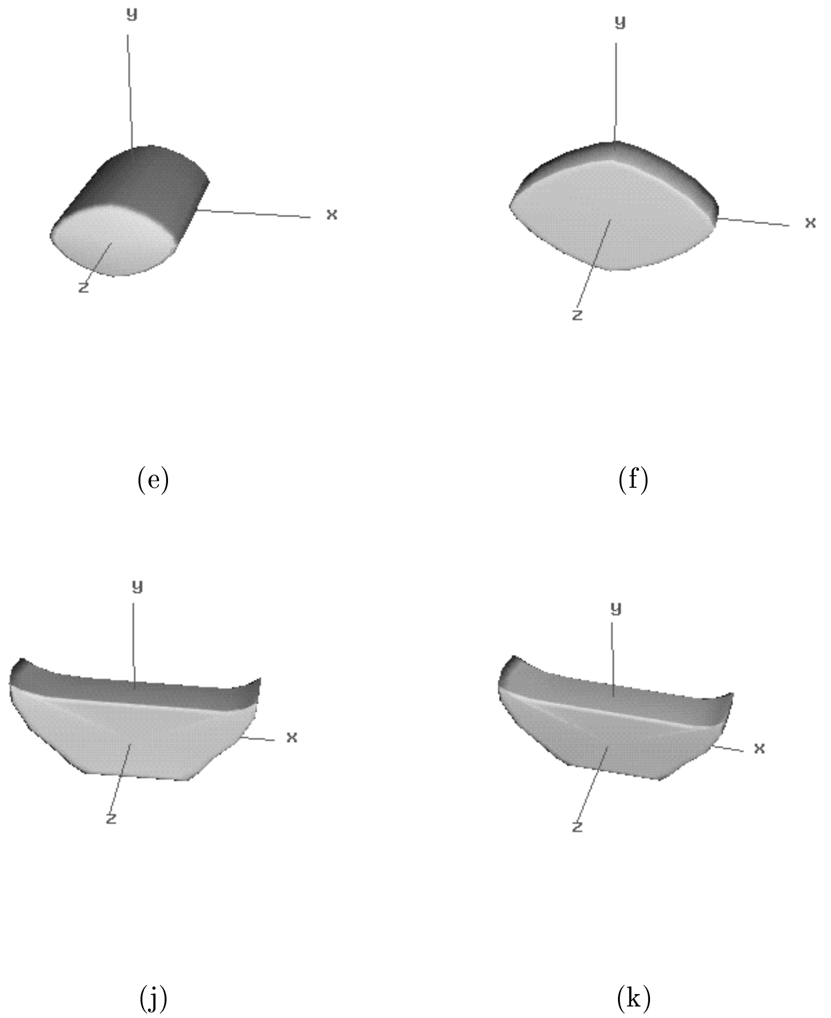


FIGURE 5.26. Superellipsoid and deformable models of a bent cuboid after 15 exploration paths and 5 trials (4th and 5th trials are shown) (e) and (j) shows the superellipsoid and deformable respectively from an initial view position of (45, -90) - 4th trial . (f) and (k) shows the superellipsoid and deformable respectively from an initial view position of (0,135) - 5th trial .

CHAPTER 6

Conclusions and Future Work

In this thesis, we have investigated the use of a deformable model representation, first developed by Pentland [**24**] with respect to an active exploration process. An exploration strategy [**36**] aids in this dynamic model building process. Such a strategy has some inherent problems, such as the stability of the model parameters, viewpoint independence, initialization of the model with the addition of data. An exhaustive study has been performed on a number of different objects to show the model uniqueness and the stability of parameters. During the experiments, exporation has been performed on objects starting from a new viewpoint at each stage of the trial. Parameters are observed over different exploation paths and trials. Observations have been made about the stability and uniqueness of the model parameters. The results indicate that the model is unique with respect to sampling and viewpoint.

We have also described a methodology for the recovery of the 3D shape of objects from surface measurements (e.g. using a laser rangefinding system). The recovered models are deformable models which provide us with a more complete description of the scene. These models evolve from superquadrics which is used as an initial estimate. The notion of superquadrics is enlarged to include non-rigid transformations such as shearing, tapering, bending and pinching to a general class of deformable superquadrics. The method for recovery of such volumetric part models is based on weighted least squares minimization of an implicit fitting function for all input range points belonging to a single-part object. The initial estimate of the model is established with a gradient descent technique and to start the minimization, only estimates for rough position, orientation and size of the object (part) have to be computed. Model's shape parameters are always initialized as being a sphere or ellipsoid and the final shape exponents are recovered through the gradient descent process. The deformation parameters are recovered using weighted least squares in a modified Newton-Raphson procedure. These deformation parameters (modes) in addition to the 11 parameters (3 estimates of position, orientation and size and 2 shape exponents) will describe our deformable superquadric model.

