Volumetric Models in Computer Vision – an Overview*

Franc Solina
Faculty of Electrical Engineering and Computer Science, University of Ljubljana, Ljubljana, Slovenia

Volumetric models are the top-level representation in computer vision applications. Volumetric models are especially suited for part-level representation on which manipulation, recognition and other reasoning can be based. The two most popular types of volumetric models are generalized cylinders and superquadrics. This paper gives an overview of research and applications of volumetric models in computer vision and robotics with emphasis on superquadrics.

1. Introduction

The goal of computer vision and robotics is to enable intelligent interaction of artificial agents with their surroundings. The means of this interaction are images of various kinds; intensity images, pairs of stereo images, range images or even sonar data and information from contact sensors (touch, force, torque). Images which at the sensory level consist of several hundreds or thousands of individual image elements must in this process be encoded in a more compact fashion. For any reasoning or acting on the surroundings, it is advantageous that this coding of images as well as the internal representation of the robot work space closely reflects the actual structure. Distinct objects, for example, should have distinct models of themselves. In this way, the labeling of individual entities, necessary for control and higher level reasoning, becomes possible.

So far, many different models have been used for modeling different aspects of objects and scenes. Models for representing 3D structures can be grouped into local and global models. Methods for local representation attempt to represent objects as sets of primitives such as surface patches or edges. Global methods on the other hand attempt to represent an object as an entity in its own coordinate system. When objects of such global models correspond to perceptual equivalents of parts, we speak of part-level models. Several part-level models are required to represented an articulated object. A part-level shape description is important for several tasks involving spatial reasoning, object manipulation, and structural object recognition. People often resort to such part description when asked to describe natural or man-made objects [35]. Such part descriptions are generally suitable for path planning or manipulation—for object-recognition, however, they are sometimes not malleable enough to represent all necessary details and several researchers are looking into extending part-level models with additional layers of details.

1.1. Generalized Cylinders

The first dedicated part-level models in computer vision were generalized cylinders [8]. A generalized cylinder, sometimes called a generalized cone, is represented by a volume obtained by sweeping a two-dimensional set or volume along an arbitrary space curve. The set may vary parametrically along the curve. Different parameterizations of the above definition are possible. In general, a definition of the axis and the sweeping set are required. To limit the complexity and simplify the recovery of generalized

* This research was supported in part by The Ministry for Science and Technology of The Republic of Slovenia (Projects P2-1122 and J2-6187).

cylinder models from images researchers often used only straight axes and constant sweeping sets. Properties of straight homogeneous general cylinders are addressed in [39]. Generalized cylinders influenced much of the model-based vision research in the past two decades—theory [33, 7] as well as actual building of vision systems [10]. In general, one can criticize the methods of recovering generalized cylinders on the count that they are not very robust because they must rely on complicated rules for grouping low level image models (i.e. edges, corners, surface normals) into models of larger granularity (i.e. symmetrical contours or cross-sections) to arrive finally to generalized cylinders. These problems are due in part to the complicated parameterization of generalized cylinders and to the lack of a fitting function that would enable a straightforward numerical examination of the model's appropriateness for the modeled image data. The latest achievements in recovery and segmentation of strait homogeneous generalized cylinders can be seen in [32].

Several other global models exist that attempt to represent an object as an entity in their own coordinate system: Spherical harmonic surfaces [41], Gaussian images and extended Gaussian images [25], Symmetry seeking models [46], Blobby model [31], Hyperquadrics [22], and Superquadrics.

Superquadric models appeared in computer vision as an answer to some of the problems with generalized cylinders [35]. Superquadrics are solid models that can, with a fairly simple parameterization, represent a large variety of standard geometrical solids as well as smooth shapes in between. This makes them much more convenient for representing rounded, blob-like shaped parts, typical for objects formed by natural processes.

The rest of the paper is divided as follows: in the next section we define superquadrics. In the second section we talk about recovery of superquadrics and applications in computer vision. The third section is on usage of superquadrics in robotics. In the conclusions we summarize the advantages of superquadric models and point to possible future developments, especially in human-computer interface design.

2. Superquadrics

A superellipse is a closed curve defined by the following simple equation:

$$\left(\frac{x}{a}\right)^m + \left(\frac{y}{b}\right)^m = 1,$$

where a and b are the size (positive real number) of the major and minor axes and m is a rational number

$$m = \frac{p}{q} > 0, \quad \text{where} \begin{cases} p & \text{is an even integer,} \\ q & \text{is an odd integer.} \end{cases}$$

If $m = 2$ and $a = b$, we get the equation of a circle. For larger m, however, we get gradually more rectangular shapes, until for $m \rightarrow \infty$, the curve takes up the shape of a square. Superellipses are special cases of curves which are known in analytical geometry as Lamé curves [30][1]. Piet Hein, who popularized these curves for design purposes also made a generalization to 3D which he named *superellipsoids* or *superspheres* [15]. The final mathematical foundation of superquadrics was laid out by Barr [5], who generalized the whole family of quadric surfaces with the help of varying exponents, and coined a new name for them—*superquadrics*. Superquadrics are by definition a family of shapes that includes not only superellipsoids, but also superhyperboloids of one and of two pieces, as well as supertoroids.

