A Direct Part-Level Segmentation of Range Images Using Volumetric Models*

Franc Solina, Aleš Leonardis, Alenka Macerl†
Computer Vision Laboratory
Faculty of Electrical Engineering and Computer Science
University of Ljubljana
Tržaška c. 25, 61001 Ljubljana, Slovenia
Franc.Solina@fer.uni-lj.si, Ales.Leonardis@fer.uni-lj.si

Abstract

Volumetric part models play an important part in robotic applications such as grasping, path planning, object avoidance, and modeling kinematic chains.

We present a novel method to reliable and efficient recovery of part-descriptions in terms of superquadric models from range images. In contrast to usual approaches which perform the recovery of volumetric models in several steps (from curves, surfaces to volumes), we show that a direct recovery is possible. This is achieved by combining two existing methods: recover-and-select *paradigm and recovery of superquadric models. A redundant set of superquadrics is initiated in the image and only the recovered models resulting in the simplest overall description are selected. We show the results on several real range images.*

1 Introduction

Superquadrics have been used as part models for several different robotic applications such as grasping, modeling kinematic chains, and modeling the robot workspace.

Allen *et al.* [2] used superquadrics for obtaining initial global estimates of object's gross shape as part of a larger system for 3-D shape recovery and object recognition using touch and vision. For grasping by containment the Utah-MIT hand equipped with tactile sensors was used. The contact points were the input for the superquadric recovery algorithm [13]. In a similar way, Choi *et al.* [7] studied the use of superquadrics for picking up rock samples with a robotic arm mounted on a Mars Rover vehicle. What makes superquadrics particularly relevant for haptic recognition is that the models are volumetric in nature, which maps directly into the psychophysical perception processes suggested by grasping by containment.

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 which are related by homogeneous transformations. Using the implicit 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.

Volpe and Khosla [15] 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. Superquadric potential functions were applied 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 robotic vision system also employs superquadric models, the integration of visual feedback into the robot control would not be difficult. Using the same type of shape models in all parts of a robotic system would clearly benefit a fast and reliable data fusion.

*This work was supported by the Ministry for Science and Technology of the Republic of Slovenia (Project P2-1122).

†Alenka Macerl is now with Hermes-Soft d.o.o, Ljubljana, Slovenia.

1050-4729/94 $03.00 © 1994 IEEE

1.1 Motivation

The significance of detecting and representing part-descriptions in terms of superquadric models has long been realized in vision and robotics community. The major obstacle to successful applications is the difficulty of reliably extracting superquadric models from the data.

Previous approaches attempt to solve this problem by first pre-segmenting the data (range images in most cases) based on various criteria and then fit superquadrics into chunks of data [6]. These approaches, in fact, isolate segmentation stage from the representation stage, and significant efforts are necessary to combine, usually surface type descriptions into volumetric models. We argued that segmentation and shape recovery have to be performed simultaneously [3]. In this paper we show that a *direct* recovery of superquadric models is possible.

One of the authors of the paper developed a *recover-and-select* paradigm for recovery of geometric parametric structures from data [8] and demonstrated its applicability on a variety of domains: surface models in range data [9] and curve-models in intensity images [10]. The encouraging results motivated us to attempt the recovery of superquadrics in the *recover-and-select* paradigm to achieve *direct* segmentation into part-level models. For the recovery of superquadrics we use the method developed by another author of this paper [13].

The paper is organized as follows. In section 2 we outline the *recover-and-select* paradigm. In section 3 we talk about the recovery of superquadrics. Experimental results are shown in section 4, followed by conclusions and work in progress.

2 Recover and Select

Due to space limitations we present here only a general outline of the recover-and-select paradigm. The choice of specific models, in our case superquadrics, imposes certain constraints on the recovery of these structures from images, which will be described in section 3.

2.1 Model Recovery

Recovery of parametric models is a difficult problem because we have to find image elements that belong to a single parametric model ***and*** we have to determine the values of the parameters of the model. For image elements that have already been classified we can determine the parameters of a model using standard statistical estimation techniques. Conversely, knowing the parameters of the model, a search for compatible image points can be accomplished by pattern classification methods. Thus we propose to solve this problem by an iterative method, conceptually similar to the one described by Besl [5], which combines data classification and model fitting.

