SHAPE DESCRIPTION

Recovery of Volumetric Object Descriptions From Laser Rangefinder Images

F.P. Ferrie J. Lagarde P. Whaite

Abstract

This paper describes a representation and computational model for deriving three dimensional, articulated volumetric descriptions of objects from laser rangefinder data. What differentiates this work from other approaches is that it is purely bottom-up, relying on general assumptions cast in terms of differential geometry.

1 Introduction

The ability of a robot to correctly perceive its environment is essential to tasks involving interaction and navigation. Descriptions computed by the perceptual system must reflect the characteristics of the world; that objects take up space, are often composed of many parts, and can be articulated in a number of different ways. This paper is about computing such descriptions from the bottom up. That is, beginning with estimates of surface points obtained with a laser rangefinder, we will describe how an articulated volumetric description of an object can be obtained through a succession of intermediate representations and computational steps. Some of the building blocks that we shall use are well-known, but it is the way in which these are tied together and the computational aspects of the problem that are the principal contributions of this paper.

Our approach follows a traditional bottom-up transition from surfaces to parts to objects in a hierarchical fashion [11, 12, 19]. Darboux frames, which describe the orientation, principal curvatures and directions at a point on a surface [7, 22, 23] are used as a local representation for a surface at each sample point. Initial estimates of these measures are readily computable using a number of different approaches [3, 8, 12]. But the technical difficulty is to refine the initial estimates, which are often corrupted by noise and quantization error, into robust measures. We have adopted a computational strategy based on the minimization of a residual form that measures the total deviation from an implicit model of surface curvature [22, 23, 9, 10]. This approach allows us to reconstruct a surface from sensor data and, more importantly, to make the features that are needed for the parts decomposition explicit (e.g. negative local minima of curvature, orientation and jump discontinuities) [11, 12, 15].

Contours formed by these features can serve to partition a surface into regions corresponding to different parts. However this grouping problem is often difficult, especially when features are sparse or when adjacent contours are in close proximity. An elegant solution can be obtained by exploiting the directional properties of the frame[1] associated with each feature point. Following the scheme devised by Zucker et al. [26], the direction vectors of each frame are used to generate a potential field which acts on a covering of unit length snakes (energy minimizing splines) [16, 24, 25]. When this system reaches a steady-state, contours are obtained that smoothly interpolate the data,

[1]For brevity we will use frame to refer to Darboux frame or augmented Darboux frame, depending on the context.

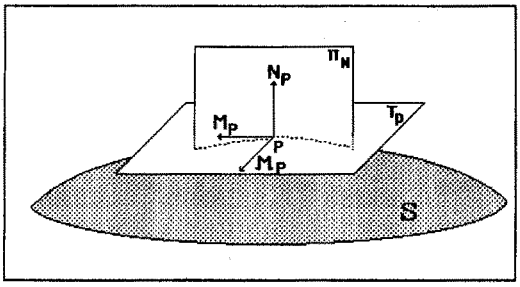

Figure 1: Local surface representation – the augmented Darboux Frame.

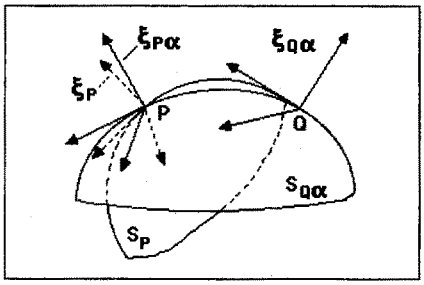

Figure 2: Local extrapolation using a parabolic quadric approximation.

and correspond to the requisite part boundaries. What is important about this scheme is that the parameters of the minimization are determined primarily by the data [26].

Each surface partition is taken to be the visible part of a single object. Part geometry is determined by a process of inference in which a suitable volumetric primitive is fit to each surface patch obtained from the parts decomposition [12, 10]. Our approach is in contrast to the minimal length encoding strategy used by Pentland [19]. He begins with a with a fixed parametric model (a superquadric) and uses a minimization that seeks a best fit to data using a minimal set of primitives. We view Pentland's approach as a top down strategy and ours bottom up. The advantages offered by the latter scheme are a reduction in computational complexity, and a representation that is not tied to any specific model. Depending on what needs to be made explicit, any volumetric model can be used to characterize a part without changing the interpretation of the object.

The paper is structured along the computational steps that define our procedure for building volumetric object models from sensor data. Section 2 describes the local representation for surfaces and the minimization procedure used to compute reliable estimates of its parameters. The task of identifying the feature points that make up the part boundaries and the interpolation procedure for aggregating these into contours using the potential field method are described in Section 3. To obtain comparative results with Pentland's approach, we used superquadric primitives to represent part geometry. The technique used to fit the primitives is similar to others [2, 19, 1, 6, 14], and is described in Section 4. Finally, the performance or our scheme on real data is shown in Section 5. Articulated models of two objects are derived from range maps acquired with a laser rangefinder.