The explicit superellipsoid equation, defined by the following surface vector, is

$$\mathbf{x}(\eta, \omega) = \begin{bmatrix} a_1 \cos^{\varepsilon_1}(\eta) \cos^{\varepsilon_2}(\omega) \\ a_2 \cos^{\varepsilon_1}(\eta) \sin^{\varepsilon_2}(\omega) \\ a_3 \sin^{\varepsilon_1}(\eta) \end{bmatrix} \quad \begin{array}{c} -\pi/2 \leq \eta \leq \pi/2 \\ -\pi \leq \omega < \pi \end{array}, \tag{1}$$

where a_1, a_2 and a_3 determine size, and ε_1 and ε_2 determine global shape. The alternative, implicit superellipsoid definition, also called the *inside-outside* function is

$$\left(\left(\frac{x}{a_1}\right)^{\frac{2}{\varepsilon_2}} + \left(\frac{y}{a_2}\right)^{\frac{2}{\varepsilon_2}}\right)^{\frac{\varepsilon_2}{\varepsilon_1}} + \left(\frac{z}{a_3}\right)^{\frac{2}{\varepsilon_1}} = 1. \tag{2}$$

[1] Lamé curves are named after the French mathematician Gabriel Lamé, who was the first who studied them in *Examen des différentes méthodes employées pour résoudre les problémes de geometrie,* Paris, 1818.

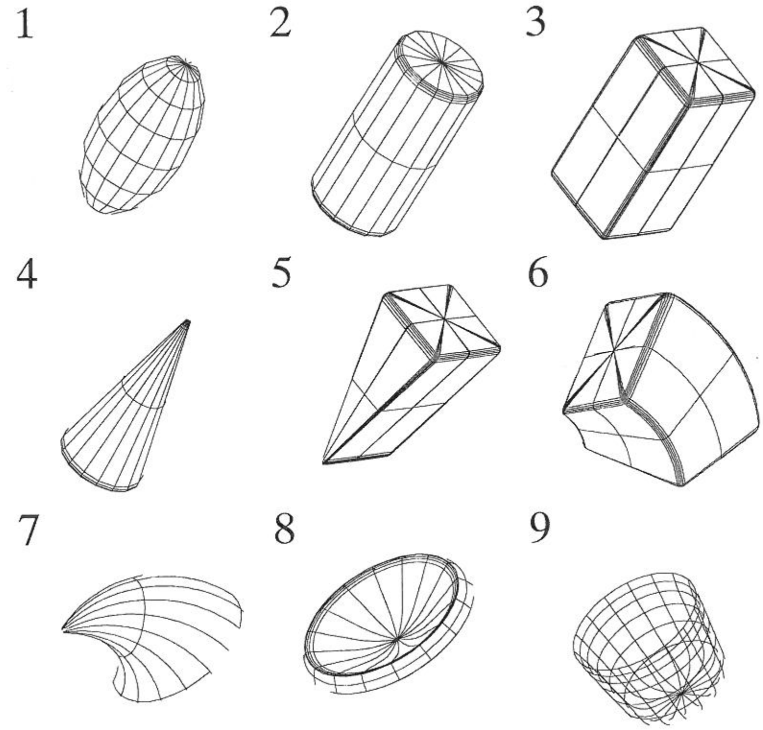

Fig. 1. Superquadric models enhanced with global deformations. (From [45].)

Points x, y, z that correspond to the above equation are on the surface of the superellipsoid.

For numerical calculation, it is easier to assume that exponents ε_1 and ε_2 can be any positive real number. Then, one should assume that exponentiation in equations 1 and 2 means

$$x^p = \text{sign}(x)|x|^p = \begin{cases} x^p & x \geq 0 \\ -|x|^p & x < 0 \end{cases}$$

to avoid complex numbers. For applications in computer vision, the values for ε_1 and ε_2 are normally bounded: $2 > \varepsilon_1, \varepsilon_2 > 0$, so that only convex shapes are produced. For a superquadric in canonical position one needs to set the value of 5 parameters (3 for size in each dimension, 2 for shape defining exponents). For a superquadric in general position 6 additional parameters are required to define the translation and rotation of the model. Parametric deformations typically require a few more parameters. For treatment of parametric deformations see [6, 45].

Barr saw the importance of superquadric models in particular for computer graphics and for three-dimensional design. Superquadric models, which compactly represent a continuum of useful forms with rounded edges, and which can easily be rendered and shaded due to their dual normal equations, and deformed by parametric deformations, are very useful in computer graphics. Parametric deformations [6] such as twisting, bending, tapering, and their combinations can be applied directly to superquadric surfaces. The surface vectors and the normal vectors of a deformed model are calculated by a simple matrix multiplication from the surface and normal vectors of the undeformed superquadric surface.

Pentland [35] was the first who grasped the potential of the superquadric models and parametric deformations for modeling natural shapes in the context of computer vision. He proposed to use superquadric models in combination with global deformations as a set of primitives which can be molded like lumps of clay to describe the scene structure at a scale that is similar to our naive perceptual notion of *parts*. Pentland presents several perceptual and cognitive arguments to recover the scene structure at such a part-level since people seem to make heavy use of this part structure in their perceptual interpretation of scenes. The superquadrics, which are like phonemes in this description language, are deformed by stretching, bending, tapering or twisting, and then combined, using Boolean operations to build complex objects.

2.1. Recovery of Superquadrics

The problem of recovering superquadrics from images is an overconstrained problem. A few model parameters (i.e. 11 for non-deformed superquadrics) must be determined from several (i.e. a few hundred) image features (range points, surface normals or points on occluding contours). By its parameterization the superquadrics impose a certain symmetry and in this way place some reasonable constraints on the shape of the occluded portion of a three dimensional object.

In the first article on the use of superquadrics in computer vision, Pentland [35] proposed an analytical method for recovery of superquadrics using the explicit equation (1). Except for some simple synthetic images, this analytical approach did not turn out to be feasible. Pentland [36] later proposed another method which combined recovery with segmentation and was based on a coarse search through the entire superquadric parameter space for a large number of overlapping image regions. The major objection to this method is its excessive computational cost.

