On the Recovery of Superellipsoids

Terrance E. Boult and Ari D. Gross
Columbia University Department of Computer Science
New York City, New York, 10027. tboult@cs.columbia.edu. ari@sylvester.columbia.edu

Abstract

Superellipsoids are parameterized solids which can appear like cubes or spheres or octahedrons or 8-pointed stars or anything in between. They can also be stretched, bent, tapered and combined with boolean to model a wide range of objects. Columbia's vision group is interested in using superquadrics as model primitives for computer vision applications because they are flexible enough to allow modeling of many objects, yet they can be described by a few (5-14) numbers. This paper discusses research into the recovery of superellipsoids from 3-D information, in particular range data.

This research can be divided into two parts, a study of potential error-of-fit measures for recovering superquadrics, and implementation and experimentation with a system which attempts to recover superellipsoids by minimizing an error-of-fit measure.

This paper presents an overview of work in both areas. Included are data from an initial comparison of 4 error-of-fit measures in terms of the inter-relationship between each measure and the parameters defining the superellipsoid. Also discussed is an experimental system which employs a nonlinear least square minimization technique to recover the parameters. This paper discuss both the advantages of this technique, and some of its major drawbacks. Examples are presented, using both synthetic and range-data, where the system successfully recovers superlliposids, including "negative" volumes as would occur if superellipsoids were used in a constructive solid modeling system.

1 Introduction

As a modeling primitive, superquadrics provide a natural extension to traditional CAD models, [Barr 81]. Recently they have become the focus of a few computer vision research efforts, see [Pentland 86a], [Pentland 87], [Bajcsy and Solina 87b], [Solina 87], and [Boult and Gross 87b]. Most of the research to date has concentrated on a subclass of superquadrics called superellipsoids (hereafter SEs). The growing research interests in SEs can partially be attributed to the fact that they can mocel a wide range of object in a very compact form. Another contributing factor is the underlying mathematical formulation which provides convenient tools for their recovery.

The next section of this paper presents a mathematical definition of superellipsoids and a little motivation for using SE's in computer vision. Section 3 defines various error-of-fit measures which could be used in the recovery of SE's, and presents an initial comparison of these measures.

The current system is described in Section 4. Briefly it employs a nonlinear least square minimization technique on the inside-outside function to recover the parameters. Oddly, this function is one of the poorer error-of-fit measures considered, yet it is sufficient to allow for reasonable (qualitive) recovery of SEs from both synthetic and range data. A few examples of the results of the system are then presented. The examples include the use of very sparse data, synthetic multiple views and the recovery of a negative superellipsoid from range data.

Previous papers by the authors, [Boult and Gross 87b] and [Boult and Gross 87a], discuss a few other difficulties that are encountered when modeling objects with superellipsoids, or attempting to recover superquadrics from 3D data. These difficulties include the general problems associated with non-orthogonal representations, the difficulties of dealing with objects which are not exactly representable with CSG operations on the primitives, the need to recognize negative objects, and certain numerical instabilities.

Some pros and cons of the approach as well as few conclusions, and a discussion of work in progress at Columbia's vision lab appear in the final section.

The main result in this paper is that there are problems and biases with currently used error-of-fit measures, as well as with two newly proposed measures. It also suggests that even better measures await definition, especially ones measuring error using knowledge of the sensor direction. Another important result of the research is that least square minimization, (even using a poor error of fit measure), allows recovery of both positive and negative instances of superellipsoids from depth data.

2 Background and Motivation

Mathematically, SE solids are a spherical product of two superellipse curves (see [Gardiner 65]). Superellipse curves are similar to traditional ellipses except the terms in the definition are raised to parameterized exponents (not necessarily integers), i.e.: $(\frac{x}{a})^{\frac{2}{\epsilon}} + (\frac{y}{b})^{\frac{2}{\epsilon}} = 1$. When ϵ, the *relative shape parameter* is 1, the curve describes an ellipse. As the relative shape parameter varies from 1 down to 0, the shape becomes progressively squarish; as it varies from 1 toward 2, the shape transforms from a ellipse to a diamond shaped bevel. When the parameter is greater than 2, the shape becomes pinched and as the parameter approaches

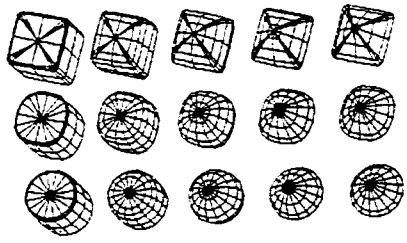

Figure 1: Superellipsoids with relative shape parameter ϵ_1 having values .1, .5, 1, 1.5 and 2 (left to right), and ϵ_2 having values .1, .6 and 1 (top to bottom).

infinity, the shape approaches a cross.

The result of the spherical product of two such curves is conveniently represented in a parametric form, e.g., a superellipsoid can be represented as:

$$\vec{s}(\eta,\omega) = \begin{bmatrix} a_x \cdot \cos^{\epsilon_1}\eta \cdot \cos^{\epsilon_2}\omega \\ a_y \cdot \cos^{\epsilon_1}\eta \cdot \sin^{\epsilon_2}\omega \\ a_z \cdot \sin^{\epsilon_1}\eta \end{bmatrix}, \quad \begin{matrix} -\frac{\pi}{2} \le \eta \le \frac{\pi}{2} \\ -\pi \le \omega \le \pi, \end{matrix} \qquad (1)$$

for any fixed positive $a_x, a_y, a_z, \epsilon_1$, and ϵ_2.

The parameters a_x, a_y, a_z affect the size of the superellipsoid along the x, y and z axes (respectively) in the object centered coordinate system. The parameters ϵ_1 and ϵ_2 effect the relative shape of the superellipsoid in the latitudinal (xz) and longitudinal (xy) directions, see figure 1. When the 5 parameters are all unity, the superellipsoid is the unit sphere.

Superellipsoids can also be defined implicitly as the the volume where the SE inside-outside function is negative, where the SE inside-outside function is given by:

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

where

$$\text{if} \begin{cases} f(\hat{x},\hat{y},\hat{z}) = 1, \text{ then } \hat{x},\hat{y},\hat{z} \text{ is on the surface,} \\ f(\hat{x},\hat{y},\hat{z}) < 1, \text{ then } \hat{x},\hat{y},\hat{z} \text{ is inside the volume,} \\ f(\hat{x},\hat{y},\hat{z}) > 1, \text{ then } \hat{x},\hat{y},\hat{z} \text{ is outside the volume.} \end{cases} \qquad (2)$$

These equations define superellipsoids (hereafter referred to as SEs) which are probably the simplest of a family of surfaces called superquadrics. SE's are the only superquadrics which can be made into a convex solid. One can also define superhyperboloids of one or two sheets and supertoroids, see [Barr 81]. In addition to the variety of shapes defined by the basic superquadrics, Barr also discusses the application of angle-preserving transforms which allow translation, rotation, bending and twisting. With the addition of tapering and traditional boolean combination operations, superquadrics become a powerful modeling tool, see Figure 2 for some examples generated using SuperSketch, [Pentland 86b].

While translation, rotation, bending and tapering are important modifications to basic SE primitives, they will not be considered in the remainder of the paper. Future work will examine the

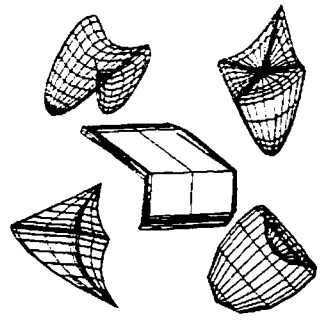

Figure 2: Examples of SE's deformed by bending and tapering

effect of these deformations on error-of-fit measures and potential implementations.