2 Local Representation of a Surface

The local surface representation at a point P, the augmented Darboux frame $\mathcal{D}(P)$ [7, 22, 23] serves three purposes: (1) It facilitates the task of surface reconstruction, (2) makes explicit the features necessary for identifying putative boundary points, and (3) provides the direction frames used to smoothly interpolate contours on the surface. It is explained as follows. Let the local neighbourhood of a point P on a surface S be represented by a parabolic quadric of the form $w = au^2+cv^2$, with origin at P and the w axis aligned with the surface normal at P, Np, as shown in Figure 1. The orientation of this local frame is such that the u and v coordinate axes align with two special directions on S at P. These are the directions for which the normal curvature at P (a directional property) takes on maximum and minimum values, κ_{MP} and $\kappa_{\mathcal{M}P}$, and are referred to as the principal directions M_P and $\mathcal{M}_P$ respectively [7]. The scalar quantities κ_{MP} and $\kappa_{\mathcal{M}P}$ are similarly referred to as the principal curvatures at P. Following the convention of [22, 23], we refer to $\mathcal{D}(P) = (p, M_P, \mathcal{M}_P, N_P, \kappa_{MP}, \kappa_{\mathcal{M}P})$ collectively as the augmented Darboux Frame at P. The problem is to estimate $\mathcal{D}(P)$ from laser rangefinder images of the form $z = f(x, y)$. Local least-squares estimation is sometimes sufficient to determine κ_{MP}, $\kappa_{\mathcal{M}P}$, and N_P, but rarely M_P and $\mathcal{M}_P$ [3, 12]. However, the latter directional properties are essential to the inference of discontinuities, occluding contours, and part boundaries

[26]. Our approach is to use local methods to obtain a first estimate of $\mathcal{D}(P)$, and then apply a second stage of minimization to obtain a stable reconstruction of S. The former problem is not addressed in this paper, but typical approaches are described in [3, 8, 12].

2.1 Iterative Refinement of the Darboux Frame

Local consistency of curvature, subject to orthogonality constraints on $\mathcal{D}(P)$, is the basis of our minimization algorithm [18, 23]. The method was first introduced by Sander and Zucker in the context of C.T. image reconstruction [22, 23]. Aside from the application to range data, there are a number of important technical details which differentiate our work from that of Sander and Zucker. However, the motivation is similar and can be explained with the aid of Figure 2. Because $\mathcal{D}(P)$ is a dual form of a parabolic quadric, one can extrapolate outward from a point Q to its neighbour P to get an idea of what the surface at P looks like according to its neighbour Q. By performing this operation for each neighbour of P, one obtains a set of frames, $\xi_{P\alpha}$, each providing an estimate of P from its associated neighbour [22, 23]. This provides a mechanism for setting up a minimization which enforces the local model over the surface, somewhat analogous to the constant curvature assumption in [18]. What the algorithm does in practice is to iteratively update each $\mathcal{D}(P)$ with the least-squares estimate computed from $\xi_{P\alpha}$.

Among the considerations in formulating the algorithm are, (i) the particular form of extrapolation along the surface (related to parallel transport [7, 22, 23]) to obtain $\xi_{P\alpha}$ and (ii) the form of minimization functional applied in updating $\mathcal{D}(P)$. For the results presented in this paper, we followed [22, 23] and used a parabolic quadric to estimate $\xi_{P\alpha}$ from the surrounding neighbourhood[2]. But this does not enforce the constant curvature constraint according to [17]. A better choice is a toroidal patch that has constant curvature along its principal directions. However, with densely sampled range data, the particular form of parallel transport does not appear to be critical.

The functional itself is set up to minimize the variation in $\mathcal{D}(P)$ subject to constraints on M_P, $\mathcal{M}_P$, and N_P. These are,

$$(N_P \cdot N_P) = 1 \quad (M_P \cdot M_P) = 1 \quad (M_P \cdot N_P) = 0. \tag{1}$$