Iterative methods based on non-linear least squares fitting techniques using different distance metrics were proposed [3, 9]. Solina and Bajcsy [3, 43, 45] formulated the recovery of deformed superquadric models from pre-segmented range data as a least-squares minimization of a fitting function. An iterative gradient descent method was used to solve the non-linear minimization problem. Initial estimates of the superquadric model are easy to compute. Center of gravity and central moments of inertia of the input points serve as estimates for position and orientation, while shape parameters ε_1 and ε_2 are set to 1, making the model an ellipsoid. A modified superquadric implicit or inside-outside function (Eq. 2) with an additional multiplicative volume factor was used as the fitting function. The volume factor is used to ensure the recovery of the *smallest* superquadric model that fits the range data in the least squares sense. To make the inside-outside function more suited for rapid convergence during minimization the inside-outside function (Eq. 2) was raised to the power of ε_1. Although this error metric varies across the surface when $\varepsilon_1 \neq \varepsilon_2$ and when size parameters change, it turned out to be efficient and robust (see Fig. 2). To the standard superquadric model, which requires 11 parameters, linear tapering, bending, and a cavity deformation were added, which adds up to a total of 18 parameters. Recovery of a single superquadric model requires on the average about 30 iterations (see Fig. 2).

Pentland [37] proposed another superquadric recovery method. Segmentation was first achieved by matching 2D silhouettes (2D projections of 3D superquadric parts of different shapes and of different orientations) to the image data. After part segmentation, superquadric models were fitted to range data of individual part regions. Superquadric fitting based on modal dynamics [38] used as the error metric the squared distance along the depth axis z between the range data and the projected volume's visible surface.

Hager [21] proposed a novel approach to sensor-based decision making that combines the estimation (recovery) process with the decision-making process (i.e. graspability, categorization). Usually both processes are divorced in the sense that first a recovery process is performed and then a decision is made based on the recovered models. Convergence on complex and non-linear problems are difficult to ensure on one hand and the amount spent on the fitting stage may be inappropriate for the

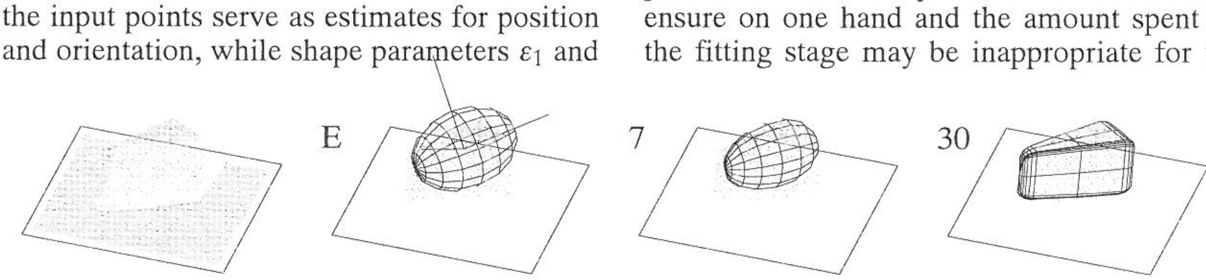

Fig. 2. Recovery of a tapered superquadric from pre-segmented range data. (From [45].) From left to right: original range image, E–initial estimate, models after the 7th and 30th iteration.

(a)	Actual	Init. est.	Computed
a_1	0.5	1	0.70
a_2	0.7	1	0.51
a_3	1	1	1.00
ε_1	1	1	0.98
ε_2	0.1	0.1	0.12
φ	0.9	0	−0.60
ϑ	0.4	0.01	0.37
ψ	0	0	−0.07

(b)	Actual	Init. est.	Computed
a_1	0.2	1	1.00
a_2	0.5	1	0.88
a_3	0.7	1	0.26
ε_1	0.3	0.1	0.70
ε_2	0.3	0.1	1.63
φ	1	0	-0.50
ϑ	1.9	1.58	-1.94
ψ	0	0	1.71

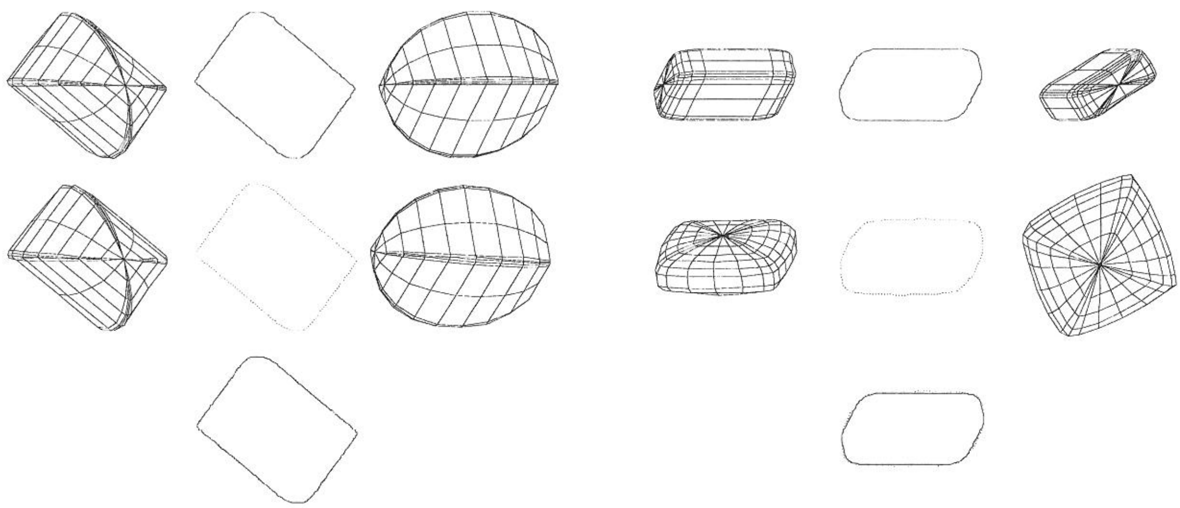

Fig. 3. Recovery of superquadrics from occluding contours (From [48]). The top row shows the actual models and the input contours. The second row shows the recoverd models. In the bottom row the difference between the input contour and the contour of the recovered model are displayed. In example (a), a very close match between the original and the recovered model was obtained. In example (b), the contour of the recovered model closely matches the input contour, however, the actual model is quite different.