Definitions in hand, let us now discusses some of the rational for using SE's for object recognition. In short, they are flexible enough to represent a wide class of objects, but are simple enough to be recovered from 3-D data. Their functional form and inside-outside function provides a useful tools to aid in their recovery.

Traditional constructive solid modeling systems use boolean operations to combine primitives, such as, spheres, cylinders, and rectangular solids. By extending the primitive shapes to SEs, we allow the user to easily produce a continuum of forms, from spheres, to cuboids with rounded corners, to cubes, to diamond shapes. Such objects are difficult to model with traditional constructive solid geometry (hereafter CSG) systems. Making them primitives simplifies the job of the designer for any problem that contains objects with these properties.[1]

However, adding flexibility above that of a traditional CSG system is not the only reason to choose superquadrics. For added flexibility, one could make the primitives of the system Generalized Cylinders or Generalized Cones, see [Brooks and Binford 80]. The problem with GCs is that recovering them is a difficult process, partially because each GC may require hundreds of parameters to describe it. The number of parameters necessary to define a GC depends on the complexity of the cross section function, the spine, and sweeping function. As in [Brooks 85], [Rao and Nevatia 86], or [Ponce, Chelberg and Mann 87], one generally has to greatly restrict the class of GCs allowed before one can reliably recover them.

It is interesting to note that superellipsoids and supertoroids can be defined as a subclass of GC; SEs are those straight spined GCs with their cross section and sweeping functions defined by superellipses.[2] If the SEs are bent or tapered, these deformations are applied to the spine and sweeping function respectively.

[1] A simple CAD system has been developed by A. Pentland, see [Pentland 86b] that combines SEs (constrained to only those that are convex solids) with CSG type operations plus bending and tapering. The ease with which people became accustomed to modeling with this system may be tied to their own internal representation of the objects. [Pentland 86a] presents arguments to motivate the use of superquadrics as modeling primitives by showing correspondences to human vocalization of object descriptions.

[2] The sweeping rule is highly nonlinear, and always goes to zero. The authors do not know of any system which can recover a subset of GC containing this class.

Thus superquadrics provide a restriction which may allow efficient recovery as well as a simpler set of "limb equations"(after [Ponce, Chelberg and Mann 87]). In addition, they provide mathematical properties (the explicit parametric definitions, inside-outside function (implicit definition) and a normal surface duality principle) which do not exist for most classes of GCs.

The inside-outside function provides a useful tool for recovery, because it provides a simple way to determine which of the data points are inside the surface. Furthermore, the value of the function grows as points are moved further from the surface. Thus it can be used in conjunction with a minimization technique to recover the surface. However, the inside-outside function is not a euclidean "error of fit" measure (see the following section), and therefore direct minimization of it will not necessarily provide satisfactory results. (However, in two independent implementations this technique has produced satisfactory results, see [Bajcsy and Solina 87b] and [Boult and Gross 87b], and section 4.)

Canonical superquadrics (except the supertoroids) have a desirable duality property. The superquadric normal vectors lie on a dual superquadric form; that is, if the normal vectors were translated to the origin, they would generate another superquadric of the same class such that the normal of the new surface (if translated to the origin) would produce a copy of the original superquadric (except for a translation). One can easily derive the form of the normal surface as a spherical product as well as its own inside-outside function, see [Barr 81]. Note that the inside-outside function for the surface normals might be used to find superquadrics fitting surface normal information as would be available from shape-from-X methods.

3 Comparison of four "Error of fit" measures

The recovery of superellipsoids from depth data has recently been the subject of research in various labs, e.g. see [Pentland 87], [Bajcsy and Solina 87b], and [Boult and Gross 87b]. Each of the approaches has something in common: at some point they define some measure of the "error of fit" (hereafter EOF), and they use a nonlinear minimization technique to recover the parameters of the superellipsoid starting from an initial estimate.[3] It is well known, from work on fitting curves to points in the plane, that different EOF functions can result in radically different curve estimations, e.g. see [Pratt 87], [Sampson 82], [Bookstein 79] and [Turner 74]. On close examination of Columbia's initial system to recover SEs, it became apparent that the EOF function being used was not even proportional to sensor errors (determined using synethic data). The purpose of this section is to examine some of the measures of the error of fit used by these researchers as well as some measures not yet used in minimization procedures. While

[3]There are many differences in the details of the algorithms. The work of [Bajcsy and Solina 87b] and [Boult and Gross 87b] differ mostly in minor details, and both are driven mainly by the minimisation approach with only minor effort to derive good initial estimates of the parameters. The approach in [Pentland 87] is radically different using a detailed nonadaptive initial search of the parameter space (at 64,000 locations), but still maintains the minimization of some error of fit measure after that initial search through the parameter space.

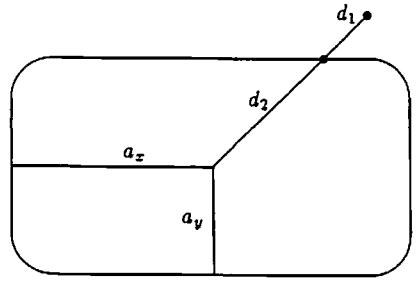

Figure 3: Drawing showing the intuitive meaning of measures. (This is an x-y cross section of an SE) $EOF_1 \approx \left(\frac{d_1}{d_1+d_2}\right)^{\frac{2}{\epsilon_1}}$, $EOF_2 \approx \left(\frac{d_1}{d_1+d_2}\right)^2$, $EOF_3 \approx (a_x \cdot a_y \cdot a_z) \cdot \left(\frac{d_1}{d_1+d_2}\right)^2$, $EOF_4 = d_1$

rotation and translation are not going to be considered herein, they are easily added to any of the measures under consideration.

In particular, this section considers 4 measures of error of fit. This section defines each of these mathematically and also gives a brief intuitive description in 2 dimensional physical/geometrical terms.

3.1 EOF_1: Mean-square-value of the inside-outside function

The first measure to be considered is, in some sense, the simplest. It is directly based on the inside-outside function which can be used to implicitly define the SE.

$$EOF_1 = \frac{1}{n} \sum_{i=1\ldots n} \left(1 - \hat{f}(x_i, y_i, z_i; \epsilon_1, \epsilon_2, a_x, a_y, a_z)\right)^2 \qquad (3)$$

where x_i, y_i, z_i are the given depth data points converted to a single viewpoint, and where $\hat{f}(x_i, y_i, z_i; \epsilon_1, \epsilon_2, a_x, a_y, a_z) = f(x, y, z)$ where $\epsilon_1, \epsilon_2, a_x, a_y, a_z$ are parameters in equation 2. A rough intuitive definition of this measure, see figure 3, is as a power of the ratio of the length of two line segments. The exponent to which the ratio is raised depends on the squareness-roundness in the xz direction and is given by $\frac{2}{\epsilon_1}$. The denominator of the ratio is the length of the line segment between the data point and the center of the SE. The numerator is the length of that portion of the above mentioned line segment connecting the center of the SE with that point on the surface in the direction of the data point. Given this intuitive definition, it is obvious that the measure can decrease if the object size grows even if the actual distance to the surface increases. Additionally, for small values of ϵ_1, this measure can be quite large (or small depending on ratio), and computationally unstable. This will be borne out through the computational evaluation in later sections.