As formulated in [22, 23], the minimization consists of two terms corresponding to (1) the surface normal N_P and principal curvatures κ_M and $\kappa_{\mathcal{M}}$, and (2) the principal direction M_P[3]. To simplify the analysis, each is minimized independently. The first term, E_1, follows directly from [22, 23]:

$$E_1 = \sum_{\alpha=1}^{n} \|N_P - N_{P\alpha}\|^2 + (\kappa_M - \kappa_{MP\alpha})^2 + (\kappa_{\mathcal{M}} - \kappa_{\mathcal{M}P\alpha})^2 + \lambda((N_P \cdot N_P) - 1) \tag{2}$$

where $\xi_P = (P, \kappa_{MP}, \kappa_{\mathcal{M}P}, M_P, \mathcal{M}_P, N_P)$ and $\xi_{P\alpha} = (P_\alpha, \kappa_{MP\alpha}, \kappa_{\mathcal{M}P\alpha}, M_{P\alpha}, \mathcal{M}_{P\alpha}, N_{P\alpha})$. Using standard methods, one obtains the following updating functionals for N_P, κ_{MP}, and $\kappa_{\mathcal{M}P}$:

$$N_P^{(i+1)} = \frac{(\sum_{\alpha=1}^{n} N_{x_{P\alpha}}^{(i)}, \sum_{\alpha=1}^{n} N_{y_{P\alpha}}^{(i)}, \sum_{\alpha=1}^{n} N_{z_{P\alpha}}^{(i)})}{\sqrt{(\sum_\alpha N_{x_{P\alpha}}^{(i)})^2 + (\sum_\alpha N_{y_{P\alpha}}^{(i)})^2 + (\sum_\alpha N_{z_{P\alpha}}^{(i)})^2}} \qquad \kappa_M^{(i+1)} = \sum_{\alpha=1}^{n} \frac{\kappa_{MP\alpha}^{(i)}}{n} \qquad \kappa_{\mathcal{M}}^{(i+1)} = \sum_{\alpha=1}^{n} \frac{\kappa_{\mathcal{M}P\alpha}^{(i)}}{n} \tag{3}$$

where the superscript i refers to the current iteration step.

Because M_P and $\mathcal{M}_P$ are *directions*, there is a 180° ambiguity in orientation. For this reason the formulation of minimization term E_2 needs to be re-cast from that described in [22, 23]. We avoid the ambiguity by minimizing the difference of directions in the tangent plane at point P as follows. Express M in tangent plane coordinates as

$$M_P = \bar{b}_1 \cos\theta + \bar{b}_2 \sin\theta, \ (0, 2\pi) \text{ s.t. } \bar{b}_1, \bar{b}_2 \in T_P, \ \|\bar{b}_1\| = \|\bar{b}_2\| = 1, \text{ and } (\bar{b}_1 \cdot \bar{b}_2) = 0 \tag{4}$$

[2]The technical details of how this is accomplished are described in [22].

[3]Since $\mathcal{M}_P$ is orthogonal to both M_P and N_P, it need not be considered.

Then

$$E_2 = \min_{\theta} \sum_{\alpha=1}^{n} \left[1 - (M_P(\theta) \cdot M_{P\alpha})^2\right] \tag{5}$$

$M_P^{(i+1)}$ is found by substituting the value of θ that minimizes (5), back into (4). Again, using standard methods, one obtains the following updating functional for θ:

$$\theta^{(i+1)} = \tan^{-1}\left[\frac{(A_{22} - A_{11}) + \sqrt{(A_{11} - A_{22})^2 + 4A_{12}^2}}{2A_{12}}\right], \quad A_{ij} = \sum_{\alpha=1}^{n} (M_{P\alpha} \cdot \bar{b}_i)(M_{P\alpha} \cdot \bar{b}_j). \tag{6}$$

Note that this also determines the solution for $\mathcal{M}_P^{(i+1)}$ for the reason cited earlier.

Control over iteration is maintained by tracking the convergence of the derivative of a composite measure R_S, which is the sum of local difference measures computed over the surface,

$$R_S^{(i)} = \sum_j R_j(\xi_P^{(i)}, \xi_{P\alpha}^{(i)}) = \sum_j E_{j1}^{(i)} + E_{j2}^{(i)}, \quad P_j \in S. \tag{7}$$

The algorithm is allowed to interate until the difference $|R_S^{(i)} - R_S^{(i-1)}|$ falls below a specified threshold. A discussion of the convergence properties is beyond the scope of this paper and is addressed in [17, 22, 23]. However, we have confirmed empirically over a large number of experiments that the algorithm produces stable results quite rapidly, generally within 5 iterations.

2.2 Identifying Feature Trace Points

The point of the minimization strategy is to obtain a description of the surface S that is *stable* with respect to further interpretation [4]. This allows for a more direct interpretation of features and specifically avoids having to deal with the problem at the level of feature interpretation, e.g. [5]. The determination of features used to partition the surface is a case in point.