Correspondence of the data-points to the model while fitting is an important concern. Several correspondence routines have been developed to account for the different datasets, such as, partial, complete 3D data, etc. A lot of care has to be taken so that the data points which do not have a major influence on the model are given very little weights, otherwise, the model might deform out of shape during fitting. For our purpose, finding the closest data point along the model surface seemed to suffice. Other sophisticated correspondence routines could be developed to take care of the problem.

During experimentation, lots of difficulties were faced due to the problem of non-uniform range data density and a large number of singular views in range images. For example, if only one face of a cube is seen, a very thin parallelepiped is recovered which fits to that face. Images taken with a passive range-finding system, which uses triangulation (like our's 3.1), have more singular views than normally associated with intensity images. The larger the distance (angle) between the source of illumination and the camera of the range finder, the better is the accuracy, but more singular views, since every recorded range point must be illuminated and seen by the camera at the same time. Singular views can be resolved by taking into account the structure of the surrounding scene. Objects normally rest on some support, they can touch but usually do not penetrate each other. This type of constraint could be incorporated in the part recovery procedure by a method devised by Witkin *et al* [**37**]. They developed an elegant method for describing geometric constraints between parts and objects in terms of energy. Nonuniform range point density causes the parts with higher density to have more influence on the shape of the recovered model than parts with lower density. However, the model recovery method is quite robust in this regard - note that no range data is available from the occluded parts to begin with. During the correspondence process, we used the same weights for all range data points and all model parameters if the data points are uniformly sampled. In case range data is partial, a greater weight is attached to the range points closer to the model so that a convex-hull like model can be recovered.

For simple objects like cubes, cylinders, our experimental results shows that the correspondence process yields accurate and unique object descriptions (section 4.3). The computation of the ϕ matrix effectively linearizes the object rotation, so it is imperative that the elliptical coordinate system established by the moments' method be sufficiently close to the object's true coordinate system. A variation of $\pm$ 15 degrees of the initial estimate of orientation with respect to the actual orientation is acceptable [**24**]. For complex objects, such as scene descriptions or articulated models made of many parts, this correspondence process is a nonlinear process, so it is necessary to establish the attachment positions iteratively. In the fitting procedure, this is accomplished by fitting a rough model based on the moments' method and then use this model to accurately determine the other spring attachment positions. A few iterations are normally sufficient for both the correspondence and the fitting procedure as is shown by our experimental results (Chapter 4).

The experiments performed in the previous chapter are constrained by the accuracy obtained and the limitations of the equipment used for data acquisition like the laser-range finding system and the rotary table. During the autonomous exploration of a scene description, the objects are lying on the table, so it is impossible to get data from under the object thereby constraining the volumetric model to the data obtained from other views. Autonomous exploration was performed on real objects to infer deformable models. The gaze planning strategy (i.e. where to look next) was based on the uncertainity surface of the part closest to the rangefinder. For models which can be characterized by superellipsoids, this strategy works very well. But for objects with non-rigid transformations like tapering, bending, etc.. this strategy runs into problems. Maybe, the gaze planning, in this case, should be based on the uncertainity surface of the deformable model. Use can be made of the covariance matrices of the deformable parameters to give an indication of the uncertainity of the parameters to guide the exploration strategy.

APPENDIX A

1. Generalized Implicit Functions

An implicit function representation [**29**] defines a surface as a level set of a function f, most commonly the set of points for which $f(x) = 0$. In the case of an superquadric ellipsoid, the inside-outside function, before rotation, translation or deformation is an implicit function of the form

$f(x) =$

$$((x^{2/\epsilon_2} + y^{2/\epsilon_2})^{\epsilon_2/\epsilon_1} + z^{2/\epsilon_1})^{\epsilon_1/2} - 1$$

A solid defined in this way can easily be positioned and oriented, by transforming the implicit function

$$\bar{\mathbf{x}} = \mathbf{M}\mathbf{x} + \mathbf{b} \tag{A.1}$$

where $\mathbf{M}$ is a rotation matrix, and $\mathbf{b}$ is a translation vector. Now the implicit function's positioned and oriented inside-outside function becomes

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{M}^{-1}(\bar{\mathbf{x}} - \mathbf{b})) \tag{A.2}$$

