Robust Feature Detection for 3D Object Recognition and Matching

Sharath Pankanti, Chitra Dorai and Anil K. Jain
Pattern Recognition and Image Processing Laboratory
Department of Computer Science, Michigan State University
East Lansing, MI 48824-1027

Abstract

Salient surface features play a central role in tasks related to 3D object recognition and matching. There is a large body of psychophysical evidence demonstrating the perceptual significance of surface features such as local minima of principal curvatures in the decomposition of objects into a hierarchy of parts. Many recognition strategies employed in machine vision also directly use features derived from surface properties for matching. Hence, it is important to develop techniques that detect surface features reliably.

Our proposed scheme consists of (a) a preprocessing stage, (b) a feature detection stage, and (c) a feature integration stage. The preprocessing step selectively smoothes out noise in the depth data without degrading salient surface details and permits reliable local estimation of the surface features. The feature detection stage detects both edge-based and region-based features, of which many are derived from curvature estimates. The third stage is responsible for integrating the information provided by the individual feature detectors. This stage also completes the partial boundaries provided by the individual feature detectors, using proximity and continuity principles of Gestalt. All our algorithms use local support and, therefore, are inherently parallelizable.

We demonstrate the efficacy and robustness of our approach by applying it to two diverse domains of applications: (a) Segmentation of objects into volumetric primitives and (b) Detection of salient contours on free-form surfaces. We have tested our algorithms on a number of real range images with varying degrees of noise and missing data due to self-occlusion. The preliminary results are very encouraging.

1 Introduction

Conventional object recognition strategies have two primary components: a representation scheme and a matching technique. A good representation scheme attempts to capture the essential structure of the 'world' which will help in discriminating among the object categories in an effective manner. The abstraction of the world structure is formulated in terms of features of the objects in the domain and relations among them. Recent advances in surface reconstruction algorithms and the commercially available, reliable range sensors for obtaining depth data have spurred an increased interest in 3D object representation and matching. An important aspect of 3D object recognition is the extraction of surface features that can be used for effective representation and matching of objects using range data.

3D object recognition and model building tasks depend on either the explicit recovery or indirect inference of 3D features embodied in depth data. One of the most popular 3D object representation schemes for recognition systems employing range data is the representation in terms of its surface attributes. For most of the applications, 3D objects can be adequately modeled in terms of a collection of piecewise smooth surface patches. The features extracted from these surface patches and the spatial relationships among the surface patches capture the structure of the 3D world (restricted to the domain of application) and provide with the necessary constraints for matching. The local structural information relevant to a 3D shape is captured in the surface attributes such as principal curvatures, normals, and principal curvature direction fields. Features such as surface orientation, orientation discontinuity, depth discontinuity, surfaces curvatures etc. have been extensively used in machine recognition systems. In more restricted object domains, some of the surface features help group surface patches into higher level features, and thus, allow for reduction in the complexity of matching. For instance, minima of negative principal curvatures have been shown to play a perceptually significant role in 'part' decomposition of the sensed data. Other surface features such as the mean and the Gaussian curvatures also aid in determining the presence of an instance of a model object in the scene [1]. Directional properties such as the principal curvatures can be used in grouping the feature points into boundary contours and surface patches [2].

2 Robust Feature Detection

In differential geometry, it has been shown that a general smooth surface is uniquely characterized by the first and second fundamental forms [3]. The first fundamental form is primarily related to the surface normals. The surface curvature is a function of both these forms. The curvature at any point on the surface is also viewpoint invariant. In addition, interesting surface properties such as (i) jump boundaries (which are surface depth discontinuities), (ii) crease edges obtained from surface orientation discontinuities, and (iii) ridge lines that are smooth local extrema of curvature can also be inferred from the *zero-crossings* and extremal values of surface curvatures. Thus there is sound basis for computing salient features based on surface curvatures. In particular, we use invariant surface characteristics such as the mean and the Gaussian curvatures derived from the principal curvatures [3] to classify surface patches. Note that these invariants are local surface properties which allow them to be used

in situations involving occlusion of objects.

When the domain is restricted to objects having a definite part structure, the contours derived from local minima of negative principal curvature often constitute "part" boundaries [4]. These local minima of curvature and the boundaries detected from the discontinuities in the surface depth and surface normal field play a significant role in obtaining our proposed part decomposition.