3.2 EOF_2: "Corrected" mean-square inside-outside measure

The second measure to be considered is a modified version of the first to remove some its most undesirable properties: the rapid grown of EOF_1 for small values of ϵ_1, and some of its biases. The new measures is given by:

$$EOF_2 = \frac{1}{n} \sum_{i=1...n} (1 - \hat{f}^{\epsilon_1}(x_i, y_i, z_i, \epsilon_1, \epsilon_2, a_x, a_y, a_z))^2 \quad (4)$$

where terms are as defined for EOF_1

The intuitive definition of this measure is as the same ratio as for EOF_1 except that the power to which the ratio is raised is simply 2. Thus, this measure gives a type of "relative error" in a radial direction, however, it greatly effects measurements in the z direction (the drawing above ignores that direction). As before, it is obvious that the measure can decrease if the object size grows even if the actual distance to the surface increases. Note that to date this particular measure has not been used in a SE recovery system.

3.3 EOF_3: a " minimal volume" function

Another EOF function, used by [Bajcsy and Solina 87b], is

$$EOF_3 = \frac{1}{n} \sum_{i=1...n} \left(\sqrt{a_x \cdot a_y \cdot a_z} \cdot (1 - \hat{f}^{\epsilon_1}(x_i, y_i, z_i, \epsilon_1, \epsilon_2, a_x, a_y, a_z)) \right)^2 \quad (5)$$

where terms are as defined for EOF_1.[4]

A variant of this error of fit measure was heuristically first motivated in [Bajcsy and Solina 87b] by the idea that there are many surfaces which might fit data from a single view, and the system should recover the surface of minimal volume. The actual equations 10 and 11 in [Bajcsy and Solina 87b], as well as similar equations in [Bajcsy and Solina 87a], suggest a multiplicative factor of $a_x \cdot a_y \cdot a_z$ rather than $\sqrt{a_x \cdot a_y z \cdot a_z}$. However, this is not in keeping with the heurstic definition, and the latter form appears in [Solina 87]. While not reported here, experimental comparisons have found the the factor of $a_x \cdot a_y \cdot a_z$ results in a meausre which has similar properties, but exaggerates the poor behavior of EOF_3.

The rough intuitive definition of EOF_3 is simply the square of the "relative error" in the radial direction multiplied by the square-root of volume of the cube bounding the SE (which is meant to be a crude approximation to the SE itself). The advantage of this is that in a pure relative error measure, the absolute error in the radial direction was divided by a term containing the the radial distance to the surface in some direction. Thus with the points inside the object, the object could grow and reduce EOF_1 or EOF_2 (with respect to one point) while the absolute error would grow. By multiplying by the volume of the bounding rectangular solid, this behavior is curtailed. However, this behavior is replaced with the possibility of reducing the measure by keeping the relative error in the radial direction constant, but reducing the size of the SE (i.e., the measure is sensitive to scale).

[4] Actually, there a few minor differences between this definition and the ones in [Bajcsy and Solina 87b]. The first difference is the multiplication by $\frac{1}{n}$ which normalises by the number of data points. The second difference is that they use a homogeneous transformation to deal with rotation and translation, thus the derivation of EOF_3 in that paper is not given in terms of the $\hat{f}$ above, but this formulation can be easily shown to be equivalent.

3.4 EOF_4: A measure based on true euclidean distance

A final test measure is defined as the mean distance between each data point and the corresponding point on the surface of the superquadric on the line connecting the data point and the center of the SE. In equation form this becomes:

$$EOF_4 = \frac{1}{n} \sum_{i=1...n} |\vec{x_i} - \vec{q}| = \frac{1}{n} \sum_{i=1...n} \sqrt{(x_i - q_{i_x})^2 + (y_i - q_{i_y})^2 + (z_i - q_{i_z})^2} \quad (6)$$

where

$$\vec{q_i} = \vec{s}(\eta, \omega), \ \omega = \left(tan^{-1} \left(\frac{a_x \cdot y_i}{a_y \cdot x_i} \right)^{\frac{1}{\epsilon_2}} \right), \quad \eta = \left(tan^{-1} \left(\frac{a_z \cdot z_i}{a_x \cdot x_i} \cdot sin^{\epsilon_2}(\omega) \right)^{\frac{1}{\epsilon_1}} \right), \quad (7)$$

and $\vec{s}(\cdot, \cdot)$ is defined as in equation 1, and the relative signs of x_i, y_i, z_i, the i^{th} data point, are used to determine the correct spherical quadrant of the arctan.

This measure overestimates the minimal distance of a point from the SE, especially for squarish ones. However, it is a true euclidean distance measure, and thus is not effected by overall scaling of the problem. There are probably better euclidean based measures, however, the only better measures currently known to the authors do *not* have a functional form defining them, and thus are difficult to minimize.

3.5 Evaluation of the Different Error of Fit Measures

This section begins the analysis of the various error of fit measures. The main vehicle for presentation will be graphical, with running commentary giving *one* interpretation of the results. The reader is advised to draw her/his own conclusions.

The evaluation is in terms of two properties of the measures: parameter bias, and shape of crossections of the error surface. The parameter bias makes it difficult to compare two potential solutions and also generally affects the minimization by shifting the location of the minima. The relation of the measure (in terms of the scale of values) to the RMS error of data is a by product of the experiment to determine bias. The shape of cross sections of the error surface is important because it affects both the possibility of minimization (in particular, is it relatively concave?), and the rate of convergence of iterative minimization techniques (how large is the gradient?).

3.5.1 Determination of Measure Bias

To compare the parameter bias, each measure will be computed using the "correct" values of the 5 parameters (i.e. the parameters used to generate a surface). Of course, because of the noise in the

data points, these parameters will generally not be those parameters for which a given measure is minimized. (More about that later). The different cases considered in this section should give the reader a good intuitive feeling for how the different measures are biased (with respect to the true RMS error) as the parameters defining an SE are varied. To cut down on the amount of analysis, certain symmetries have been anticipated, in particular a_y is assumed to be identical to a_x. The discussions in this section are a review of the computational experiment, and only represent a small sample of the experimental data.

Each of the "bias" graphs show a sequence of computational experiments where a single parameter is varied over some range. The values are scaled, and the pattern of the scaled values often provide a good feel for the bias of measure. The scale values themselves show the relative difference in the size of the error measure (using the "correct" answer) with respect to the RMS error added to the data.

The data in each graph was computed as follows:

Step 1 An initial set of parameters for the SE, error range, and sensor direction (x, y, or z) are chosen. (Only x direction presented here). Also, a parameter is chosen to vary in the iterations. Each iteration generates one data point on the graph for each of the different measures, computed in steps 2 through 5.

Step 2 For a given set of SE parameters (i.e. each iteration), 500 points are generating on the surface of the SE assuming a uniform pseudo-random distribution of angles on the viewing hemisphere for the chosen sensor direction.

Step 3 For each data point, a pseudo-random noise value is computed assuming a uniform distribution in a predetermined range. The noise is added to the data point in the direction of the sensor. The value is directly added to the variable used to compute the RMS measure.

Step 4 Using the exact value of the set of parameters for this iteration, and the noisy data computed in steps 2 and 3, the value of each of the measures EOF_1, EOF_2, EOF_3 and EOF_4 is computed and recorded.

Step 5 For each measure (including RMS) the recorded values are examined and the maximum of a measure is used to re-scale its value into [0,1] and the value is entered on the graph.

Before beginning the description of the initial results, it is necessary to make a few comments about the graphical presentation of the results. First note that each graph contains 5 different marks which are used to show the scaled value of the measures. Also, each graph has a legend *with scaling factors*. Thus in figure 4, the computed value of EOF_1 for $\epsilon_{xy} = .5$ is approximately $.64 \times .176$, and when $\epsilon_{xy} = 1.8$ is approximately $.91 \times .176$. Looking at the tick-marks allows one to obtain a feel for the measure's bias as the parameter varies, and the scaling provides a means of relating the measure to the actual RMS error. For space considerations, only a few examples are presented here. However, the running commentary in this section is based on the analysis of over 500 of the above computational experiments.