One of the crucial problems is where to find the initial estimates (seeds) in an image since their selection has a major effect on the success or failure of the overall procedure. We propose that a search for the points that could belong to a single parametric model is performed in a grid-like pattern of windows overlaid on the image. Thus, the requirement of classifying all data points of a certain model is relaxed to finding only a small subset. However, there is no guarantee that every seed will lead to a good description since some initial models can be constructed over areas which belong to different models. As a remedy we propose to *independently* build *all possible* models using all statistically consistent seeds and to use them as hypotheses that could compose the final description.

Having an initial set of points (a seed) we estimate the parameters of the model. If sufficient similarity between the model and the data is established, ultimately depending on the task at hand, we proceed with a search for more compatible points. An efficient search is performed in the vicinity of the present border points of the model. New image elements are included in the data set and the parameters of the model are updated. The new goodness-of-fit is computed and compared to the old value. This is followed by a decision whether to perform another iteration or terminate the procedure.

The main features of model-recovery procedure are: a high degree of resistance to outliers since the performance of the fitting is constantly monitored and an independent and (possibly) parallel execution of the recovery procedure for individual models.

The final outcome of the model-recovery procedure for a particular model M_i consists of: a set R_i of data elements that belong to the model M_i (n_i is the cardinality of this set), a set of parameters of the model (N_i denotes the number of parameters), and the error value ξ_i which describes the conformity between the data and the model. This output is subsequently passed to the model-selection procedure.

2.2 Model Selection

The redundant representation obtained by the model-recovery procedure is a direct consequence of the decision that a search for parametric volumetric models is initiated everywhere in an image. Several of the models are completely or partially overlapped. The task of combining different models is posed as a selection procedure which selects those models that produce the simplest description, i.e., the one that describes the data with the minimum number of models while keeping the deviations between data points and models low.

The objective function $F(\mathbf{m})$, which is to be maximized in order to produce the "simplest" description in terms of models, has the following form:

$$F(\mathbf{m}) = \mathbf{m}^T \mathbf{Q} \mathbf{m} = \mathbf{m}^T \begin{bmatrix} c_{11} & \dots & c_{1M} \\ \vdots & & \vdots \\ c_{M1} & \dots & c_{MM} \end{bmatrix} \mathbf{m} \ , \tag{1}$$

where $\mathbf{m}^T = [m_1, m_2, \dots, m_M]$. m_i is a *presence variable* having the value 1 for the presence and 0 for the absence of the model M_i in the final description. The diagonal terms express the cost-benefit value for a particular model M_i

$$c_{ii} = \mathrm{K}_1 n_i - \mathrm{K}_2 \xi_i - \mathrm{K}_3 N_i \ , \tag{2}$$

while the off-diagonal terms handle the interaction between the overlapping models

$$c_{ij} = \frac{-\mathrm{K}_1 \Gamma(M_i, M_j) + \mathrm{K}_2 \xi_{i,j}}{2} \ . \tag{3}$$

$\Gamma(M_i, M_j) = |R_i \cap R_j|$ is the number of points that are explained by both models, and $\xi_{i,j}$ is defined as

$$\xi_{i,j} = \max(\sum_{R_i \cap R_j} d^2_{M_i}, \sum_{R_i \cap R_j} d^2_{M_j}) \ . \tag{4}$$

The error terms $d^2_{M_i}$ and $d^2_{M_j}$ are calculated in the region of intersection and correspond to deviations from the i-th and j-th model, respectively. K_1, K_2, and K_3 are weights which can be determined on a purely information-theoretical basis (in terms of bits), or they can be adjusted in order to take into account the signal-to-noise ratio.

Maximizing the objective function $F(\mathbf{m})$ belongs to the class of problems known as combinatorial optimization (Quadratic Boolean problem). Since the number of possible solutions increases exponentially with the size of the problem, it is usually not tractable to explore them exhaustively. The exact solution has to be sacrificed to obtain a realizable solution.

Various methods have been proposed for finding a "global extreme" of this class of non-linear objective functions. However, they are in general too time consuming to be applicable in most situations. It turns out that in our case a reasonable solution can be obtained by a direct application of the ***greedy algorithm*** which at any individual stage selects the option which is "locally optimal". In other words, the models are selected in the sequence that corresponds to the size of their contributions to the objective function, which is equivalent to applying at each stage of the algorithm the *winner takes all* principle.

2.3 Model Recovery and Model Selection

In order to achieve a computationally efficient procedure we combine the model-recovery and model-selection procedures in an iterative fashion. The recovery of currently active models is interrupted by the model-selection procedure which selects a set of currently optimal models which are then passed back to the model-recovery procedure. This process is repeated until the remaining models are completely recovered.