Hoffman & Richards [15] argue that a natural basis for surface decomposition is the principle of transversality regularity. Simply stated, the interpenetration of two arbitrarily shaped surfaces (i.e. corresponding to different parts) results in a contour of concave discontinuity of their tangent planes. In the context of smooth surfaces this translates into the partitioning of S into parts at loci of negative minima of each principal curvature along its associated family of lines of curvature [15]. We will refer to such loci as *critical points*. Thus, it is important to have stable estimates of the principal curvatures and directions at each point on S. The following procedure is used to determine the loci of critical points on smooth surfaces.

Let $\kappa_M(x,y)$ and $\kappa_{\mathcal{M}}(x,y)$ represent stable estimates of the principal curvatures of S sampled on the discrete grid (x,y), with corresponding principal directions $M(x,y)$ and $\mathcal{M}$ (x,y). The directional derivatives in these directions are $\kappa'_M(x,y)|_M$ and $\kappa'_{\mathcal{M}}(x,y)|_{\mathcal{M}}$ respectively. Then P is deemed to be a critical point iff

$$\kappa'_M(x,y)|_M = 0 \quad \text{AND} \quad \kappa_M(x,y) < 0 \qquad \text{OR} \qquad \kappa'_{\mathcal{M}}(x,y)|_{\mathcal{M}} = 0 \quad \text{AND} \quad \kappa_{\mathcal{M}}(x,y) < 0. \tag{8}$$

As presently implemented, the curvature consistency algorithm does not have an explicit representation for orientation discontinuities, but does make use of such information (i.e. an externally computed discontinuity map) in the updating procedure. For example, a local edge operator can be used to provide an estimate of surface discontinuities [13]. While this does not solve the problem of correctly localizing all discontinuities on a surface, it can be used to significantly reduce edge smoothing in the reconstruction procedure. As far as the identification of critical points due to concave discontinuities is concerned, these will be smoothed into negative local minima and can be identified as outlined above. But jump discontinuities, caused either by self-occlusions of the object or occlusions by other objects in the scene, are also necessary for the partitioning task.

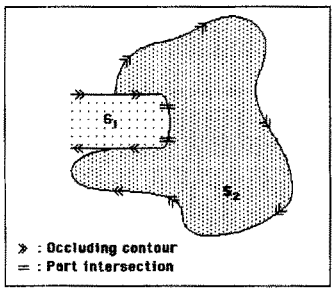

Figure 3: Partitioning contours on a surface

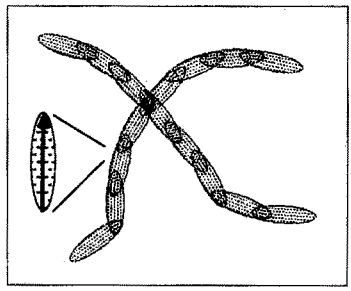

Figure 4: Tangents and the potential field

Such points can be identified from ξ_P by considering the angle between the surface normal N_P and the view vector V. Furthermore, if S is assumed to be smooth and in orthographic projection to the viewer, then the Z component of N_P will roll off to zero along occluding contours. It is this latter property that we use in identifying the trace points of the occluding contour. In fact, because of the stability of ξ_P as a result of reconstruction, even a thresholding of Z component values can suffice in localizing these points.

3 Piecing Together Partitioning Contours

Critical points, orientation discontinuities, and jump discontinuities are not themselves a solution to the parts decomposition problem, but they can provide the materials necessary to infer the partitioning contours. We will refer to such points collectively as the set of *trace* points $\{P_n^i\}$, where n is an index on the set. The second computational task is, given $\{P_n^i\}$ and the associated set of frames $\{\xi_{P_n}^i\}$, find the set of integral curves, $v_k(s) = (x_k(s), y_k(s), z_k(s)), \quad k \in (1, \text{no. contours})$ that partition the surface S (Figure 3).

The approach used to solve this problem involves the use of energy-minimizing spline fitting, [16, 24, 25], using the strategy devised by Zucker et al. [26] for finding a global covering of plane curves through a 2-D tangent field. This strategy is best understood by analogy. Consider what happens when iron filings are distributed on a piece of paper with a magnet placed beneath. With a bit of shaking the filings eventually align with the magnetic field and smoothly interpolate the lines of flux. In the tangent field model, the iron filings become unit length snakes that align to a potential field generated by the set of trace points and their tangent directions. That is, for each trace point, one associates a local potential with an asymmetric Gaussian profile as shown in Figure 4. The sum of these local potentials defines the field.

In principle, one can extend the tangent field model to three dimensions by associating a 3D asymmetric Gaussian envelope with each frame of $\{\xi_{P_n}^i\}$ and including a torsion component in the minimization functional. However, for surfaces acquired from single viewpoint and without large foreshortening, it is often sufficient to assume that $z_k(s) \approx 0$. That is, one can use the planar model by projecting each frame ξ_P onto the view plane[4], where the principal direction M is used to generate a planar potential field as in the tangent field model. The resulting plane curves are used to segment a depth map of the surface into regions corresponding to parts.