The interest in bias factors stems from their importance in comparing the realtive quality of two potential solutions. For example, if the bias in a parameter P is toward larger values then if two candidate sets of parameters differ only in parameter P, the it can occur that the set with the larger P value may report a larger measure but have a smaller actual RMS error. Thus to properly compare two potential solutions, one needs to be able to estimate the bias in each of the parameters in which they differ.

The reader should recall that these graphs are *not* showing the shape of the error surface near the correct solution, and the non-monotonic behavior does not mean that minimization will be difficult.

The first graphical display for the bias experiment is in figure 4, and involves the bias of the measures with respect to the parameter ϵ_{xy}, the squareness-roundness in the xy direction. The first thing that should strike you about the graph is that all of the measures clump together, and none of them show a particularly smooth bias. There is a general trend toward having a large measure for ϵ_{xy}, but locally it is far from monotonic. Thus, the result of the minimization process on all the measures will generally report some value which overestimates the error (even if it could be scaled) and the overestimate is a generally increasing (but nonmonotonic) function ϵ_{xy}.

One interpretation of the clumping of the measures is that the observed bias is caused by measurement in the radial direction when the sensor error is in an axial direction. This interpretation suggests that no radial measure will be bias-free and provides an impetus to attempt to derive measures in the sensor direction.

The scales of the measured values were generally close to the RMS error, with EOF_1 very close (by accident), and EOF_3 EOF_4 off by a factor of 2. Although not evident from the examples presented here, EOF_3 is greatly effected by scaling. If all sizes in the example were multiplied by a factor of 100 (say meters into mm) then EOF_3 would increase by a factor of 1000000. On the otherhand, EOF_1 would decrease (by a much smaller amount) as the scale was increased.

The second example in the bias determination experiment is shown in figure 5. The most striking feature of this graph radically different behavior of the measures. Note that EOF_4 follows the RMS error rather well, showing that it is not biased with respect to ϵ_{xz}. EOF_2 and EOF_3 (which differ only in scale) show the same nonmonotonic but generalling increasing bias as they did for ϵ_{xy}. (Close examination shows qualitatively similar small scale perturbations in EOF_2, EOF_3 and EOF_4. Again, these are probably biases inherent in any axial measure.) While the other measures show a bias for larger values, EOF_1 has a significant, and apparently monotonic, bias for smaller values.

A second fact to consider in this example is the relative scales of the measure. A glance at the scaling factors in the legend suggest that only EOF_4 is closly linked to the actual RMS error. The other measures are all off by at least an order of magnititude, or more.

The third example in the bias determination experiment is shown in figure 6. The most striking feature of this graph is the clear separation of the measures. EOF_2 tracks the RMS error rather well, showing that it is not biased over this range of a_x.

EOF_4 is also rather unbiased, with the exception of very small values of a_x, which have a negative bias. Both of the aforementioned measures show non-monotonicity in what bias they do have, but the deviations (when appropriately scaled) from the RMS error are not very significant. EOF_1 on the other hand, shows a strong, almost monotonic, bias toward smaller values of a_x, as does EOF_3. The bias of EOF_3 for smaller values might be attributied to the "volume" factor which is the only difference between it and EOF_2. While not shown here, the above patterns are still present when the sensor direction is changed to the z axis, although the rates of change of the bias (i.e. slopes of curves) are smaller.

The final example from the bias experiment, is shown in figure 7. Again there is a rather noticeable difference between the measures. This time it is EOF_1 which display virtually no particular bias, although there are some obvious local non-monotonicities. The other three measures all show a strong bias for larger a_x values, with the bias in EOF_3 slightly less than in the other measures. While one might have predicted a bias in EOF_3 for small values of a_x, the greater bias of the underlying EOF_2 factor dominates the "volume" term (but the latter probably provides the mitigating factor in the bias as compared with EOF_2 and EOF_4). To the eye, the bias trend in all but EOF_1 appear linear, and anticipation/compensation might be possible. Again, the experiments with the sensor direction along the z axis have similar overall characteristics, though the slopes of the bias "curves" were smaller.

As mentioned, there were also numerous test cases included in the experiment which were not reported here. To summarize the results, little characteristic difference was found to be caused by changes in the sensor direction, although many minor variations did occur. As mentioned above, many of the measures also displayed a bias in terms of overall scale of the problem.

A final comment about the bias experiments. The graphical presentation makes it very easy to overlook the absolute scaling parameters in the legend. However, if one were interested in converting from one measure into an approximate RMS measure, even knowing the bias characteristic would not be useful unless one could also compute the scaling factors. As shown in the next section, many of the measures can have significant scale changes in the neighboorhood of the correct solution, thus estimation of the scale factors from this experiment is unlikely.

3.5.2 Analysis of cross sections of error-of-fit surface

This section discusses experiments which attempts to determine the structure of the error-of-fit surface by examining one dimensional slices through that multi-dimensional surface. The experiment considered both the structure in a small neighborhood of the correct parameter value and the larger scale structure (say variations of 10-100%). The local structure provides insight into how the measures err (i.e. it shows location of measures minima). The large scale structure provide insight into the likelyhood of successful minimization when reasonable initial estimates are used as starting values. Because global properties are of more importance, only the experiments using a wide range of values will be reported here, even though some of the conclusion are more readily made by examination of the experiments in small neighborhoods.

As in the reporting of the bias experiments, the presentation will be graphical with commentary. Each graph was generated as follows:

Step 1 Parameters defining SE and sensor direction are chosen. Using these parameters, 500 points are generated on the surface of the SE assuming a uniform pseudo-random distribution of angles on the viewing hemisphere for the chosen sensor direction. For each data point, a pseudo-random noise value is computed assuming a uniform distribution in a predetermined range. The noise is added to the data point in the direction of the sensor. The value is directly added to the variable used to compute the RMS measure.

Step 2 A parameter is chosen to vary, as well as the range of variation. For each iteration point (10-20 spanning the chosen range) the experimental values for each measure are generated using the current value of the variable parameter, and the correct values of the remaining parameters, and are recorded.

Step 3 For each measure (including RMS) the recorded values are examined and the maximum of a measure is used to rescale its value into the interval [0,1] and the value is entered on the graph.

Note the above procedures result in the correct location of the minima occuring in the center of each graph.

Probably the most surprising result is that when rotations and translations are ignored (as in the results reported here), all of the measures generated relatively concave cross sections. Thus, minimization of any of them (given correct rotations and translations) should be straight forward. The main difference between the measures was the degree of concavity, which would effect the efficiency of the minimization.

The first presented crossection, figure 8, shows the cross section as ϵ_{x_1} is varied. The clear separation of EOF_4, and higher curvature suggest that it is the preferred measure in this example. Note that since a_x, a_y and a_z are fixed, the structure of EOF_2 and EOF_3 is identical; they differ only in a scaling factor. Finally, EOF_1 does show some concavity, but the dramatic scale differences make the minima very shallow. While not shown, the separation of the measures is even greater when the errors are along the z axis.

It is worth pointing out that EOF_2, EOF_3 and EOF_4 all have their minima to the left (i.e. smaller parameter value) than the underlying surface, while EOF_1 overestimates the parameter. (The detailed analysis of the local region which is not presented shows that the minima of EOF_4 is the closest to the parameter value defining the actual underlying surface.) The reader might want to compare this with the results of the bias experiment for the same parameter.