While the efficacy of a recognition system depends upon a prudent choice of the features of the surfaces and their spatial relations constituting a representation scheme, the realization of such a system crucially depends on the robust estimation of these features and their relationships. Noise in the sensed data is often a primary reason for inaccurate estimates of the feature values. This necessitates smoothing of the sensed data before extracting features. The surface features detected from the smoothed data, however, are not perfect due to several reasons: low signal to noise ratio in the sensed image, artifacts of the feature detectors themselves, etc. It is desirable that the complementary strengths of the individual feature detectors and the contextual information be integrated to produce a refined feature map.

We propose a three-stage system for extracting robust features from the sensed data: (a) preprocessing, (b) feature detection, and (c) feature integration. The preprocessing step selectively smoothes out noise in the depth data without the loss of salient surface details and permits reliable local estimation of the surface features. The feature detection stage detects both edge-based and region-based features, of which many are derived from curvature estimates. The third stage is responsible for integrating the information provided by the individual feature detectors. This stage also completes the partial boundaries provided by the individual feature detectors, using proximity and continuity principles of *Gestalt*.

In many applications, the assumption of second order smoothness ($\mathcal{C}^2$) for each of the surface patches is not overly restrictive. Hence, some form of smoothing imposing a piecewise ($\mathcal{C}^2$) continuity improves the accuracy of estimated surface. The curvature consistency algorithm of [2] is one such method of adaptive smoothing.

The key idea behind curvature consistency is to smooth the surface while preserving its local structure described by surface normal orientation, principal curvatures and their directions at a point on a surface. It can be viewed as the second stage of processing applied to the estimates of surface attributes from depth values using local methods. It is formulated as an optimization problem in which the objective is to minimize a functional form embedding a minimum variation of curvature in a local neighborhood [2]. The three components involved in this process are: (i) a local description of the surface at a point P, (ii) a *transport model* that describes the change in the local description as P is moved to an adjacent point Q and vice versa, and (iii) a function that prescribes how the local description at P can be updated in order to be compatible with the descriptions of its local neighbors once they have been moved from Q to P by the transport mechanism. In the implementation of this algorithm, the surface normals and curvatures are estimated using a local neighborhood in the first step, and these local estimates are smoothed based on the consistency of curvatures in the neighborhood by an iterative process. The algorithm is found to converge rapidly. See [2] for more details on this method.

It should be noted that the curvature consistency algorithm provides the smoothed data, surface normals, and surface curvatures for the smoothed data. From these initial features, the method for obtaining the desired feature map, the feature integration scheme, and derivation of final representation will vary from application to application depending on the task at hand and the object domain. We present details of the application of our robust feature detection technique to (i) the task of deriving part-based description and (ii) the task of detecting salient contours on free-form surfaces.

All our algorithms use local support and, therefore, are inherently parallelizable. We demonstrate the efficacy and robustness of our approach on several real images.

3 Volumetric Primitives from Range Data

Our objective here is to implement a feature–based 3D object recognition system inspired by Biederman's *RBC* theory [5]. Based on psychophysical evidence, this theory hypothesizes that the perceptual recognition of objects is conceptualized to be a process in which an input image is segmented at regions of deep concavity into a few, simple volumetric primitives (*"parts"*), called *geons*. An object can be efficiently recognized if an arrangement of two or three components can be recovered from the image [5, 6, 7, 8].

We make the following simplifying assumptions about the object domain.

- The objects are rigid and opaque.
- The part decomposition of each of the object is unique and unambiguous. Each *part* of the object is piecewise smooth and not very convoluted nor very eccentric.
- The objects are simply connected in a topological sense. We exclude objects like doughnuts since their part decomposition is not conceptually simple.
- Further, we assume that all the prominent features of a given object can be detected by the sensor [9]. We exclude *origami* world objects.

Most of the objects around us have a definite part structure and well within the boundaries laid down by this set of assumptions. Many complex industrial objects can be thought of as a composition of a small number of well-defined volumetric primitives.

Problem Definition : Given a 2.5D sensed data of an industrial scene consisting of an unoccluded object from a non-accidental viewpoint, we desire to derive a description of the object in terms of volumetric primitives and to obtain the most feasible interpretation [10] of the sensed data in terms of the objects stored in the model database.