The formulation of the minimization problem is as follows. Let $v(s,t) = (x(s,t), y(s,t)), \quad 0 \leq s \leq 1$ represent a deformable curve with kinetic energy functional $T(v)$ defined as

$$T(v) = \frac{1}{2} \int_0^1 \mu |v_t|^2 ds, \tag{9}$$

[4]The XY plane, assuming orthographic projection.

where μ is the (constant) mass density, and the potential energy functional $U(v)$ defined as

$$U(v) = \frac{1}{2}\int_0^1 \Big(\omega_1(s)|v_s|^2 + \omega_2(s)|v_{ss}|^2 + I(v) + S(v)\Big)ds, \tag{10}$$

where $\omega_1(s)x_s$ controls the tension of the curve, $\omega_2(s)x_{ss}$ controls the rigidity of the curve, $I(v)$ is the potential field coming from $\{\xi^t_{P_n}\}$, and $S(v)$ is a force between neighbouring curves that operates when they are in close proximity.

The space curves that we seek are described by those functions $x(s,t)$ and $y(s,t)$ for which

$$\int_{t_0}^{t_1} T(v) - U(v) \quad dt \tag{11}$$

is a minimum. Zucker et al. [26] describe a solution for the 2D case obtained from the calculus of variations. This method was used in the experiments described in Section 5. Surface regions corresponding to parts are obtained from the set of contours $\{v_k(s)\}$ using a conventional region labeling algorithm.

This model of surface partitioning will not suffice for a single view without an additional assumption. That is, the silhouette contour is assumed to close a part boundary for those cases where the endpoints of the boundary terminate on the contour. In other words, parts can be cut out of surface by a segment drawn between two points on the silhouette contour. This will become apparent in the experimental results presented in Section 5.

4 Fitting Part Models

The set of contours $\{v_k(s)\}$ partition the surface S into a set of regions $\bigcup_l S_l$. For each l = (1, no. parts), we now seek to infer a corresponding volumetric element V_m that best characterizes the 3-D shape of the part. Different subscripts are used to signify the fact that one or more surface patches S_l can map to a single volumetric element V_m, e.g. where a surface is occluded or where multiple viewpoints are involved. For the purposes of this paper, however, it is assumed that $l = m$. Given a set of volumetric primitives Γ, the final computational task consists of (1) finding $V_m \in \Gamma$ that best characterizes a particular S_l, and (2) determining the parameters of V_m by minizing

$$|V_m(x,y,z) - S_l(x,y,z)|. \tag{12}$$

There are a number of different approaches to solving this problem. They range from simple geometric approximations [12] to more complex fits using superquadric models [1, 6, 14, 19]. Pentland demonstrated how superquadrics could be used to advantage in representing a wide variety of shapes. His success with this representation has motivated others, including ourselves, to investigate the use of superquadrics for object and part models. But the task is not at all straightforward. The minimization represented by (12) does not necessarily possess a single global minimum, i.e. the problem is underdetermined, and additional constraints are needed to select a suitable minimization. For example, it is not possible to determine the depth of a only one face, but if we add the additional constraint that the volume be minimized, then the only possible volumetric solution is a thin plate.

Our approach is somewhat of a hybrid method. A simple fit to an ellipsoidal model is used to provide the starting point for a subsequent iterative fit to a superquadric

$$F = \left(\left|\frac{x}{a_x}\right|^{\frac{2}{\epsilon_2}} + \left|\frac{y}{a_y}\right|^{\frac{2}{\epsilon_2}}\right)^{\frac{\epsilon_2}{\epsilon_1}} + \left|\frac{z}{a_z}\right|^{\frac{2}{\epsilon_1}} = 1. \tag{13}$$

For a superquadric in an arbitrary position 11 parameters have to be found: three translation parameters; three rotation parameters; two shape parameters (ϵ_1 and ϵ_2); and three extent parameters (a_x, a_y, and a_z). The goal is to attempt to restrict the search and solution space by starting the minimization in the correct neighbourhood.

4.1 Initial Fit to an Ellipsoid Model

In our experience the iterative fitting procedure will converge to an acceptable solution provided it has a good initial estimate of the rotation and translation parameters of the volumetric primitive V_m. Only a rough estimate of the extent parameters is required. The shape parameters are not critical and can be initialized so that the superquadric starts as an ellipsoid ($\epsilon_1 = 1$, $\epsilon_2 = 1$).

We initialize the translation parameters to the centroid of the points in the surface patch S_l. Like Ferrie [11] the initial rotation parameters are found by aligning the axes of the ellipsoid along the principal moments of inertia of S_l about the centroid. The extent parameters are set to the maximum projected length of a vector from the centroid to a point on S_l for each ellipsoid axis.