The above set of equations can be generalized further by defining an appropiate set of global deformations D with parameters u. For particular values of u the new deformed surface is defined using a deformation matrix $\mathbf{D_u}$

$$\bar{\mathbf{x}} = \mathbf{M}\mathbf{D_u}\mathbf{x} + \mathbf{b} \tag{A.3}$$

The inside-outside function now becomes

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{D_u^{-1}}\mathbf{M}^{-1}(\bar{\mathbf{x}} - \mathbf{b})) \tag{A.4}$$

This implicit function is valid as long as the inverse deformation $\mathbf{D_u^{-1}}$ exists. $\mathbf{D_u}$ is known as the modal deformation matrix which acts like a transformation of the original undeformed data points to the final observed deformation positions. The entries of the matrix are polynomials which mimic the free vibration modes found in the real objects. A linear superposition of these deformation polynomials allows an accurate description of the dynamic behavior of real objects.

The commonly used polynomials as entries of $\mathbf{D_u}$ are

$$
\begin{aligned}
\mathbf{d_{00}} &= \mathbf{u_6 + yu_{12} + zu_{15} - (u_{13} + u_{16})sgn(x) - u_{14} - u_{17}} \\
\mathbf{d_{01}} &= \mathbf{u_{11} + 2y(u_{13} + sgn(x)u_{14})} \\
\mathbf{d_{02}} &= \mathbf{u_{10} + 2z(u_{16} + sgn(x)u_{17})} \\
\mathbf{d_{10}} &= \mathbf{u_{11} + 2y(u_{19} + sgn(x)u_{20})} \\
\mathbf{d_{11}} &= \mathbf{u_7 + xu_{18} + zu_{21} - (u_{19} + u_{22})sgn(y) - u_{20} - u_{23}} \\
\mathbf{d_{12}} &= \mathbf{u_{11} + 2z(u_{22} + sgn(y)u_{23})} \\
\mathbf{d_{20}} &= \mathbf{u_{10} + 2x(u_{25} + sgn(z)u_{26})} \\
\mathbf{d_{21}} &= \mathbf{u_9 + 2y(u_{28} + sgn(x)u_{29})} \\
\mathbf{d_{22}} &= \mathbf{u_8 + xu_{24} + yu_{27} - (u_{25} + u_{28})sgn(z) - u_{26} - u_{29}}
\end{aligned}
\tag{A.5}
$$

Here $\mathbf{u} = [\mathbf{u_0}, \mathbf{u_1}, ..., \mathbf{u_{p-1}}]^T$ is a $\mathbf{p}$ x 1 vector of modal amplitudes, and $\mathbf{x} = [\mathbf{x}, \mathbf{y}, \mathbf{z}]^T$ is the coordinate of a point in undeformed space. The parameters $\mathbf{u_i}$ have intitutive meanings. The lowest frequency modes are the rigid-body modes of translation (u_0 through u_2) and rotation (u_3 through u_5) which specify the values of $\mathbf{b}$ and $\mathbf{M}$. u_6 through u_8 control the x, y, and z radii, u_9 through $u_1 1$ are shears about the x, y, and z axes, and u_{12+3j} through $u_{12+3j+2}$ are tapering, bending an pinching around the j^{th} pairwise combination of the x, y, and z axes. Since the rigid body modes are calculated in the object's coordinate system, they are rotated to the global coordinates.

REFERENCES

[1] Tal Arbel, Peter Whaite, and Frank P. Ferrie. Recognizing volumetric objects in the presence of uncertainty. Technical Report TR-CIM-94-03, Center for Intelligent Machines, McGill University, Montréal, Québec, Canada, March 1994. Available via ftp at `ftp.cim.mcgill.ca` in `/pub/3d/papers/tr-cim-94-04.ps.gz`.

[2] A. H. Barr. Superquadrics and angle-preserving transformation. *IEEE Computer Graphics Applications*, 1:11–23, 1981.

[3] A. H. Barr. Global and local deformations of solid primitives. *Computer Graphics*, 18(3):21–30, 1984.

[4] K. J. Bathe. *Finite Element Procedures in Engineering Analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1982.

[5] T. Binford. Visual perception by computer. Proc. of the IEEE conference on Systems and Control, Miami, FL, December 1971.

[6] F. Blais, M. Rioux, and J. A. Beraldin. Laser range finder development for 3-d vision. *Vision Interface '89*, June 1989. Canadian Image Processing and Pattern Recognition Society.

[7] T. E. Boult and A. D. Gross. Recovery of superquadrics from depth information. volume 1987 of *Proc. Spatial Reasoning and Multi-Sensor Fusion Workshop*, pages 128–137.