3.1 Representation

Our representation scheme is based on a set of geon-like volumetric part primitives. The original set of geons proposed by RBC theory is based on four geometric properties involving shape of axis, shape of the cross-section, and the variation of cross-section along the axis. Recognizing

that the computation of cross-section is somewhat problematic, we have defined a reduced set of 12 *geon*-like volumetric primitives based on the following three geometric properties of the "parts": straightness or curvedness of their axes, straightness or curvedness of the boundaries of their cross-section, and the variation of the area of cross-section along their axes (constant, increasing, and increasing-decreasing). An object can be represented as an attributed graph whose vertices represent the geons of the object and the arcs between the vertices represent adjacency relationship between the corresponding geons of the object. Let V be a set of vertices, each representing a geon in an object. Let $E \subseteq V \times V$ be the set of arcs representing adjacency relations among the pairs of geons of the object such that an ordered pair $(v_i, v_j) \in E$ iff geons v_i and v_j are adjacent (share a border). We also define following terms:

- **Distance function,** $d : V \times V \rightarrow \mathcal{R}$, maps each pair of geons to a real number representing the distance separating their center of masses.
- **Type function,** $t : V \rightarrow T$, where T is the set of types of geons. Type of a geon is determined by the shape of its axis, the shape of boundaries of its cross–section, and the variation of the area of cross-section along its axis.
- **Angularity,** $l : \{(v_i, v_j, v_k)\} \rightarrow \mathcal{R}$ defined over three adjacent geons (v_i, v_j, v_k) such that if (v_i, v_j), $(v_j, v_k) \in V$ and θ is the angle between the segments joining center of masses of (v_i, v_j) and that of (v_j, v_k) then $l(v_i, v_j, v_k) = \theta$.
- **Attach function,** $a : E \rightarrow \{axial, sideways\}$, defines the type of attachment between neighboring geons. Notice that a could be an asymmetric function.
- **Size function,** $s : V \rightarrow \mathcal{R}$, assigns volume to each geon in the graph.

Thus, an object is defined by a weighted, attributed, and directed (WAD) graph G represented by a seven tuple (V, E, s, d, a, t, l).

3.2 Feature Extraction

We smooth the range data using the curvature consistency algorithm. Since the boundaries are computed from the depth data, we choose to call the resulting edge map 3D boundaries. An initial map of the the 3D boundaries can be computed by a typical range edge detector; we have used an MRF-based range edge detector [11]. We also extract the local minima of the principal curvature of depth data which we refer to as *trace points*. Since the curvature values are noisy, we avoid their direct participation in computing the 3D boundary computation. As in the computation of intensity edge map, the depth gradient values do not accurately represent the presence of a 3D boundary and the edge map estimated from the gradient values needs further processing. Usually an edge is terminated prematurely due to insignificant gradient values, even though a physical edge exists at that location. We have developed a boundary completion algorithm which attempts to extend each edge termination

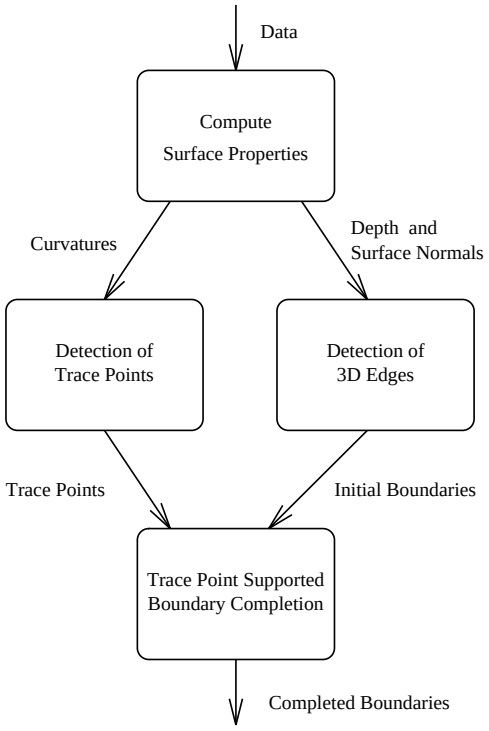

Figure 1: The surface segmentation module.

based on the curvature values and presence of other 3D edges in its spatial neighborhood.

In our algorithm, we only consider the trace points whose magnitude is above a certain threshold value. The boundary completion algorithm is essentially an implementation of the grouping strategy based on proximity and continuity. Each boundary termination is iteratively extended if evidence can be gathered in the form of presence of either a trace point or other boundary points in its restricted neighborhood. The decision about which candidate trace point to choose for the extension of the termination is based on a cost function given by:

$$c = |\theta| + d * 90.0,$$

where θ is the angle between the trace (boundary) point and the boundary termination in degrees and d is its distance from the boundary termination. Notice that this cost function prefers proximity grouping over continuity grouping. We find the concept of extension of boundary terminations using support of the trace points to be highly intuitive since this single strategy is compatible with both the concept of transversality as well as *Gestalt*.