The second cross section presented in figure 9, shows all of the measures are very asymetric about their minima. In addition, although it cannot be seen from this graph, all of the measures have their minima above the "correct" value. The dramatic changes in scale for EOF_1 make it appear flat on this graph, but it does have a shape similar to the other measures, just more dramatic.

It is also obvious (though barely visible) that EOF_4 has the

sharpest minima, although minimization of any of the measures with a starting value above the true solution looks like a good idea. Again, EOF_4 has its minima closest to the correct value. The asymmetric shape of this curve is even more pronounced when the errors in the data are along the z axis.

A comment also is in order regarding the vast differences in scales between the measures. As one can see from the legend in figure 9, the relative scales differ by orders of magnitude. From a practical point of view, those measures which routinely have a wide numerical range (EOF_1, EOF_2 and depending on scaling possibly EOF_3) present more numerical problems, and show a greater sensitivity to roundoff errors.

A third cross section, see figure 9, shows a case where the relative shape of all of the measures is almost identical, with EOF_4 just slightly outperforming the other measures in the concavity analysis. Also, note that all of the measures again overestimate the correct underlying parameters (by about 1%-2%). When the error was in the direction of the z axis, all measures were very sharply concave, and the minima were very close to the correct value.

The final cross sections, see figures 11 and 11 show the structure of the error surfaces for some of the parameters used in the bias experiments. Note that in these cases, EOF_4 is a clear winner.

4 The Current System

This section presents some of the details of our current system for the recovery of superquadrics from 3-D information. The basis of our system is the minimization of EOF_1. Note that one of the advantages of minimizing the a measure based on the inside-outside function is that it requires little extra effort to incorporate multiple views, assuming one knows the sensor position for each view (to convert points to a common coordinate system). However, such an approach will not take error distributions into account, and the errors in conversion may be exaggerated by the fitting process.

Given 3-D information about a SE in canonical position, one can use a nonlinear minimization technique to recover the 5 parameters needed to define it. Our system uses a Gauss-Newton iterative nonlinear least square minimization technique, (for example see [Hageman and Young 70]). If the SE is not in canonical position, the system must also recover estimates of the translation and rotation necessary to put the information on the surface of a canonical SE. There are obviously many approaches to deal with the translation and rotation. The two most obvious are: the use of a pair of transforms (one for translation and one for rotation) and the use of a homogeneous transform that combines both translation and rotation. Our system uses the first approach.

For the remainder of this section let $C_\theta = \cos\theta$ and $S_\theta = \sin\theta$. Thus given a canonical SE surface defined as $\vec{s} = \vec{s}(\eta, \omega)$ with an inside-outside function $f = f(x, y, z)$, the translated and rotated SE solid $\hat{\vec{s}}$ is given by

$$\hat{\vec{s}} = R_\theta R_\phi R_\psi \vec{s} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}, \tag{8}$$

where θ, ϕ, and ψ are the Euler angles expressing pitch, yaw and roll respectively, $R_\theta R_\phi R_\psi$ are the associated rotation matrices and t_x, t_y, t_z are the translation in the x, y, and z directions respectively. The definition of the rotation matrices and Euler angles can be found in many standard texts on computer graphics, e.g., [Rogers and Adams 76], as well as in the computer vision literature e.g., see [Tsai 86], or [Bolle and Murthy 86].

Given the data for a general SE the system minimizes:

$$EOF_1(\epsilon_1, \epsilon_2, a_x, a_y, a_z, t_x, t_y, t_z, \theta, \phi, \psi) = \sum_{i=1\ldots n} \left(1 - \hat{f}(x_i, y_i, z_i, \epsilon_1, \epsilon_2, a_x, a_y, a_z, t_x, t_y, t_z, \theta, \phi, \psi)\right)^2. \tag{9}$$

where x_i, y_i, z_i are the given data points from a single viewpoint, and where $\hat{f}(x_i, y_i, z_i, \epsilon_1, \epsilon_2, a_x, a_y, a_z, t_x, t_y, t_z, \theta, \phi, \psi) = f(\hat{x}, \hat{y}, \hat{z})$ and

$$\begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = R^T \begin{bmatrix} x - t_x \\ y - t_y \\ z - t_z \end{bmatrix}.$$

where

$$R^T = \begin{bmatrix} C_\theta C_\phi & S_\psi S_\theta C_\phi - C_\psi S_\phi & C_\psi S_\theta C_\phi + S_\psi S_\phi \\ C_\theta S_\phi & S_\psi S_\theta S_\phi + C_\psi C_\phi & S_\psi S_\theta C_\phi - C_\psi S_\phi \\ -S_\theta & S_\psi C_\theta & C_\psi C_\theta \end{bmatrix} \tag{10}$$

To employ the Gauss-Newton iteration, the system must compute the Jacobian of the transformation, and thus also needs the partial derivatives of $\hat{f}(x, y, z)$ with respect to the 11 parameters, (5 shape, 3 rotation and 3 translation), which were obtained symbolically.

The initial implementation of the system recovered only the 5 shape parameters. Using synthetic data with up to 10% uniformly distributed noise, the system could start the minimization procedure from a canonical position (all 5 parameters = 1) and in most cases was still able to recover the underlying surface within the error of the data. Moreover, the convergence was generally quick, requiring < 15 iterations. These results are reported because, if though some other means one was able to compute the local coordinate system of the SE (or out very good bounds on it), the stability of the algorithm would be greatly increased.

When the system was extended to handle 11 parameters (i.e. 5 shape and 6 position/orientation), things became more complicated. For most of the examples presented, even when presented with very poor estimates for starting values, the system was quickly able to find an SE that had low "error of fit". Unfortunately, the solutions proposed by the system often seemed to be nonintuitive. However, when examined closely, many of these solutions proved themselves to be reasonable (though not the most reasonable) interpretations of the data. The differences could generally be traced to the nonintutative behavior of error-of-fit measure. For example, when presented with the range data (from the Utah range database, [UTAH 85]) for a coke can minus the concave portion of the bottom end, the system initially proposed a solution with $EOF_1 = .0078$. However, the parameters

described an object that is slightly beveled along its long axis, and rather round in the other direction. The length of the object was far larger then expected. The exaggerated length was caused by two factors:

1. the model is assumed to be intersected with a negative ellipsoid at the top, and cut by the ground plane at the bottom, and
2. the "error of fit" measure used by the system is an underestimate, and biased toward larger objects, see Section 3 for further discussion of this problem.

When examined closely, the proposed object (assuming the volume terminates when it intersects another object in the scene) does seem to be a reasonable fit to the data, but still a cylinder seems intuitively to be a better fit. If forced to look for a cylindrical object, the system finds an SE with EOF_1 of .17. Oddly enough, the system converged to a can more readily if data from the negative ellipsoid was not removed. As reported in the next section, the use of a different error-of-fit measure resulted in better recovery.

Currently, the system obtains initial estimates of the translation parameters from the centroids of the original data and derives bounds from sensor information and overall data (maximum variation in any data). However, there are many problems with these estimates especially if the number of data points is small or if the system is only given a partial view of an object. These estimates are obviously much better if multiple views of the object are available. The estimates also assume that the data is segmented, an assumption with which the authors feel particularly uncomfortable. The system derives estimates of rotation angles and length scales from moments of inertia and bounds on the length parameters from sensor information and overall data. Because the moments of inertia require second order moments, the estimates are plagued with more difficulties than the estimates of the translation parameters. The estimates of both the rotation and length scales are very poor if an object (after segmentation) is the result of boolean combinations.