3 Superquadric Models

Superquadrics are an extension of basic quadric surfaces and solids. Superquadrics have been considered as volumetric primitives for shape representation in computer graphics [4], computer vision [12,13,14,11], and robotics [2,7,1,15]. Superquadrics are convenient part-level models that can be further deformed and glued together to model articulated objects.

Superquadric surface is defined by the following equation

$$F(x,y,z) = \left(\left(\left(\frac{x}{a_1}\right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_2}\right)^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}} + \left(\frac{z}{a_3}\right)^{\frac{2}{\epsilon_1}} \right)^{\epsilon_1} = 1. \tag{5}$$

When both exponents ϵ_1 and ϵ_2 equal 1, the surface forms an ellipsoid. When $\epsilon_1 \ll 1$ and $\epsilon_2 = 1$, the superquadric surface is shaped like a cylinder. Parallelepipeds are produced when both $\epsilon_1 \ll 1$ and $\epsilon_2 \ll 1$. Modeling capabilities of superquadrics can be enhanced by deforming them in different ways using either global deformations (i.e., tapering and bending [13]) or local deformations for detailed modeling [14]. We devised a robust method for recovery of isolated superquadric models which is based on a nonlinear least-squares method [13]. In the next two subsection we will first summarize the recovery method of isolated superquadric models and then describe its modifications necessary for the inclusion of superquadrics into the recover-and-select paradigm.

3.1 Recovery of Superquadrics

The implicit function (Eq. 5), defined in an object centered coordinate system, determines where a given point (x, y, z) lies relative to the superquadric surface. To recover a superquadric in general position, the implicit function for general position is used

$$F(x,y,z) = F(a_1, a_2, a_3, \epsilon_1, \epsilon_2, \phi, \theta, \psi, p_x, p_y, p_z). \quad (6)$$

This expanded "inside-outside" function has 11 parameters; a_1, a_2, a_3 define the superquadric size; ϵ_1 and ϵ_2 are shape parameters; ϕ, θ, ψ define the orientation in space, and p_x, p_y, p_z define the position in space. We refer to the set of all model parameters as $\Lambda = \{a_1, a_2, \dots, a_{11}\}$.

Suppose we have N 3-D points on a surface of an object (x_W, y_W, z_W) which we want to model with a superquadric. We want to vary the 11 parameters $a_j, j = 1, \dots, 11$ in equation (6) to get such values for a_j's that most of the 3-D points will lie on, or close to the model's surface. There will probably not exist a set of parameters Λ that perfectly fits the data. Finding the model Λ for which the distance from points to the model's surface is minimal is a least-squares minimization problem. Since due to self occlusion, not all sides of an object are visible at the same time, we introduced an additional constraint. Among all possible solutions we search for the *smallest* superquadric that fits the given range points in the least squares sense. We defined the following function which has a minimum corresponding to the smallest superquadric that fits the set of 3-D points

$$f = \sqrt{a_1 a_2 a_3}\,(F - 1). \quad (7)$$

Since for a point (x_W, y_W, z_W) on the surface of a superquadric

$$f\,(x_W, y_W, z_W; a_1, \dots, a_{11}) = 0, \quad (8)$$

we have to find

$$G = min \sum_{i=1}^{N} [f\,(x_{W_i}, y_{W_i}, z_{W_i}; a_1, \dots, a_{11})]^2. \quad (9)$$

Because f is a nonlinear function of 11 parameters $a_j, j = 1, \dots, 11$, minimization must proceed iteratively. Given a trial set of values of model parameters Λ_k, we evaluate equation (7) for all N points and employ a procedure to improve the trial solution. The procedure is then repeated with a set of new trial values Λ_{k+1} until the sum of least squares (9) stops decreasing, or the changes are statistically insignificant. In most cases 15 iterations are more than sufficient. We use the Levenberg-Marquardt method for nonlinear least squares minimization since first derivatives $\delta G/\delta a_i$ for $i = 1, \dots, 11$ can be computed analytically.

Only very rough initial estimates of object's true position, orientation, and size suffice to assure convergence to a local minimum that corresponds to the actual shape. Initial values for both shape parameters, ϵ_1 and ϵ_2 are 1, which means that the initial model Λ_E is always an ellipsoid. Position in the world coordinates is estimated by computing the center of gravity of all range points, and the orientation is estimated by computing the central moments with respect to the center of gravity. To assure that the minimization procedure does not get stuck in a shallow local minimum we add noise during the minimization procedure.