decision sought on the other hand. Combining both stages should result in minimal work required to reach a decision. Any large system of parametric constraints under the assumption of bounded sensing error can be solved by this approach which is demonstrated in the article on two different problems, graspability and categorization, using range data and superquadric models. The approach is based on a interval-bisection method to incorporate sensor based decision making in the presence of parametric constraints. The constraints describe a model for sensor data (i.e. superquadric) and the criteria for correct decisions about the data (i.e. categorization—see also [44]). An incremental constraint solving technique performs the minimal model recovery required to reach a decision. The major drawback of the method is slow convergence when categorization is involved. Determining the shape parameters ε_1 and ε_2

Yokoya et al. [52] experimented with simulated annealing to minimize a new error-of-fit measure for recovery of superquadrics from pre-segmented range data. The measure is a linear combination of distance of range points to the superquadric surface and difference in surface normals (first proposed by Bajcsy and Solina in [3]). Several hundred iterations were needed to recover models from range data.

Vidmar and Solina [48] studied the recovery of superquadrics from 2D contours (see Fig. 3). For a given contour several possible superquadric interpretations are derived. To a human observer some of these interpretations are obviously more natural than others, although all recovered models have a very tight fit to the contour data. Perceptually better solutions could be selected by using just a few additional pieces of information (a few range points or shading information).

Horikoshi and Suzuki [24] multiplied the objective function with a weighting function for robust estimation (based on whether the point is closer to the median value of the inside-outside function, or far from it in either directions).

Consequently, the model is less sensitive to outliers.

3. Application of Superquadrics in Vision

This section explores the role of superquadrics as volumetric shape primitives for object classification and recognition, model evaluation, and segmentation. As is the case with any shape primitive, superquadrics have a limited shape vocabulary. They can be used to capture the global coarse shape of a 3D object or its constituent parts. The addition of global deformations increases the expressive power of superquadrics, but still limits it to the global coarse shape as opposed to local details. This lack of fine scale representation can be addressed by adding local degrees of freedom [38, 47]. However, one drawback of such locally deformable extensions is that they have too many degrees of freedom to meaningfully segment even a simple scene. The increase in expressive power also results in an increase in complexity of all the visual tasks like segmentation, representation, recognition, and classification. Consequently, all of the segmentation and classification work [37, 17, 14, 29] has used globally deformable models, limiting the role of local deformations to refine the surface details.

The earliest works on superquadrics dealt primarily with single model analysis, since segmentation of a complex scene required model recovery to be understood first. These methods focussed either on classification of single models, where the power of superquadrics as a compact parametric model was exploited [44, 23, 40], or on using superquadrics as a volumetric primitive *after* a segmentation had been obtained [36, 19, 13]. Once the model recovery was understood, more sophisticated techniques were designed to apply superquadrics to scene segmentation [17, 29, 24].

3.1. Model Classification and Recognition

Solina and Bajcsy [44] recovered objects in the postal domain and categorized them as flats, tubes, parcels, and irregular packages based on the shape and size parameters of the segmented recovered models. Gupta et. al. [19] extended Solina's approach to work on a cluttered scene by segmenting the range image using an independent edge-based scheme, and then recovering individual postal objects after reasoning about the physical supporting plane to constrain the 3D shape of the object.

Horikoshi and Kasahara [23] partitioned the superquadric parametric space between $0 < \varepsilon_1 \leq 2.0$ and $1.0 \leq \varepsilon_2 \leq 3.0$ to develop a shape indexing language. They mapped the representation space to verbal instructions like "rounder", "pinch", "flatten", etc., and developed a man-machine interface to construct object models. They also described an indexing scheme where complex objects were stored as superquadric models and indexed by model parameters.

Raja and Jain [40] conducted experiments on mapping superquadric shapes to 12 shape classes corresponding to a "collapsed" set of 36 different geons. Geons are qualitative descriptions of shapes, classified only on the basis of axis shape, cross-section shape, cross-section sweeping function, and cross-section symmetry [7]. These qualitative geometrical properties could prove to be very useful in indexing object databases. Raja and Jain used the five shape and deformation parameters of superquadrics for classification into 12 geon classes using binary tree and k-nearest-neighbor classifiers.

Model-driven recognition (with superquadric part-primitives) has not so far been exploited despite the compact representation. The reason is that it is very difficult to recover "canonical" representations of objects from real data. Instead of recovering canonical descriptions, most researchers have followed the data-driven bottom-up strategy of fitting superquadric models to the data. The recognition problem then reduces to matching superquadric parts from a library of a continuum of shapes to a cluttered scene.

3.2. Model Evaluation, Validation and Active Exploration

Using a parametric shape model for vision requires that the model evaluation be built into the segmentation and recognition systems. To this end, it is imperative to study the residuals of shape models by comparing them against the given data. Gupta et al. [18] describe qualitative and quantitative residuals for the objective evaluation of the recovered superquadric models. Whaite and Ferrie [51] have recently described a decision theoretic framework to evaluate the fitted models. They extended their earlier work [50] on uncertainty in model parameters to develop three lack-of-fit statistics. Their technique is embedded under the umbrella of "active exploration," which evaluates the recovered model as more data is collected. Gupta and Bajcsy [17] used global *and* local distribution of residuals to determine model fitness and segmentation options on static data.