A connected component analysis then determines closed regions from the boundaries detected by the boundary completion algorithm.We discard all boundaries which do not participate in the formation of closed boundaries. Part segmentation of the object is then derived by identifying all minimal closed regions bounded by edges which either (i) are external boundaries of the object or (ii) represent crease edges due to concavities. Each

of these closed regions is then hypothesized to represent a 'part' of the sensed object [4]. We then fit a deformable superquadric to the depth data of each 'part' thus identified [12]. The features derived from each superquadric determine the attributes of the geons and their relations resulting in the final object representation.

3.3 Matching and Verification

The recognition scheme proposed earlier involves matching of the geon structure (WAD) of each object in the model database to the geon structure extracted from the sensed data. Our approach to matching two attributed graphs is largely inspired from the concept of *interpretation tree*. Suppose that $(n-1)$ geons have been matched between the two attributed graphs. Then the n^{th} match is found based on locally consistent geon relations. The local consistency of a match between two geons from two attributed graphs is defined based on the attributes of the geons and the relation between the given geons and the geons adjacent to it. At present, we use angle and distance functions between the neighboring geons to determine local consistency.

Verification is undertaken to test the hypothesis generated by the local matching scheme. Our verification algorithm evaluates the global consistency of the hypothesized match by exhaustively evaluating the angle and distance constraints.

3.4 Experimental Results

The proposed strategy for part segmentation was implemented on a Sun SPARC-2 workstation and tested on more than a dozen range images of real objects scanned by a Technical Art's White scanner. Some of these input range images are shown in figures 2(a), 3(a), 4(a), 4(b), and 4(c). To highlight the depth information in these images, we have rendered the surface orientation as an intensity image; darker intensities denote surfaces oriented farther away from the direction of camera. There were small regions in the image where the sensor detected spurious depth values. These small regions were discarded before applying the part segmentation algorithm.

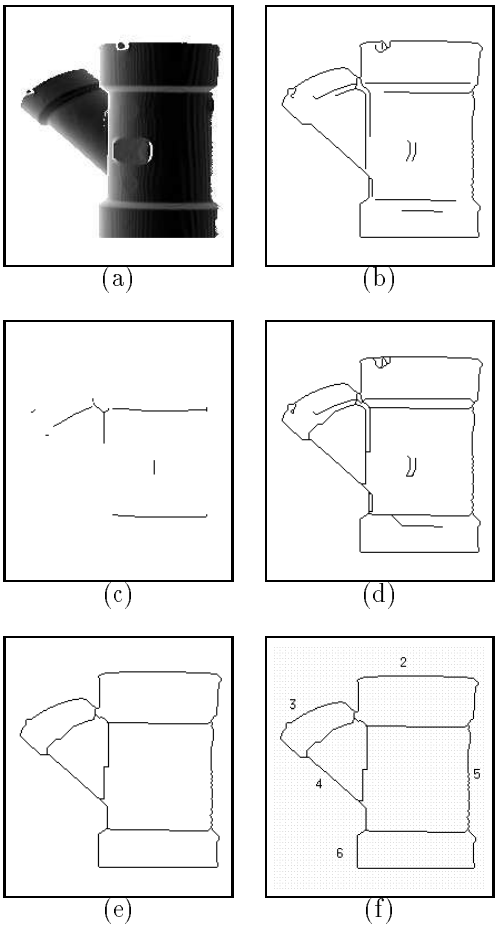

Figure 2: Intermediate representations for 'pipe1': (a) Input range image (size 240x240); (b) Output of the MRF edge detector; (c) Trace points detected; (d) Completed boundary edges; (e) Closed regions after merging small regions; (f) Part segmentation.

- **Part segmentation:** Figure 2 shows representations derived at various stages of the proposed part segmentation process. The output of MRF edge detector is shown in figure 2(b). Figure 2(c) shows the thinned trace points surviving the 90 percentile threshold on the magnitude of curvature. Figure 2(d) shows the boundaries completed by using the trace point supported boundary completion algorithm. Finally, small neighboring regions were merged if their size is less than 150 pixels. Figure 2(e) shows the set of all boundary points detected by the segmentation process. For the purpose of parts segmentation, we only use the silhouette edges which are defined by the local minima of the negative curvature, or their extensions. We don't permit participation of ridge boundaries defined by convexities of the objects in the part segmentation process. The result of the part segmentation for pipe1 is shown in figure 2(f).

  The segmentation obtained by the proposed method appears reasonable even in the presence of low to moderate laser "shadows". While in this present implementation, we have discarded some of the estimated boundaries, it appears that they can serve a useful role in refining the coarse part structure.