The system also estimates bounds on the maximum values for each parameter, deriving these estimates from knowledge of the "sensor" and the data values. If during the minimization process, any parameter attempts to stray beyond its allowed boundary, the system stochastically "pushes" it back toward its initial value.

4.1 A Few Examples from Our Current System

The first example in figure 13, is a synthetic SE with noisy synthetic data from multiple views. The actual parameters of the superquadric are (1.59, .39, 1, 2, 3, 1.5, 2.5, 3.5, .1, .1, .1).[5] The noise in the underlying object was uniformly distributed over the interval [-.15, +.15] and then added to the z (depth) value of a point. The 1000 data points were randomly distributed on the surface before noise was added. The system recovered an SE with parameters (1.8, .3, 1, 2, 3.58, 1.49, 2.5, 3.47, .079, .09, .101). The error (as measured by EOF_1) of the reconstruction wasf .079. The system required 7 iterations from the initial values to find the solution.

Figure 14, is the recovery of the same synthetic SE as in example 1. However, this time this system was given 1000 data points from one view of the object. Under these conditions, the recovered parameters were (2.09, .67, 1, 1.94, 3.36, 1.43, 2.53, 4.3, .075, .07, .08). The error (as measured by EOF_1) of the reconstruction was .169. Unsurprisingly, the reconstruction from multiple views is superior.

The final three examples presented, show the fitting of actual range data from the Utah range database. Figure 15 shows the elliptical indentation on the bottom of a soda can. The object which is defined by 590 data points. The system recovered the parameters (1.29, .955, .939, .930, .277, -.096, -1.55, 2.07, -.05,0,0), and had EOF_1 = .0293. Figure 16 shows a quasi-spherical object which is defined by 859 data points. The system recovered the parameters (.994, .951, 1.19, 1.13, 1.13, .4729, 1.437, -1.457,-.05, -.04, 0) and EOF_1 = .021. Figure 17 shows the cylindrical portion of a soda can defined by 1645 data points. When using the above described estimations techniques, the system recovered the parameters (2.0, .88, 1, 1.3, 19.59, -.19, -1.69, -.834, -.04,.05, 0) and with EOF_1 = .0079. When the inside-outside function was modified to include an extra multiplicative factor of $a_x a_y a_z$, the result was the bottom object in figure 17.

5 Conclusions and Future Direction

The main result in this paper is that there are problems and biases with currently used error-of-fit measures, as well as with two newly proposed measures. Still, the newly proposed measures often fared better than currently used measures

The analysis also suggests that even better measures await definition, especially error of fit measures which make use of knowledge of a sensor error model (direction).

The paper also summarizes results which demonstrate that even with a poor EOF measure the system can recognize both positive and negative superquadrics from depth data on only part of the surface. (The latter an important consideration if SE's are to be used with CSG operations.)

Future plans for the system also include extensions to incorporate surface derivative information. This will be accomplished by minimizing a sum with (some variant of) both the inside-outside function and a differentiated form of the inside-outside function.

Of course, one of the most important avenues for future research will be attacking the segmentation problem, our current plans are to attempt at least two approaches: pure growing of superquadrics from small data patches, and a skeletonization (to find axis) followed by both growing and splitting of superquadric solids.

A final avenues of research will be in using SE's for integration of multi-sensor (possible multi-modal) information, including some model of sensing errors. Part of this work will undoubtedly look into the use of edges in an intensity image /range image as

[5] In this example and hereafter the parameters will always be presented as a 11 vector with association implied by order given by: $\epsilon_1, \epsilon_2, a_x, a_y, a_z, t_x, t_y, t_z, \theta, \phi, \psi$.

they relate to limb-equations.

Acknowledgments

This work was supported in part by Darpa Contract #N00039-84-C-0165. The authors are grateful to Alex Pentland, Frank Solina and R. Bajcsy for their discussions on this subject. An special thanks to Alex Pentland (and SRI) for allowing us to use Supersketch, which (with modifications) was used to generate the renderings of SE's in the paper.

References

[Bajcsy and Solina 87a] R. Bajcsy and F. Solina. *Range image interpretation of Mail pieces with Superquadrics.* Technical Report MS-CIS-87-18, GRASP LAB 98, University of Pennsylvania, March 1987.

[Bajcsy and Solina 87b] R. Bajcsy and F. Solina. Three dimensional object representation revisted. In *Proceedings of the IEEE Computer Society International Conference on Computer Vision*, pages 231–240, IEEE, June 1987.

[Barr 81] A. H. Barr. Superquadrics and angle preserving transformations. *IEEE Computer Graphics and Applications*, 1:11–23, January 1981.

[Bolle and Murthy 86] R.M. Bolle and S.S. Murthy. *Curvature extraction from approximations to image or range data.* Technical Report, IBM Research, AI Systems Group, May 1986. IBM Thomas J. Watson Research Center.

[Bookstein 79] F. L. Bookstein. Fitting conic sections to scattered data. *Computer Vision, Graphics, and Image Processing*, 9:56–71, 1979.

[Boult and Gross 87a] T.E. Boult and A.D. Gross. Recovery of superquadrics from 3d data. In *Proceedings of the SPIE Conference on Intelligent Robots and Computer Vision: Sixth in a Series*, SPIE, November 1987. to appear, Order # 848-55.

[Boult and Gross 87b] T.E. Boult and A.D. Gross. Recovery of superquadrics from depth information. In *Proceedings of the AAAI Workshop on Spatial-Reasoning and Multisensor Integration*, AAAI, October 1987.

[Brooks 85] R.A. Brooks. Model based 3-d interpretation of 2-d images. In A. Pentland, editor, *From Pixels to Predicated*, Ablex Publishing Co., Norwood, N.J., 1985.

[Brooks and Binford 80] R.A. Brooks and T. O. Binford. Representing and reasoning about specific scenes. In *Proceedings of the DARPA Image Understanding Workshop*, pages 95–103, DARPA, April 1980.

[Gardiner 65] M. Gardiner. The superellipse: a curve that lies between the ellipse and the rectangle. *Scientific American*, September 1965.

[Hageman and Young 70] L.A. Hageman and D.M. Young. *Applied Iterative Methods.* Academic Press, NYC, NY, 1970.

[Pentland 86a] A. Pentland. *Recognition by Parts.* Technical Report 406, SRI International, December 1986.

[Pentland 86b] A. Pentland. Towards an ideal 3-d cad system. In *Proceedings of the SPIE Conference on Machine Vision and the Man-Machine Interface*, SPIE, Janurary 1986. Order # 758-20.

[Pentland 87] A. P. Pentland. Recognition by parts. In *Proceedings of the IEEE Computer Society International Conference on Computer Vision*, pages 612–620, IEEE, June 1987.

[Ponce, Chelberg and Mann 87] J. Ponce, D. Chelberg, and W. Mann. Invariant properties of the projection of straight homogenous generalized cylinders. In *Proceedings of the IEEE Computer Society International Conference on Computer Vision*, pages 631–635, IEEE, June 1987.

[Pratt 87] V. Pratt. Direct least-squares fitting of algebraic surfaces. In *Proceedings of SIGGRAPH '87: Computer Graphics 21:4*, pages 145–152, ACM, July 1987.

[Rao and Nevatia 86] K. Rao and R. Nevatia. Generalized cone descriptions from sparse 3-d data. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 256–263, IEEE, 1986.

[Rogers and Adams 76] D.F. Rogers and J.A. Adams. *Mathematical Elements for Computer Graphics.* McGraw-Hill, New York, 1976.