Real situations demand that the parametric shape models must be recovered on partial and noisy single viewpoint data. Data could be missing due to other objects occluding the view, or due to the shadows in scanner geometry, or due to self-occlusion in single viewpoint data. Noise in 3D measurements is inevitable and most difficult to model. While the symmetry constraints of superquadrics are useful in predicting the missing information, the downside of parametric models is the lack of uniqueness in describing incomplete and noisy data within an acceptable error of tolerance. This fact is borne out in the experiments conducted by [50], where they derive an ellipsoid of confidence within which all the acceptable models lie. Model validation is a crucial step in scene segmentation.

3.3. Segmentation

A common underlying task of most recognition applications is building the scene description in terms of symbolic entities. A challenging problem in scene understanding is segmentation, where each piece of information must be mapped either to a shape primitive or discarded as noise. At the same time, there should be a minimum number of such primitives applied, so as to get as compact a description as possible. The absence of the domain knowledge further makes it more difficult, as ambiguities arise due to multiple representations, incomplete data, and the multiple degrees of freedom of superquadrics. There have been various approaches to scene segmentation when using superquadrics to model parts in images. The approaches to segmentation can be broadly classified into two categories:

1. *Segment-then-fit* schemes.
2. *Segment-and-fit* schemes.

3.3.1. Two-stage Segment-then-fit Methods

These methods decouple segmentation and model recovery. First, segmentation is performed, then superquadrics are fitted to resulting regions.

Pentland [37], for example, used matched filters to segment binarized image data into part regions. The best set of binary patterns that would completely describe a silhouette is selected. The 3D data corresponding to each of the selected patterns was then fitted with a deformable superquadric based on modal dynamics.

Gupta et al. [19] used an edge based region growing method to segment range images of compact objects in a pile. The regions were segmented at jump boundaries, and each recovered region was considered a superquadric object. Reasoning was done about the physical support of these regions, and several possible 3D interpretations were made based on various scenarios of the object's physical support. A superquadric model was fitted and classified corresponding to each recovered object.

Ferrie et al. [14] used differential geometric properties and projected space curves modeled as snakes for segmenting range data. An augmented Darboux frame is computed at each point by fitting a parabolic quadric surface, which is iteratively refined by a curvature consistency algorithm.

Another qualitative shape recovery method using geon theory was proposed by Metaxas and Dickinson [34] to recover superquadrics on intensity data. Their integrated method uses Dickinson et al. [12]'s geon-based segmentation scheme (into ten geon classes) to provide orientation constraint and edge segments of a part. This is then the input to the physically-based global superquadric model recovery scheme developed by Terzopoulos and Metaxas [47].

All two-stage approaches suffer from the problem that the results of segmentation might not correspond tightly to any superquadric model. Thus, model recovery on such part domain will be uncertain about the shape, size and orientation of the model. To describe adequately a scene with a particular shape language, one must use this language also for partitioning the scene. A method that can accommodate the domain to the orientation, shape, and size of the superquadric model must be used. The integration of model recovery with segmentation was proposed in [4]. Solina [42] had originally attempted part-segmentation by recursively splitting the domains and rejecting extraneous points during model recovery. It is, however, extremely difficult to constrain the single model recovery to take part-structure into account. Clearly, segmentation into part models must recover parts by hypothesizing parts and testing (evaluating) them.