- **Fitting:** In order to reliably obtain the best superquadric fit determined by the global minimum of the error function, it is necessary that we initiate the algorithm from a reasonable neighborhood of the global minima. The strategy of deriving a good initial estimate of the superquadrics was guided by the following heuristic: Initial estimate of the major axis is in the direction of one of the eigenvectors of the data points deviating least from the mean direction of the minimal principal curvature of the visible surface [8].

  Figures 4(d), (e), and (f) show the wire frame diagrams of the superquadrics fitted to each part extracted by the part segmentation for the range images in figures 4(a), (b), and (c). In a few cases where the objects were comprised of parts with bent axes

and with a non-monotonic (increasing-decreasing) change in area of cross-section along their axes, the superquadratic fits were not good. Fifteen parameters were extracted from each superquadric fit to determine the type of geon. These parameters include relative orientation and position of each superquadric (ϕ, ψ, ω, c_x, c_y, and c_z), shape parameters (ϵ_1, ϵ_2), and taper parameters (K_x and K_y) [8]. Estimation of the attributes of geons from superquadric is motivated by [13].

Following observations could be made from the perusal of the various estimated parameter values of L-pipe images. The estimated value of the ϵ_2 is consistently greater than 0.7, indicating that cross section edges of each extracted part are curved. The radius of curvature of the superquadric axes is relatively large and the taper parameters (K_x and K_y) indicate that area of cross section of each part is constant along its major axis. Thus without an exception, all parts extracted from these range images can be classified as (straight axis, curved cross section area, constant area along the axis) type of geons. The features extracted for the part structure of an object include *distance, angle, type*, and *size* functions defined elsewhere. The estimates of angle between the axes of two parts are consistently very accurate. Figure 3 shows results of part segmentation and superquadric fitting for a range image of a doll.

- **Matching:** To evaluate the proposed recognition strategy, we designed the following matching experiment. We obtained six range images of an L pipe captured from a White scanner, four of which are shown in figures 2(a), 4(a), 4(b), and 4(c). All images were captured under identical sensing conditions (position of camera, calibration setup etc.). The only variable across the sequence of these images was the relative orientation of the L pipe. The extent of "shadows" cast on the surface of the sensed object in a White scanner setup depends upon the surface characteristics of the object as well as its geometry. As a result, the extent of deterioration of the sensed data is different in each range image. For instance, the magnitude of missing depth data shown in figure 2(a) is less than those shown in 4(a), 4(b), and 4(c). We then attempted an exhaustive pairwise matching between the representations derived from each pipe image.

  The recognition is carried out by a simple attributed graph matching algorithm. We have evaluated our matching algorithm by pairwise exhaustive matching of the part structures estimated from the six range images. Thus, in all 36 matching experiments were performed and in all cases no more than 8 hypotheses (out of 120 possible) were presented to the verification stage. The verification stage always produced the correct part correspondence and the correct 3D transformation required for orienting the sensed data to the model data.