[8] R. Brooks. Model-based 3-d interpretations of 2-d images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 5(2):140–150, 1983.

[9] F. P. Ferrie, J. Lagarde, and P. Whaite. Darboux frames, snakes, and super-quadrics: Geometry from the bottom-up. volume 15, pages 771–784, 1993.

[10] F. P. Ferrie, S. Mathur, and G. Soucy. *3D Object Recognition Systems*, chapter Feature extraction for 3D model building and object recognition. Elsevier, Amsterdam, 1993.

[11] M. Gardiner. The superellipse: a curve that lies between the ellipse and the rectangle. *Scientific American*, September 1965.

[12] A. Gupta and R. Bajcsy. Volumetric segmentation of range images of 3d objects using superquadric models. *Computer Vision Graphics and Image Processing*, 58:302–326, 1993.

[13] B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker. Shapes, shocks, and deformations, i: The components of shape and the reaction-diffusion space. Technical report, Brown University, May 1992.

[14] A. Lejeune and F. P. Ferrie. Partitioning range images using curvature and scale. Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York City, New York, June 15-17 1993.

[15] D. Marr. *Vision*. Freeman, San Francisco, CA, 1982.

[16] D. Marr and H. K. Nishihara. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. Roy. Soc. London B*, 200:269–294, 1978.

[17] L. R. Nackman and S. M. Pizer. Three-dimensional shape description using symmetric axis transform theory. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 7(2):187–202, 1985.

[18] R. Nevatia and T. O. Binford. Description and recognition of complex-curved objects. *Artificial Intelligence*, 8:77–98, 1977.

[19] A. Pentland. Perceptual organization and representation of natural form. *Artificial Intelligence*, 28(3):293–331, 1986.

[20] A. Pentland. Automatic extraction of deformable part models. *International J. Computer Vision*, pages 107–126, 1990.

[21] A. Pentland. Computational complexity versus virtual worlds. *Computer Graphics*, 24(2):185–192, 1990.

[22] A. Pentland, I. Essa, M. Freidmann, B. Horowitz, and S. Sclaroff. 'the thingworld modeling system: Virtual sculpting by modal forces. *Computer Graphics*, 24(2):143–144, 1990.

[23] A. Pentland and B. Horowitz. Recovery of nonrigid motion and structure. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13:730–742, 1991.

[24] A. Pentland and S. Sclaroff. Closed-form solutions for physically-based shape modeling and recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13:715–729, 1991.

[25] A. Pentland and J. Williams. Good vibrations: Modal dynamics for graphics and animation. *Computer Graphics*, 23(4):215–222, 1989.

[26] A. Pentland and J. Williams. The perception of non-rigid motion: Inference of material properties and force. Proc. International Joint Conf. on Artificial Intelligence, August 1989.

[27] A. P. Pentland. Recognition by parts. International Conference on Computer Vision, pages 621–620, London, 1987.

[28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, New York, NY, 1988.

[29] S. Sclaroff and A. Pentland. Generalized implicit functions for computer graphics. *Computer Graphics*, 25(4):247–250, 1991.

[30] F. Solina and R. Bajscy. Range image interpretation of mail pieces with superquadrics. Proc. AAAI-87, pages 733–737, Seattle, WA, 1987.

[31] F. Solina and R. Bajscy. Recovery of parametric models from range images: The case of superquadrics with global deformations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 12(2):131–147, 1990.

[32] G. Soucy. View correspondence using curvature and motion consistency. Master's thesis, McGill University, Dept. of E.E., 1991.

[33] D. Terzopoulos and D. Metaxas. Dynamic 3d models with local and global deformations: Deformable superquadrics. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(7):703–714, 1991.

[34] D. Terzopoulos, A. Witkin, , and M. Kass. Symmetry-seeking models and 3d object reconstruction. *Int. J. Comput. Vision*, 1:211–221, 1987.

[35] P. Whaite and F. P. Ferrie. From uncertainty to visual exploration. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(10):1038–1049, 1991.

[36] P. Whaite and F. P. Ferrie. Autonomous exploration: Driven by uncertainity. IEEE Computer Society Conference on computer vision and pattern recognition, Seattle, Washington, June 1994.

[37] A. Witkin, A. Barr, and K. Fleischer. Energy constraints on parametrized models. *Computer Graphics*, 21(4):225–232, 1987.

[38] Kenong Wu and M. D. Levine. 3-d object representation using parametric geons. Technical report, McGill University, 1993.

Document Log:

Manuscript Version 0

Typeset by $\mathcal{AMS}$-LaTeX — 21 May 1996

SANTANU ROYMOULIK