No attempt has been made to compensate for the fact that the patch corresponds to only a partial view of the surface. Others, e.g. [11], have imposed symmetry constraints to improve the initial estimate, but in practice such methods have little impact on the final result.

4.2 Iterative Fit to a Superquadric Model

We have used the Levenberg-Marquardt gradient descent procedure [20] to minimize the error of fit between a superellipsoid surface V_m and a surface patch of range data S_l. The method requires calculation of the Jacobian of an error of fit function with respect to the adjustable parameters for each point of range data. Unlike others, [1, 6], we have not used Poisson distributed "jitter" to avoid local minima during the fitting process. Generally the solutions reached have been acceptable but in a small number of cases the procedure would benefit from the technique.

We have tried the true euclidean measure of fit error as suggested by Gross and Boult [14] but found that while it tended to fit the range data well, there was a tendency for the volume of the superquadric to become very large, especially when the surface patch was flat. We have found that the minimum volume measure motivated by Bajeszy and Solina produces more intuitive results [1]. Gross and Boult [14] defined a modified error of fit based on this measure as follows

$$R = \sqrt{a_x a_y a_z}\,(F^{\epsilon_1}(\bar{x}_w, \bar{a}) - 1)\,, \tag{14}$$

where $F(\bar{x}_w, \bar{a})$ is the translated and rotated version of the inside-outside function shown in (13) and is 1 when the data lies exactly on the surface. The factor $a_x a_y a_z$ is a measure of the superquadric volume, so, given any set of superquadrics models that fit data equally well as measured by $F(\bar{x}_w, \bar{a})$, the smaller members of that set will have the lower overall error and will be preferred by the fitting procedure. Raising the inside-outside function to the power ϵ_1 shapes the error of fit to be more quadratic and more suited to rapid convergence under the assumptions of the Levenberg-Marquardt method.

5 Experiments

The results of two experiments are now presented that show how our strategy works for computing articulated volumetric descriptions of objects from laser rangefinder data. Range images used in these experiments are part of a standardized database available through the National Research Council of Canada [21]; each has been averaged down to a 256 × 256 by 12-bit resolution. The majority of the computation performed was done on a Symbolics 3650 Lisp machine. Rendering of the object models was done using the SuperSketch modeling system.

5.1 The Doll

Figure 5a shows an image rendered from a depth map of a toy doll which is lying face down. The first stage of processing uses the curvature consistency algorithm outlined earlier to reconstruct

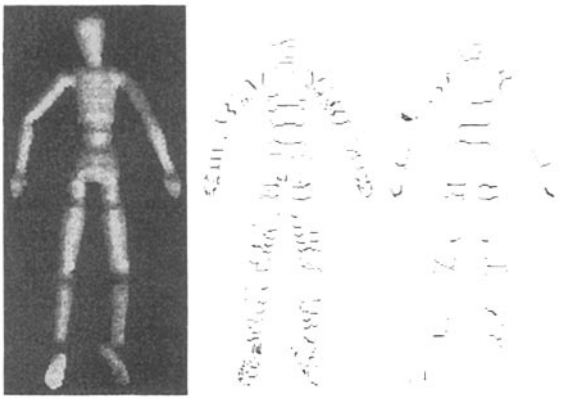

Figure 5: (a) Range image of the doll (b) Initial trace points (c) Trace points after 5 iterations

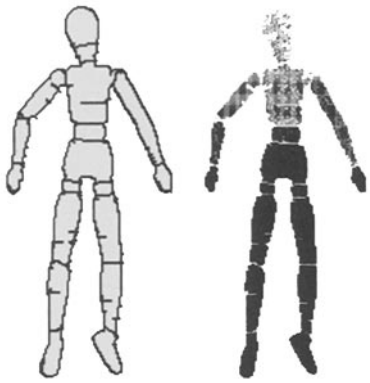

Figure 6: (a) Snake contours at steady state. (b) Regions computed from the snake contours

the surface, computing a stable intermediate description in the process. This is best illustrated by comparing Figures 5b and 5c, which show the trace points computed from the initial frame estimates, and after 5 iterations of the algorithm respectively. Other elements of $\xi_{P\alpha}$ are similarly stabilized. In examining Figure 5c, notice how the trace points serve to delineate the part boundaries.