[Sampson 82] P.D. Sampson. Fitting conic sections to "very scattered" data: an iterative refinment of the bookstein algorithm. *Computer Vision, Graphics, and Image Processing*, 18:97–108, 1982.

[Solina 87] F. Solina. *Shape recovery and segmentatin with deformable part models.* PhD thesis, University of Pennsylvania, Department of Computer Science., 1987. available as Tech. Report MS-CIS-87-111.

[Tsai 86] R.Y. Tsai. An efficient and accurate camera calibration technique for 3d machine vision. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 364–374, June 1986.

[Turner 74] K.J. Turner. *Computer perception of curved objects using a televesion camera.* PhD thesis, Dept. of Machine Intelligence, University of Edinburgh, 1974.

[UTAH 85] utah. *The University of Utah range database manual.* Technical Report , University of Utah, 1985.

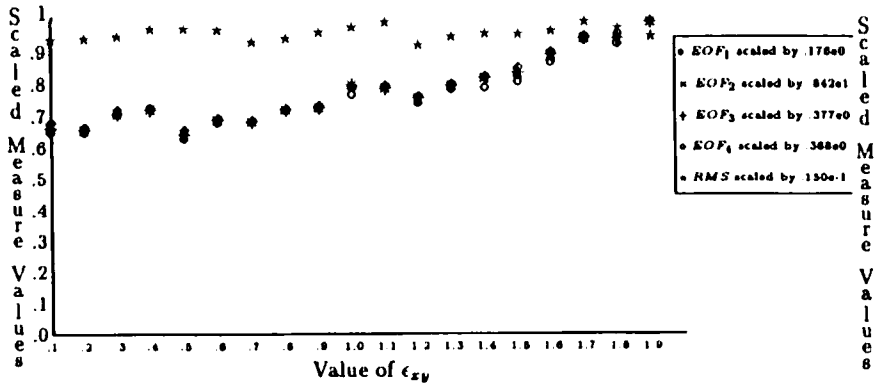

Figure 4: Bias experiments for ϵ_{xy} with $\epsilon_{xz} = .1$ $a_x = 1$, $a_y = 10$ and $a_z = 50$

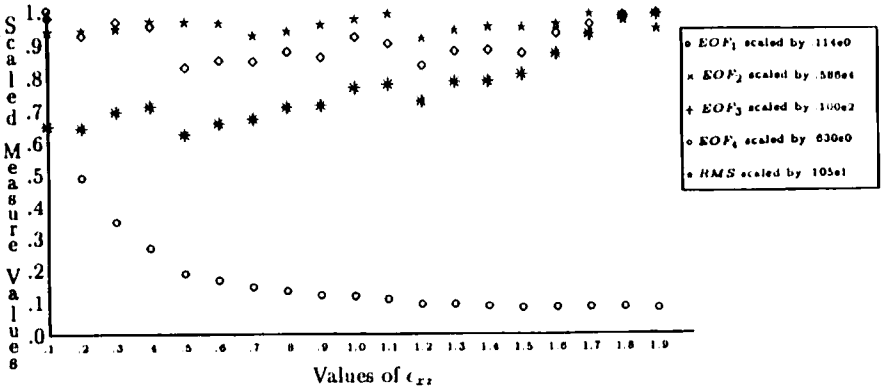

Figure 5: Bias experiments for ϵ_{xz} with with $\epsilon_{xy} = .1$, $a_x = 70, a_y = 70$ and $a_z = 70$

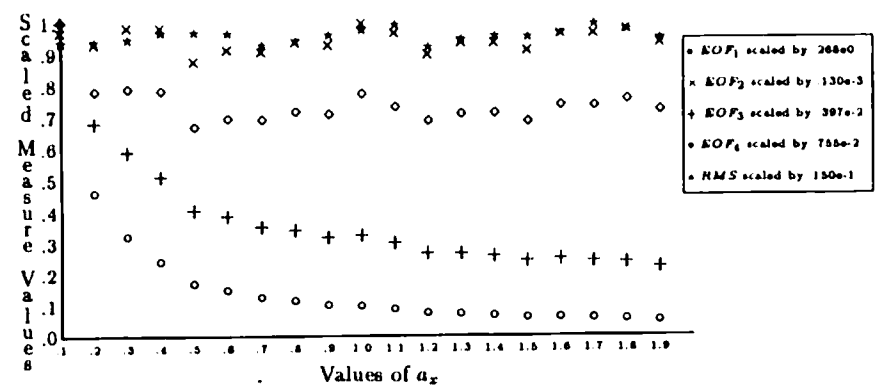

Figure 6: Bias experiments for a_x with $\epsilon_{xz} = .5$, $\epsilon_{xy} = .5$, $a_y = .1$ and $a_z = .1$

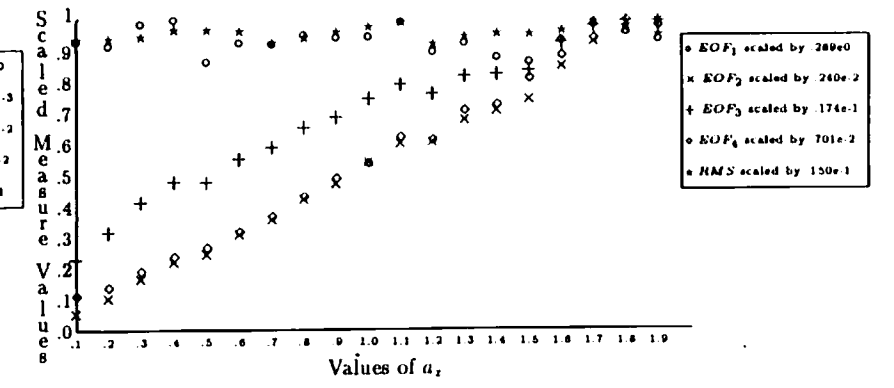

Figure 7: Bias experiments for a_z with $\epsilon_{xz} = .5$, $\epsilon_{xy} = .5$, $a_x = .1$ and $a_y = .1$.

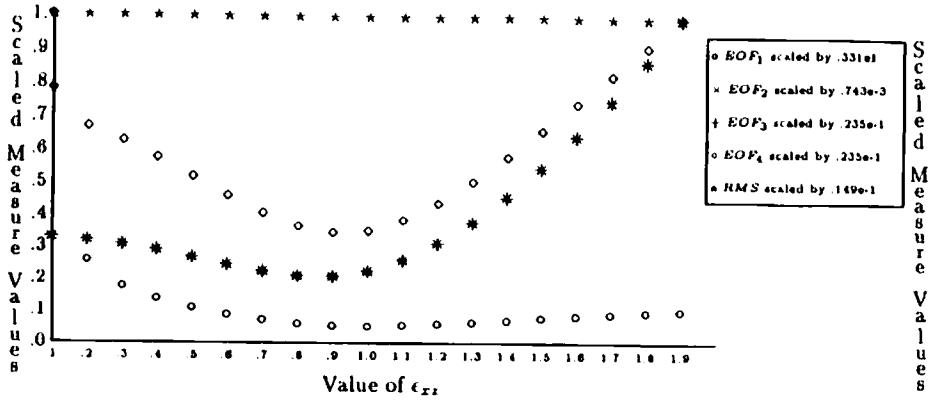

Figure 8: Cross section of error surface as one varies ϵ_{xz} with correct parameters: $\epsilon_{xz} = 1, \epsilon_{xy} = 1$, $a_x = .1, a_y = .1$ and $a_z = .1$