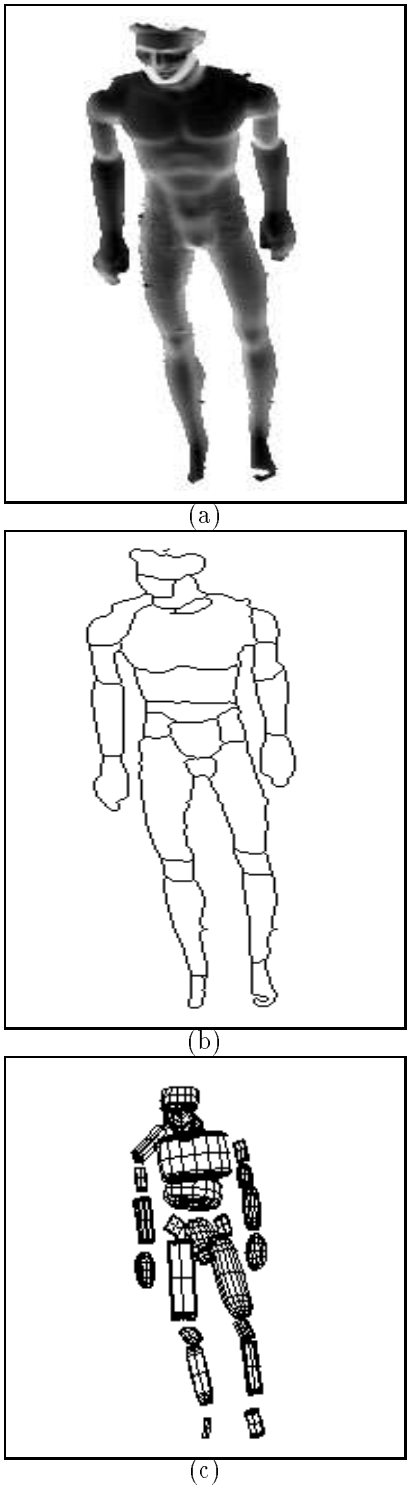

Figure 3: Intermediate representations of 'doll' image: (a) Input range image of a doll (size 240x600); (b) Part segmentation; (c) Volumetric representation.

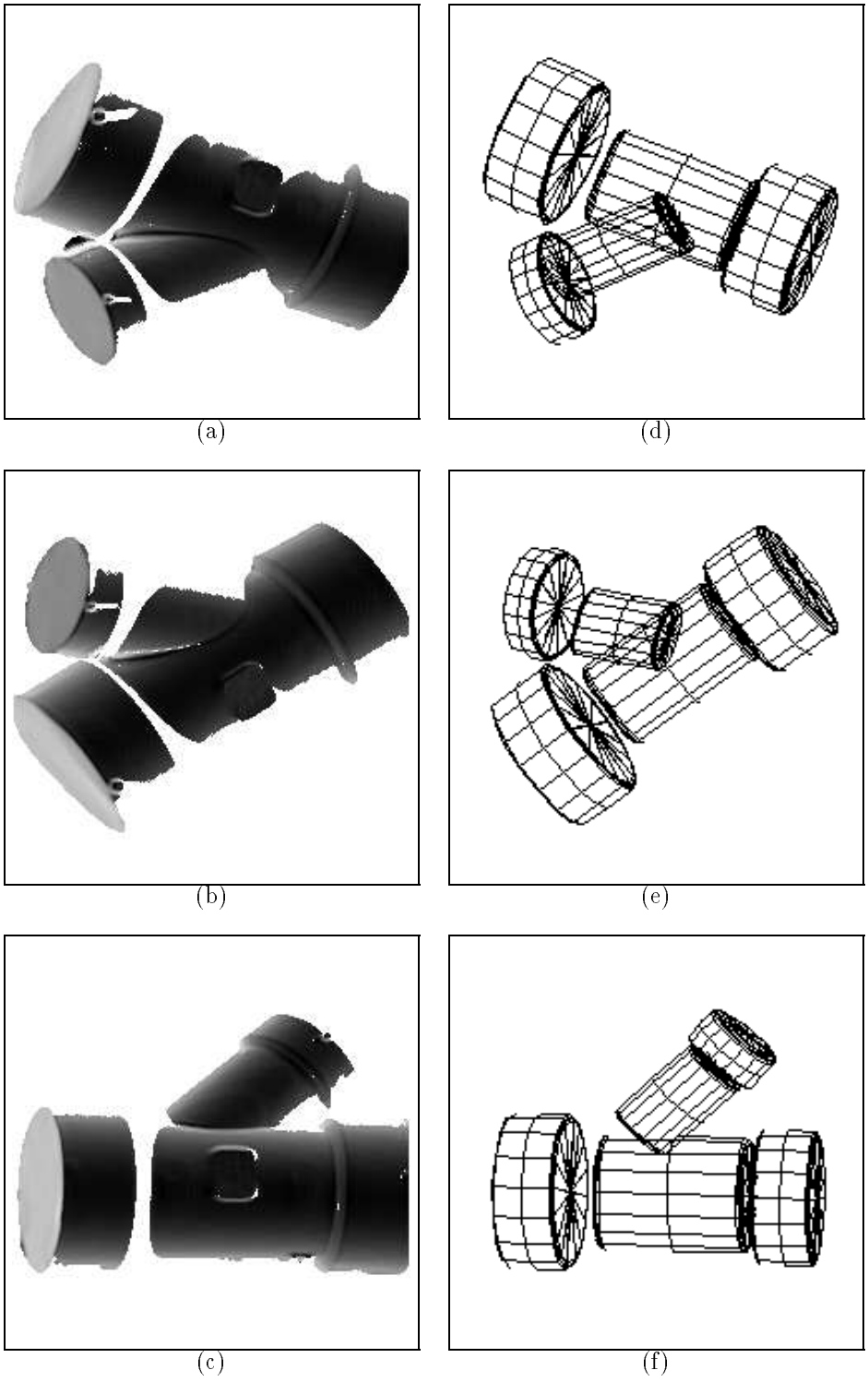

Figure 4: Input Range images of (a) pipe4 (size 192x240); (b) pipe5 (size 200x240); (c) pipe6 (size 167x240). The part-based volumetric representations derived from these range images are shown in (d), (e), and (f), respectively.