3.2 Superquadrics in the Recover-and-Select Paradigm

To include the superquadric recovery method into the recover-and-select paradigm we made some adjustments. The number of points modeled by individual superquadric models is changing since the models in the recover-and-select paradigm must grow. The growing of models is in each step achieved first by simply increasing the values of the parameters a_1, a_2, a_3 of the current model to get an enlarged model. Then all the points which are inside this enlarged model are checked on how close they are to the surface of the original model. Only those points close enough to the original model are included into the enlarged set of points. On this larger set of points a new superquadric model recovery procedure is started.

Since the fitting function f depends on the size of the models and is as such not a good measure of the true distance, we used an approximation for computing the Euclidean distance of points to the models. When distance d is small,

$$d^2((x,y,z), \Lambda) \approx \frac{f^2(x,y,z)}{\|\nabla f((x,y,z))\|^2}. \quad (10)$$

4 Results

The proposed method has been tested on a variety of range images. All the examples were run on an

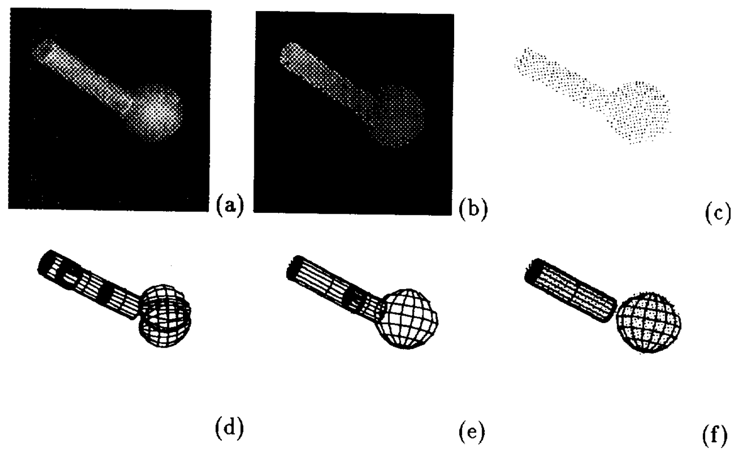

Figure 1: A sphere and a cylinder.

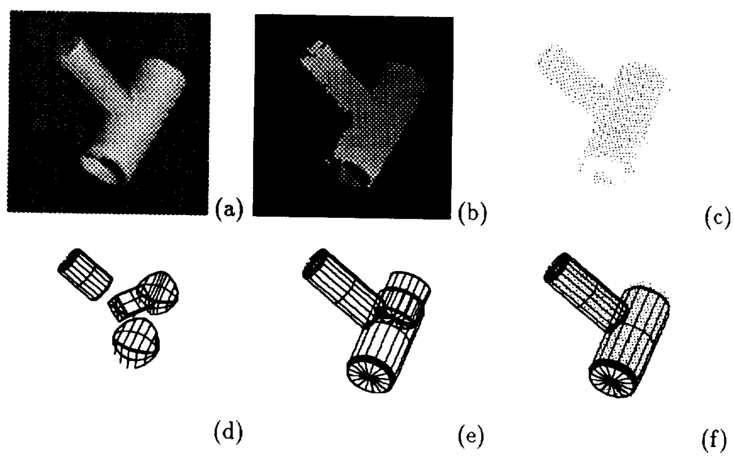

Figure 2: Two intersecting cylinders.

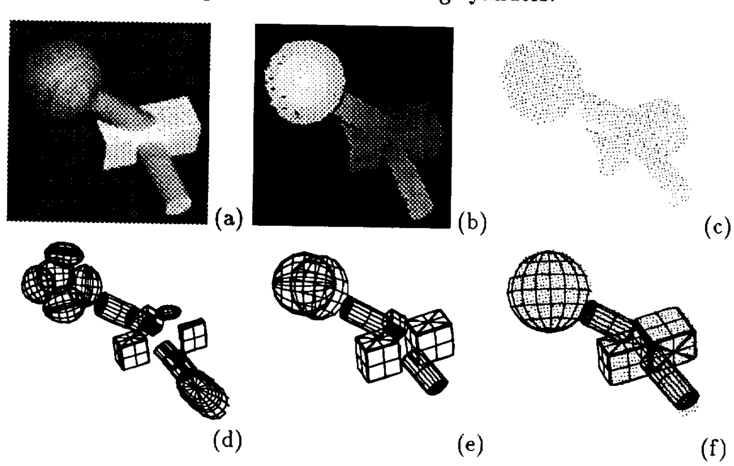

Figure 3: A complex scene (Problem: the block is represented by two superquadrics).