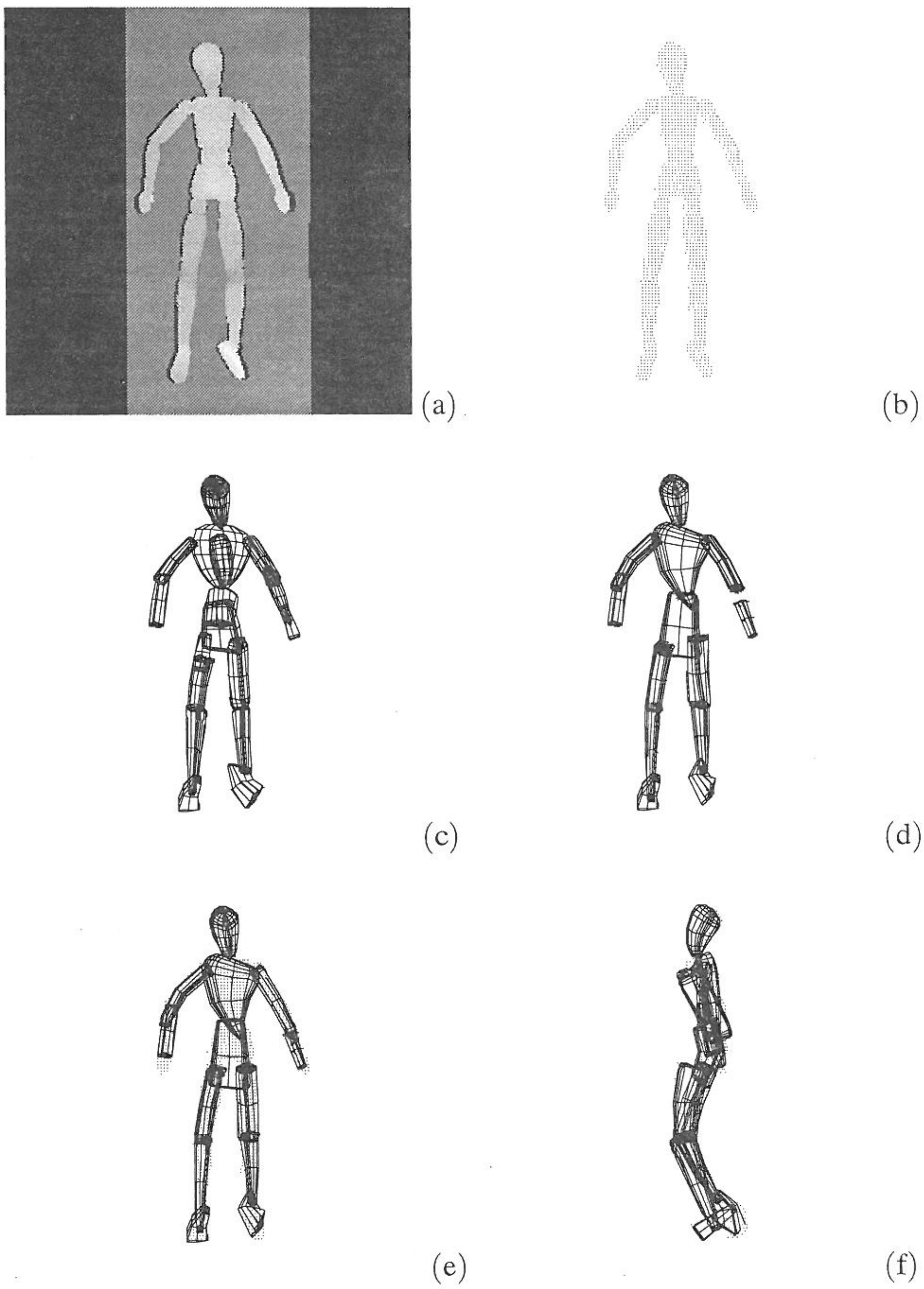

Fig. 4. Range image segmentation using deformed superquadric part models. The human form was segmented by employing the *recover-and-select* paradigm [27]. (a) – original range image, (b) – input range points, (c), (d) – models during the recovery process, (e), (f) – the final result from two views.

3.3.2. Interleaved Segment-and-fit Methods

Pentland [36] was the first one to integrate segmentation with superquadric model recovery. However, the brute-force method of searching the entire parameter space for a large number of overlapping image regions is computationally too expensive.

Gupta and Bajcsy [17], proposed a recursive global-to-local technique for part recovery. A global model is recovered for the entire data (using [45]), which is evaluated by studying local and global residuals so that the further course of action can be determined. A set of qualitative acceptance criteria define suitability of the model. If the model is found to be deficient in representing data, then additional part models are hypothesized on the un-described data (regions of surface underestimation). The global model is refitted on the remaining data. Thus, the global model shrinks while the part models grow, yielding a hierarchical part-structure.

A tighter integration of segmentation and model recovery was achieved by combining the recover-and-select paradigm developed by Leonardis [28, 27] with the superquadric recovery method by Solina [45]. This work demonstrates that superquadrics can be *directly* and *simultaneously* recovered from range data without the mediation of any other geometric models. The *recover-and-select* paradigm for the recovery of geometric parametric structures from image data [27] was originally developed for the recovery of parametric surfaces [28]. The paradigm works by recovering independently superquadric part models everywhere on the image, and selecting a subset which gives a compact description of the underlying data.

Horikoshi and Suzuki [24] proposed a segment-and-merge method to segment 2D contours (with the figure of interest separated from the background) and sparse 3D data. This recursive procedure results in a possibly overlapping convex superquadric parts. Parts are then merged to arrive at a compact description.

4. Applications of Superquadrics in Robotics

Superquadrics have been used for several different applications in robotics. Grasping can be used for manipulation by picking up objects as well as for tactile recognition. Choi et al. [11] studied the use of superquadrics for grasp planning. The robot gripper mounted on the Mars Rover developed at Carnegie-Mellon University has the task of picking up rock samples. A range image of the area in front of the Rover robotic arm is taken and areas corresponding to larger objects on the surface are isolated with a simple segmentation method. The range data of these regions serves as the input to the superquadric recovery method [45]. A robot gripper is then guided to pick up the objects.

Allen et al. [2] used superquadrics for obtaining initial global estimates of object's gross contour and volume as part of a larger system for 3D shape recovery and object recognition using touch and vision methods. For grasping by containment they used the Utah-MIT hand equipped with tactile sensors which provided a fair amount of sparse point contact data. These points were the input for the superquadric recovery algorithm [45].

Agba et al. [1] applied superquadrics to modeling kinematic chains. For control of kinematic chains such as robot manipulators and legged robots, simulation of interactions of kinematic links with objects in the environment is required. Superquadric models are used for modeling individual links of the manipulator. Using the inside-outside function, it is possible to determine distinctly whether an arbitrary point falls inside or outside of the volumes defined by the superquadrics. This modeling technique has been implemented in a hybrid simulator for undersea telerobotic manipulation for collision detection and grasp planning.

Khatib [26] and Volpe and Khosla [49] proposed superquadric potential functions for modeling obstacles and goal positions in robot work space. Manipulators and mobile robots must reach a desired destination without collision with obstacles. Artificial potential technique surrounds the obstacles with repulsive potential energy functions and places the goal in an attractive well. This approach enables real-time collision avoidance since the control of the manipulator can be done in the operational space—the space in which the task is described. Volpe and Khosla [49] applied superquadric potential functions in a newly developed robot controller in an experimental setting. In the future, visual feedback is planned that could provide obstacle data in real-time, enabling dynamic obstacle avoidance. Providing that the vision system also employs superquadric models, the integration of visual feedback into the robot control would not be difficult.

5. Conclusions

Despite initial reluctance in using superquadrics due to their nonlinear form, they have proven to be the primitives of choice for many a researcher seeking a volumetric model. Since no direct comparison of different metrics and minimization methods for superquadric recovery was made, it is difficult to rank them only on the basis of results presented in the articles. Some experimental comparisons of different error-of-fit measures are given in [16]. Gupta [18] also discusses the error-of-fit functions. What is finally important is the perceptual likeness of models to the actual objects, the speed of convergence, and last but not least, the simplicity of implementation. On this ground the method proposed in [45] received a wide acceptance since several other authors have used it in their vision or robotic systems [2, 18, 17, 11, 40, 14, 29].