4 Salient Contours on Free-Form Surfaces

Free-form surface matching, often referred to as sculpted object recognition, is gaining increasing importance in the field of computer vision. A free-form surface is defined to be smooth such that the surface normal is well defined and continuous almost everywhere, except at vertices, edges, and cusps [14]. Examples of free-form surfaces include human faces, cars, airplanes, clay models, sculptures, terrains, etc. This increased interest has motivated current research in matching arbitrary free-form surface shapes that cannot be modeled using volumetric primitives and that may or may not have easily detectable landmark features such as vertices of polyhedra, vertices of cones and centers of spheres. Free-form surfaces may also be too complicated to be approximated by simple surface patches (planar, quadric, etc.). Furthermore, the parts-based approach would not work with sculpted objects that may be essentially smooth and featureless surfaces such as warped metal sheets and turbine blades.

4.1 Problem Definition

The problem of free-form surface matching can be posed as follows: Given 3D data describing a scene shape in a sensor coordinate system, and given a model shape in a model coordinate system, estimate the optimal rotation and translation that registers the model shape and the data shape. Some of the approaches for free-form surface matching use invariant local structural descriptions that are derived from an input scene data, and match these local descriptions with the stored model descriptions. The matching schemes use indexing or search methods employing heuristics to reduce the search. Some other approaches use parametric descriptions. The difficulties in comparing a scene surface description with a model surface description using parametric forms are: (i) it may not be possible to parameterize the scene surface accurately such that the derived parameters can be directly compared with those of the model surface and (ii) the surfaces may not be aligned in 3D space.

In our feature-based approach for free-form surface matching, we propose to use point and curve features if they exist on a surface to form hypotheses about models, and also use other surface invariants to be able to handle relatively featureless (no distinguishing points) surface shapes.

4.2 Feature Extraction using Curvatures

The local geometry of surfaces can be exploited to extract points or curves on the surfaces, so that these may provide us with features for representing free-form surfaces. These structural descriptions are based on properties that can be derived from small surface patches directly without resorting to defining the surface using volumetric primitives. The features we propose are a combination of edge and point features. We propose to use point features such as the Gaussian (K) and the mean (H) curvatures, that are invariant to translations, rotations and changes in parameterization, to derive our representation.

Though they themselves may not determine the surface completely, we can find enough features in the Gaussian and mean curvature images to recognize the surface. We propose a new method to detect both the jump and crease edges reliably from the depth data using the mean curvature values. We also integrate information present in the labels of surface points, obtained from a coarse segmentation of the range image using the sign of the K-H map.

4.2.1 Reliable Curvature Estimation using Curvature Consistency

The range data provides a grid of discrete points $z(i,j)$ and the surface normals and curvatures are estimated from this data. For each point P centered in a $n \times n$ window in a depth image, the surface is locally parameterized with a parabolic quadric fit, using a least-square technique. Once the parameters of the parabolic quadric form are determined, the normal direction, the principal curvature values, and their directions are estimated from these parameters about P on the surface. Since curvatures describe the local behavior of a surface, their estimated values depend crucially on the size of the neighborhood. Curvatures are also very sensitive to the effects of noise and quantization error. Reliable estimates of principal curvatures and their directions are important for our proposed curvature based representation of free-form surfaces. The normal and curvature estimates are first estimated using a local approximation, and then refined using the curvature consistency constraint [2] as described in an earlier section. This constitutes our preprocessing step.

4.2.2 Edge Extraction using Mean Curvatures

In our feature detection method, the extraction of jump and crease edges depends on the estimates of the mean and the Gaussian curvatures. From the reliable estimates of maximum (κ_{M_P}) and minimum ($\kappa_{\mathcal{M}_P}$) principal curvatures, we compute the mean curvature (H) and Gaussian curvature (K).

$$\begin{aligned} H &= (\kappa_{M_P} + \kappa_{\mathcal{M}_P})/2 \\ K &= \kappa_{M_P} \times \kappa_{\mathcal{M}_P} \end{aligned}$$

We observe here that the curvature properties correspond to certain significant physical properties of a surface. The occluding boundary or the *jump* boundary creates a zero crossing of the curvature in a direction normal to that of the boundary. A *crease* boundary where surface normals are discontinuous causes a local extremum of the curvature at that point. Furthermore, principal curvature extrema correspond to certain distinguished points or lines on smooth surfaces.