HP-715. To simplify the first experiments, only superquadrics without deformations were used. Because the original range images are very dense, they were subsampled to speed up the recovery. Since the computation of the fitting function and its derivatives is independent for each point the procedure can be parallelized in a straightforward way. Typical processing times for examples shown were 10 minutes CPU time. The thresholds in the model-recovery procedure were determined on the basis of the analysis of the noise properties of the acquisition process.

Three examples of processed range images are shown[1]. Each figure shows the following image sequence: (a) The intensity image of the object. (b) The corresponding range image. (c) The range image resampled and transformed into the form (x,y,z), appropriate for the superquadrics recovery process. (d) The recovered volumetric models after the first model selection. (e) Volumetric models at the midpoint of the recover-and-select process. (f) Final result.

5 Conclusions

Superquadrics are finding a role in robotics as part of dedicated vision systems, or for modeling the shape of robot kinematic chains, and the robot workspace (path planning, collision detection). Using the same type of geometric models simplifies fusion of visually derived information with the information of the robot control system. Unfortunately, reliable recovery of superquadric models from unsegmented data is still difficult. In this paper we have successfully combined two existing methods, namely recovery of superquadric models [13] and the *recover-and-select* paradigm [8]. Instead of a standard segmentation approach utilizing a series of models of different granularity (lines, surfaces, volumes), which is done without direct support of the final description, we showed that it is possible to directly employ the final volumetric part-level models. The interpretation of these models is straightforward, with direct application for manipulation, object recognition, and CAD modeling (reverse engineering).

There are several open issues in this approach: the number of starting seeds, error measure, robustness, joining models of same parts which are separated by occlusion (see Fig. 3). We also plan to explore the possibility to use the proposed method for fusing multiple geometric modalities (surfaces and volumes) and different image information (range and intensity images).

[1]The range images were kindly provided by Marjan Trobina from ETH, Zürich, Switzerland.

References

[1] E. I. Agba, T. Wong, and M. Z. Huang. Solid modeling of kinematic chains. *Fifth Int. Conf. on Advanced Robotics, Robots in Unstructured Environments*, pp. 959–963, Pisa, Italy, 1991. IEEE.

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

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

[4] A. H. Barr. Superquadrics and angle-preserving transformations. *IEEE Comp. Graph. and Appl.*, 1(1):11–23, 1981.

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

[6] A. Gupta and R. Bajcsy. Volumetric segmentation of Range Images of 3-D objects using superquadrics. *CVGIP: Image Understanding* 58(3):302-326, 1993.

[7] T. Choi, H. Delingette, M. DeLuise, Y. Hsin, M. Hebert, and K. Ikeuchi. A perception and manipulation system for collecting rock samples. In *Proc. NASA Symp. on Space Operations, Appl. and Research*, Albuquerque, NM, 1990.

[8] A. Leonardis. *Image Analysis Using Parametric Models: Model-Recovery and Model-Selection Paradigm.* A Dissertation in Computer Science. University of Ljubljana, Faculty of EE & CS. May 1993.

[9] A. Leonardis, A. Gupta, and R. Bajcsy. Segmentation as the search for the best description of the image in terms of primitives. In *Proc. of the Third Int. Conf. on Computer Vision*, pp. 121–125, Osaka, Japan, 1990. IEEE.

[10] A. Leonardis and R. Bajcsy. Finding parametric curves in an image. In G. Sandini, editor, *Proc. of The 2nd European Conf. on Computer Vision—ECCV-92.* Springer-Verlag, 1992. LNCS-Series Vol. 588.

[11] D. Metaxas and D. Terzopoulos. Shape and nonrigid motion estimation through physics-based synthesis. *IEEE Trans. on PAMI*, 15(6):580–591, 1993

[12] A. P. Pentland. Perceptual organization and the representation of natural form. *Artif. Intell.*, 28(2):293–331, 1986.

[13] F. Solina and R. Bajcsy. Recovery of parametric models from range images: The case for superquadrics with global deformations. *IEEE Trans. on PAMI*, 12(2):131–147, 1990.

[14] D. Terzopoulos and D. Metaxas. Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Trans. on PAMI*, 13(7):703–714, 1991.

[15] R. Volpe and P. Khosla. Manipulator control with superquadric artificial potential functions: Theory and experiments. *IEEE Trans. on Systems, Man, and Cybernetics*, 20(6):1423–1436, 1990.