Superquadric models have shown to be useful as volumetric shape primitives for object categorization, segmentation, recognition, and representation. Segment-and-fit methods report good results on objects which would otherwise be un-segmentable with surface-based techniques. Reliable segmentation of intensity and sparse 3D data is still an open problem. The discussed methods do not use domain knowledge directly, although it is possible to build in hooks to incorporate task-level constraints [21].

Important to notice are other potential application areas for parametric models. Image compression is normally not concerned with the actual structure of the depicted scene. However, a much higher compression rate could be achieved by (a) recovering models, (b) sending only the parameters of the models, (c) and then reconstructing the image from the models on the receiving side. Such intelligent compression is important if the scene must be understood also in terms of its physical structure (i.e. teleoperation).

Image understanding is becoming more important for general computer user interfaces since even personal computers are getting equipped with cameras and fast capabilities of image digitalization. Teleconferencing, face recognition, gesture recognition, user dependent interfaces, and virtual reality are just some of the exciting new application areas for computer vision.

References

[1] E. I. Agba, T.-L. Wong, and M. Z. Huang, Solid modeling of kinematic chains. In *Proceedings Fifth International Conference on Advanced Robotics*, pages 959–964, Pisa, Italy, 1991. IEEE.

[2] P. K. Allen and P. Michelman, Acquisition and interpretation of 3-D sensor data from touch. *IEEE Transactions on Robotics and Automation*, 6(4):397–404, 1990.

[3] R. Bajcsy and F. Solina, Three dimensional shape representation revisited. In *Proceedings of the First International Conference on Computer Vision*, pages 231–240, London, 1987. IEEE.

[4] R. Bajcsy, F. Solina, and A. Gupta, Segmentation versus object representation – are they separable? In R. Jain and A. Jain, ed., *Analysis and Interpretation of Range Images*. Springer-Verlag, New York, 1990.

[5] A. H. Barr, Superquadrics and angle-preserving transformations. *IEEE Computer Graphics and Applications*, 1(1):11–23, 1981.

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

[7] I. Biederman, Human image understanding: recent research and theory. *Computer Vision, Graphics, and Image Processing*, 32:29–73, 1985.

[8] T. O. Binford, Visual perception by computer. In *Proceedings of the IEEE Conference on Systems and Control*, Miami, FL, 1971.

[9] T. E. Boult and A. D. Gross, Recovery of superquadrics from depth information. In *Proceedings of Spatial Reasoing and Multi-Sensor fusion workshop*, pages 128–137, St. Charles, IL, 1987. SPIE.

[10] R. Brooks, Model-based 3-D interpretation of of 2-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):140–150, 1983.

[11] T. CHOI, H. DELINGETTE, M. DELUISE, Y. HSIN, M. HEBERT, AND K. IKEUCHI, A perception and manipulation system for collecting rock samples. In *Proceedings NASA Symposium on Space Operations, Applications and Research*, Albuquerque, NM, 1990.

[12] S. J. DICKINSON, A. P. PENTLAND, AND A. ROSENFELD, From volumes to views: An approach to 3-D object recognition. *CVGIP: Image Understanding*, 55(2):130–1154, 1992.

[13] F. P. FERRIE, J. LAGARDE, AND P. WHAITE, Darboux frames, snakes, and super-quadrics: Geometry from the bottom-up. In *Proc. IEEE Workshop on Interpretation of 3D Scenes*, pages 170–176, Austin, TX, 1989.

[14] F. P. FERRIE, J. LAGARDE, AND P. WHAITE, Darboux frames, snakes, and super-quadrics: Geometry from the bottom up. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):771–784, 1993.

[15] M. GARDNER, The superellipse: a curve that lies between the ellipse and the rectangle. *Scientific American*, 213(2):222–234, 1965.

[16] A. D. GROSS AND T. E. BOULT, Error of fit measures for recovering parametric solids. In *Proceedings of the 2nd International Conference on Computer Vision*, pages 690–694, Tampa, FL, 1988. IEEE.

[17] A. GUPTA AND R. BAJCSY, Volumetric segmentation of range images of 3-D objects using superquadrics. *CVGIP: Image Understanding*, 58(3):302–326, 1993.

[18] A. GUPTA, L. BOGONI, AND R. BAJCSY, Quantitative and qualitative measures for the evaluation of the superquadric models. In *Proc. IEEE Workshop on Interpretation of 3D Scenes*, pages 162–169, Austin, TX, 1989.

[19] A. GUPTA, G. FUNKA-LEA, AND K. WOHN, Segmentation, modeling and classification of the compact objects in a pile. In *Proceedings of the Conference on Intelligent Robots and Computer Vision VIII: Algorithms and Techniques*, pages 98–108, Philadelphia, PA, November 1989. SPIE.

[20] A. GUPTA AND C.-C. LIANG, 3-D model-data correspondence and nonrigid deformation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, New York City, June 1993.

[21] G. D. HAGER, Constraint solving methods and sensor-based decision making. In *Proceedings of the International Conference on Robotics and Automation*, pages 1662–1667, Nice, France, 1992. IEEE.

[22] A. J. HANSON, Hyperquadrics: Smoothly Deformable Shapes with Complex Polyhedral Bounds *Computer Vision, Graphics, and Image Processing*, 44:191-210", 1988.

[23] T. HORIKOSHI AND H. KASAHARA, 3-D shape indexing language. In *Proceedings of the International Conference on Computers and Communications*, Scottsdale, AZ, 1990. IEEE.

[24] T. HORIKOSHI AND S. SUZUKI, 3D parts decomposition from sparse range data using information criterion. In *Proceedings of the CVPR*, pages 168–173, 1993.

[25] B. K. P. HORN, *Robot Vision*, MIT Press, Cambridge, MA, 1986.