The frames $\{\xi^t_{P_n}\}$ associated with the set of trace points $\{P^t_n\}$ are used to generate a potential field for the second stage of processing. Accurate determination of the frame directions is important because this field is generated from a sum of local potentials, each oriented in its respective M_P. Snakes are then deposited at each trace point and are allowed to evolve in the potential field according to (11) until the system reaches a steady state. Figure 6a shows the resulting contours obtained after running the algorithm. Because of the dense covering of trace points at the part boundaries, the result is not much different from that shown in Figure 5c. The interpolating behaviour of the snakes does become important, however, when trace points become sparser. A region map, Figure 6b, is derived from the snake contours by clustering the interior regions. Using this procedure, the algorithm was able to correctly locate the parts of the doll with one exception. The right elbow joint was not located, largely because the arm is straight. For this reason the forearm and upper arm are grouped together as a single part.

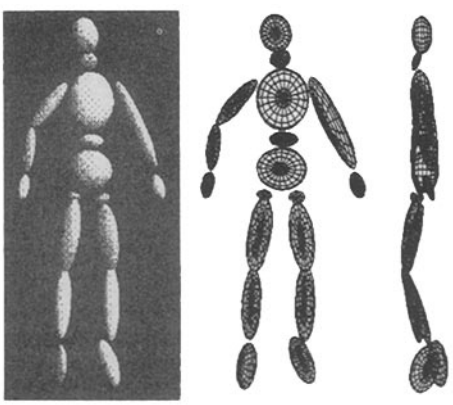

Figure 7: (a) Shaded image of the initial ellipsoid fit (b) Wire frame showing parts (front) (c) (side)

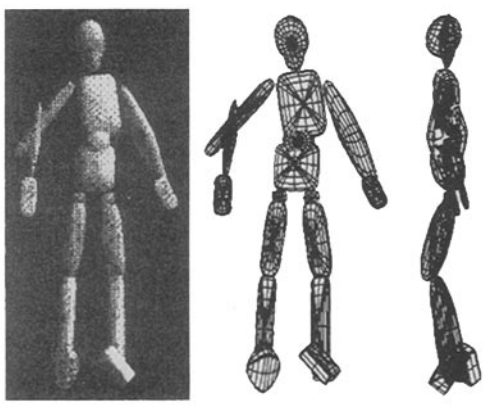

Figure 8: (a) Shaded image of the final superquadric fit (b) Wire frame showing parts (front) (c) (side)

The final stage of processing consists of fitting superquadric primitives to each of the regions

located above. This proceeds in two stages, beginning with an initial approximation using ellipsoids to define initial starting points. Figure 7a shows a rendering of these initial fits as a shaded image; front and side profiles of the same model are rendered as wire frames in Figures 7b and 7c respectively. Superquadrics are then fit to each region using the gradient descent algorithm described earlier, with initial conditions and volume constraints provided by the initial ellipsoid fits. Figures 8a–8c show the results obtained by this procedure from the same viewpoints as shown earlier in Figures 7a–7c. From a qualitative viewpoint, the results are very pleasing and capture the essential structure of the doll.

5.2 A Toy Horse

The second example is of a toy horse (actually a unicorn) shown in Figure 9a. What is particularly interesting about this example is the complex nature of its surfaces. A comparison between the initial (Figure 9b) and final (Figure 9c) principal direction fields, i.e. $\mathcal{M}_P$, shows how the curvature consistency algorithm is able to correctly recover directional properties of the surface. Using the same procedure as above for the statue, the volumetric model shown in Figure 10a is obtained. A wire frame corresponding to this model showing each part is shown in Figure 10b, and a depth map generated from the volumetric model is shown in Figure 10c.

Figure 9: (a) Range image of the toy horse (b) Principal direction field computed from initial estimates (c) Principal direction field after 5 iterations

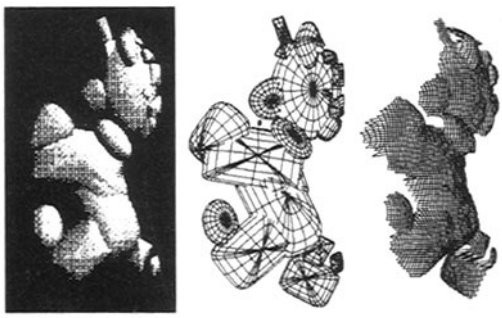

Figure 10: (a) Superquadric model rendered as a shaded image (b) Wire frame of the model showing the parts (c) Depth map computed from the model

6 Discussion and Conclusions

Although the approach we have just presented is advertized as being a method for computing volumetric models from laser rangefinder data, it is intended to provide a more general framework that includes other kinds of sensors. The augmented Darboux frame used in reconstruction can be estimated from either orientation or depth. It has already been applied to shape-from-shading [9]. In fact, sensor fusion is possible using this kind of minimization with a minimal change in the formulation. At present, no attempt is made to use cues provided by the silhouette contours in the parts decomposition. It is assumed that the observer is sufficiently active so as to choose a vantage point where part intersections are visible. A more comprehensive strategy must include analysis of contour as well as representation at multiple scales. Finally, the methodology for fitting part models needs to be extended. One of strengths of Pentland's [19] method is that model fitting takes the structure of neighbouring parts into account. By running the fitting procedure in parallel, constraints from neighbouring parts can be used to improve the local fit. These topics are currently under investigation in our laboratory.