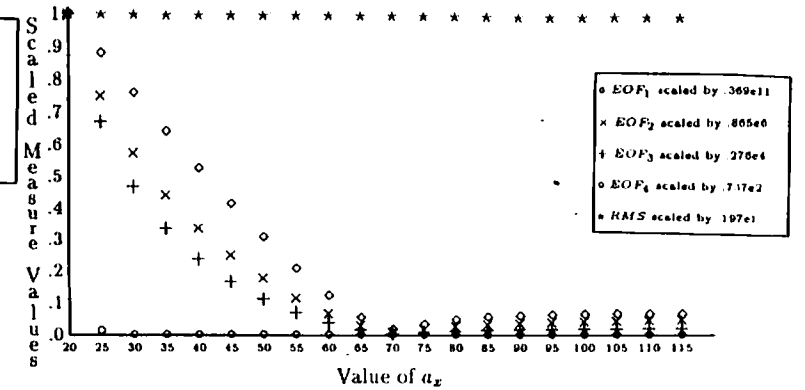

Figure 9: Cross section of error surface as one varies a_x with correct parameters: $\epsilon_{xz} = .1, \epsilon_{xy} = .1$, $a_x = 70, a_y = 70$ and $a_z = 70$

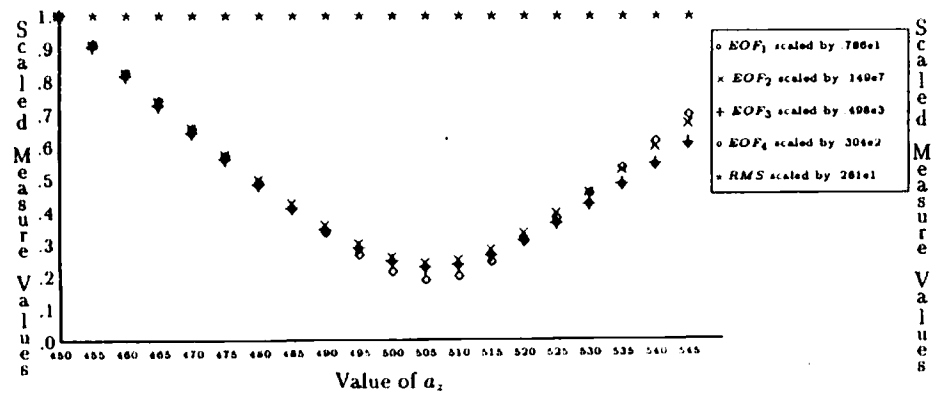

Figure
10: Cross section of error surface as one varies a_z with correct parameter: $\epsilon_{xz} = 2, \epsilon_{xy} = .1$, $a_x = 100, a_y = 200$ and $a_z = 500$

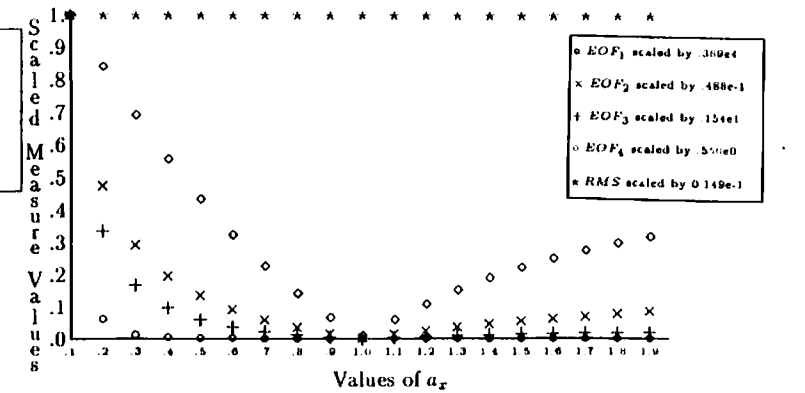

Figure 11: Cross section of error surface as one varies a_x with $\epsilon_{xz} = .5, \epsilon_{xy} = .5$, $a_x = 1.0, a_y = .1$ and $a_z = .1$

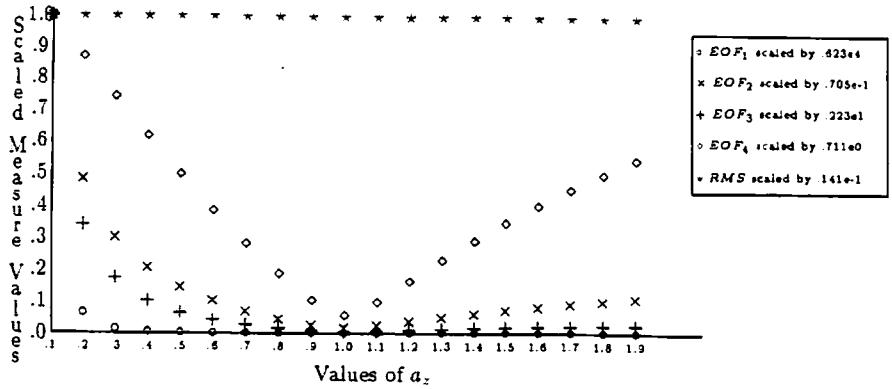

Figure 12: Cross section of error surface as one varies a_z with $\epsilon_{xz} = .5, \epsilon_{xy} = .5$, $a_x = .1, a_y = .1$ and $a_z = 1.0$

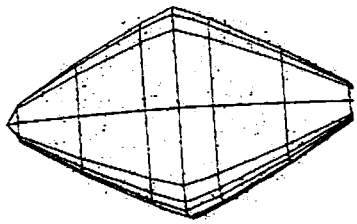

Figure 13: Example recovering 11 parameters using noisy synthetic data from single views

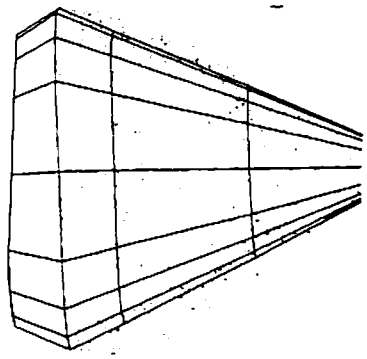

Figure 14: Example recovering 11 parameters using noisy synthetic data from single views

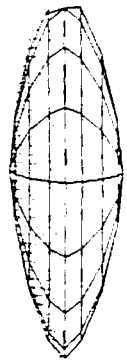

Figure 15: Reconstruction of a negative ellipsoid from real range data

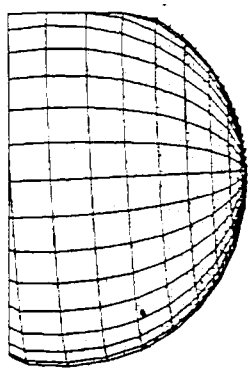

Figure 16: Reconstructed sphere from actual range data

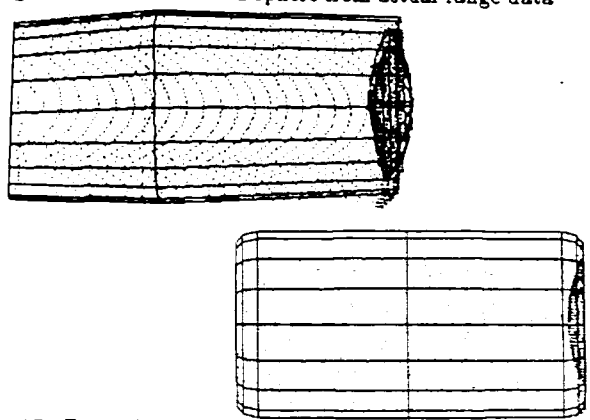

Figure 17: Example using real range data example of soda can, best fit surface with standard inside-outside function.(top) and with function multiplied by $a_x a_y a_z$ (bottom).