[26] O. KHATIB, Real-time obstacle avoidance for manipulators and mobile robots. *International Journal of Robotics Research*, 5(1):90–98, 1986.

[27] A. LEONARDIS, *Image analysis using parametric models: model-recovery and model-selection paradigm*. Ph.d. dissertation, University of Ljubljana, Faculty of Electr. Eng. and Computer Science, 1993.

[28] A. LEONARDIS, A. GUPTA, AND R. BAJCSY, Segmentation as the search for the best description of the image in terms of primitives. In *Proc. Third Int. Conf. on Computer Vision*, pages 121–125, Osaka, Japan, 1990. IEEE.

[29] A. LEONARDIS, F. SOLINA, AND A. MACERL, A direct recovery of superquadric models in range images using recover-and-select paradigm. In *Proc. Third European Conf. on Computer Vision*, pages 309–318, Stockholm, Sweden, 1994.

[30] G. LORIA, *Spezielle algebraische und transzendente ebene Kurven*. B. G. Teubner, Leipzig, Berlin, 1910.

[31] S. MURAKI, Volumetric shape description of range data using "Blobby model', *Computer graphics*, 25(4): 227–235, 1991.

[32] M. ZERROUGH AND R. NEVATIA, Segmentation and recovery of SHGCs from a real intensity image. In *Proceedings Third European Conference on Computer Vision, Volume I*, pages 319–330, Stockholm, Sweden, 1994.

[33] D. MARR AND K. NISHIHARA, Objects, parts, and categories. *Proceedings of the Royal Society, London*, B 200:269–294, 1978.

[34] D. METAXAS AND S. DICKINSON, Integration of quantitative and qualitative techniques for deformable model fitting from orthographic, perspective, and stereo projections. In *Proceedings of International Conference on Computer Vision*, pages 641–649. IEEE, 1993.

[35] A. P. PENTLAND, Perceptual organization and the representation of natural form. *Artificial Intelligence*, 28(2):293–331, 1986.

[36] A. P. PENTLAND, Recognition by parts. In *Proceedings of the First International Conference on Computer Vision*, pages 612–620, London, 1987. IEEE.

[37] A. P. PENTLAND, Automatic extraction of deformable part models. *International Journal of Computer Vision*, 4:107–126, 1990.

[38] A. P. PENTLAND AND S. SCLAROFF, Closed-form solutions for physically based shape modeling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):715–729, 1991.

[39] J. PONCE, D. CHELBERG, AND W. B. MANN, Invariant properties of straight homogeneous generalized cylinders and their contours. *IEEE Transactions and Pattern Analysis and Machine Intelligence*, 11(9):951–966, 1989.

[40] N. S. RAJA AND A. K. JAIN, Recognizing geons from superquadrics fitted to range data. *Image and Vision Computing*, 10(3):179–190, 1992.

[41] R. B. SCHUDY AND D. H. BALLARD", Towards an anatomical model of hearth motion as seen in 4-D cardiac ultrasound data. *Proceedings 6th Conference on Computer Applications in Radiology and Computer-Aided Analysis of Radiological Images*, 1979.

[42] F. SOLINA, *Shape recovery and segmentation with deformable part models.* Ph.d. dissertation, University of Pennsylvania, Department of Computer and Information Science, 1987.

[43] F. SOLINA AND R. BAJCSY, Range image interpretation of mail pieces with superquadrics. In *Proceedings AAAI-87*, volume 2, pages 733–737, Seattle, WA, 1987.

[44] F. SOLINA AND R. BAJCSY, Recovery of mail piece shape from range images using 3-D deformable models. *International Journal of Research & Engineering, Postal Applications*, Inaugural Issue:125–131, 1989.

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

[46] D. TERZOPOULOS, A. WITKIN, AND M. KASS, Symmetry-seeking models for 3D object reconstruction, *International Journal of Computer Vision*, 1(3), 1988.

[47] D. TERZOPOULOS AND D. METAXAS, Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714, 1991.

[48] A. VIDMAR AND F. SOLINA, Recovery of superquadric models from occluding contours. In R. Klette and W. Kropatsch, editors, *Theoretical Foundations of Computer Vision*, Mathematical Research, Vol. 69, pages 227–240. Akademie Verlag, Berlin, 1992.

[49] R. VOLPE AND P. KHOSLA, Manipulator control with superquadric artificial potential functions: Theory and experiments. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(6):1423–1436, 1990.

[50] P. WHAITE AND F. P. FERRIE, From uncertainty to visual exploration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10):1038–1049, October 1991.

[51] P. WHAITE AND F. P. FERRIE, Active exploration: Knowing when we're wrong. In *Proceedings of the ICCV*, pages 41–48, Berlin, Germany, 1993. IEEE CS.

[52] N. YOKOYA, M. KANETA, AND K. YAMAMOTO, Recovery of superquadric primitives from a range image using simulated annealing. In *Proc. 11th Int. Conf. on Pattern Recognition*, vol. A, pages 168–172. IAPR, IEEE Comp. Soc. Press, 1992.

Contact address:
Franc Solina
Computer Vision Laboratory
Faculty of Electrical Engineering and Computer Science,
University of Ljubljana,
Tržaška 25, 61001 Ljubljana, Slovenia
Phone: +386 (61) 1768 381,
Fax +386 (61) 264 990
E-mail: Franc.Solina@fer.uni-lj.si

FRANC SOLINA is an associate professor of computer science and head of Computer vision laboratory at Faculty of electrical engineering and computer science, University of Ljubljana. He received his dipl.ing. degree (1979) and M.S. degree (1982) in electrical engineering from University of Ljubljana, Slovenia. His Ph.D. degree in computer science (1987) is from University of Pennsylvania, PA. His research interests are modelling, segmentation and object recognition in computer vision.