References

[1] R. Bajcsy and F. Solina. Three dimensional object recognition revisited. In *Proceedings, 1ST International Conference on Computer Vision*, London,U.K., June 1987. Computer Society of the IEEE, IEEE Computer Society Press.

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

[3] P. Besl and R. Jain. Segmentation through symbolic surface description. In *Proceedings IEEE Conf. Computer Vision and Pattern Recognition*, pages 77–85, Miami Beach, Florida, June 1986.

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

[5] P. Boulanger. Label relaxation technique applied to the topographic primal sketch. In *Proceedings, Vision Interface 1988*, pages 158–162, Edmonton, Canada, June 1988.

[6] T. Boult and A. Gross. Recovery of superquadrics from depth information. In *Proceedings of the AAAI workshop on spatial reasoning and multisensor integration*, pages 128–137. American Association for Artificial Intelligence, 1987.

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

[8] T. Fan, G. Medioni, and R. Nevatia. Description of surfaces from range data using curvature properties. In *Proceedings IEEE Conf. Computer Vision and Pattern Recognition*, pages 86–91, Miami Beach, Florida, June 1986.

[9] F. Ferrie and J. Lagarde. Robust estimation of shape from shading. In *Proceedings 1989 Topical Meeting on Image Und. and Machine Vision*, Cape Cod, Massachusetts, June 1989.

[10] F. Ferrie, J. Lagarde, and P. Whaite. Towards sensor-derived models of objects. In *Proceedings. Vision Interface '89*, London, Ontario, June 19-23 1989.

[11] F. Ferrie and M. Levine. Piecing Together the 3-D Shape of Moving Objects: An Overview. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition*, pages 574–584, San Francisco, CA., June 1985.

[12] F. Ferrie and M. Levine. Deriving Coarse 3D Models of Objects. In *IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition*, pages 345–353, University of Michigan, Ann Arbor, Michigan, June 1988.

[13] G. Godin and M. Levine. Structured edge map of curved objects in a range image. In *Proceedings IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition*, San Diego, California, June 4-8 1989.

[14] A. D. Gross and T. E. Boult. Error of fit measures for recovering parametric solids. In *Proceedings, 2ND International Conference on Computer Vision*, pages 690–694, Tampa,Florida,UK, Dec. 1988. Computer Society of the IEEE, IEEE Computer Society Press.

[15] D, Hoffman and W. Richards. Parts of recognition. *Cognition*, 18:65–96, 1984.

[16] M. Kass and D. Terzopoulos. SNAKES: active contour models. *International Journal of Computer Vision*, 1:321–332, 1988.

[17] J. Lagarde. Constraints and their satisfaction in the recovery of local surface structure. Master's thesis, Dept. of E.E., McGill Univ., 1989. in preparation.

[18] P. Parent and S. Zucker. Curvature consistency and curve detection. *J. Opt. Soc. Amer., Ser. A*, 2(13), 1985.

[19] A. Pentland. Recognition by parts. In *Proceedings, 1ST International Conference on Computer Vision*, London,UK, June 1987. Computer Society of the IEEE, IEEE Computer Society Press.

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

[21] M. Rioux and L. Cournoyer. The nrcc three-dimensional image data files. National Research Council of Canada, CNRC No. 29077, June 1988.

[22] P. Sander. *Inferring Surface Trace and Differential Structure from 3-D Images*. PhD thesis, Dept. Elect. Eng., McGill University, Montréal, Québec,Canada, 1988.

[23] P. Sander and S. Zucker. Inferring surface trace and differential structure from 3-d images. *IEEE Trans. PAMI*, 1990. To appear.

[24] D. Terzopoulos. On matching deformable models to images. *Topical Meeting on Machine Vision, Technical Digest Series*, 12:160–163, 1987.

[25] D. Terzopoulos. On matching deformable models to images: Direct and iterative solutions. *Topical Meeting on Machine Vision, Technical Digest Series*, 12:164–167, 1987.

[26] S. Zucker, C. David, A. Dobbins, and L. Iverson. The Organization of Curve Detection: Coarse Tangent Fields and Fine Spline Coverings. In *Proceedings, 2ND International Conference on Computer Vision*, Tampa,Florida,USA, Dec. 1988. Computer Society of the IEEE, IEEE Computer Society Press.