Since the mean curvature at a surface point is the average of the two principal curvature values at the point, it captures the underlying discontinuities in depth and surface normals just as well as the principal curvatures themselves. We use the Laplacian edge detector with a 4x4 mask to detect the changes in the values of mean curvatures. The output of the Laplacian edge detector is then thresholded to retain only the upper 10 percentile of the gradient values to yield the edges. The threshold is determined empirically. Since thick edges are formed as a result of thresholding, we apply a thinning operator on this thresholded edge output to obtain a single pixel thick edge map.

4.2.3 Integration of Segmentation Labels and Edges

During the integration step, we classify the edges as jump and crease edges based on the surface labels. We integrate the region information provided by the segmentation of the surface based on the sign of the Gaussian and mean curvatures, along with the raw edge map obtained above. We first segment the surface into regions based on the sign of the K-H map, and label each surface pixel as one of the following eight types. The labeling is given below [15]:

- H negative, K positive – peaked surface
- H positive, K positive – cupped surface
- H positive, K negative – saddle valley surface
- H negative, K negative – saddle ridge surface
- H positive, K zero – valley surface
- H negative, K zero – ridge surface
- H zero, K zero – plane surface
- H zero, K negative – minimal surface

It can be seen that the crease edges which are the local extrema of surface curvatures are in fact, the boundaries between the peaked and saddle valley regions, where the sign of mean curvatures undergoes a change. We label each edge pixel based on the segmentation labels of the surface points in a small neighborhood around the edge pixel, thus obtaining support from its local neighboring region. If there is enough support for the presence of surface points belonging to the saddle valley region in the neighborhood, then the edge pixel is labeled as crease edge. Otherwise, it is labeled as a jump edge.

We tested this edge detection method on several range images of real objects. We found that significant occluding and crease boundaries are reliably detected irrespective of the orientation of the object in the scene. While the jump boundaries would be detected by any normal edge detector as well, our method detects both types of edges in an integrated manner and guarantees the coherence between jump and crease edge maps.

4.3 Matching and Pose Estimation

We extend the approach of [16] to compare the edge maps containing the jump and crease edges obtained from the range images of different objects. The key idea behind this approach is the comparison of the model edge map with the scene edge map using the Hausdorff distance to find the transformation specifying the best alignment of the model in the scene. Two-dimensional geometric structures as represented by the edge maps are compared to measure the difference between the two shapes using the minimum Hausdorff distance under the transformation group G that includes translation in x and y directions. We extended this idea to include rotation about the z-axis as well.

Our matching scheme thus compares the edge maps obtained from the range images. Given a shape (its edge map) as a point set P and its transformed point set Q, under the action of some transformation group G, the distance between P and Q is the minimum difference between the shapes under all possible transformations of P with respect to Q denoted by

$$D_G(P,Q) = \min_{g \in G} H(g(P),Q)$$

where $H(g(P),Q)$ is given by

$$H(g(P),Q) = max(h(g(P),Q), h(Q,g(P)))$$

and

$$h(A,B) = \max_{p \in A} \min_{q \in B} \|p-q\|.$$

The Hausdorff distance, $H(g(P),Q)$, is a max-min distance for comparing sets. It measures the degree to which each point of the edge map P is closer to some point of Q and vice versa. It simply measures the proximity of points in the two edge maps and does not require a correspondence of the points in one edge map with points in the other. In order for the distance $D_G(P,Q)$ to be small, there must be a transformation in G that brings all of one object near some part of the other and vice versa. The method is implemented as a search strategy augmented with a heuristic to prune the transformation space wherein a possible "match" of the model to the scene image can be found.

4.4 Experimental Results

We have carried out several experiments with real range images of facial masks, to test our feature detection and matching schemes. The three-stage feature extraction method was employed to extract the edge maps containing jump and crease edges from these range images. We found that the initial smoothing using curvature consistency criterion resulted in denser and consistent principal curvature directions and this aided in obtaining consistent edge maps for the same object in different orientations.

A sample result using the range image of a face mask is shown in figure 5. Figure 5 shows all the processing steps of our feature extraction method. The jump and crease edges were extracted from the mean curvatures. The threshold required to obtain the edges using the Laplacian mask was set to upper 10 percentile. The smoothing effect of the curvature consistency criterion is evident from the figure 5(d). Figure 5(i) shows the salient features present in the image of the facial mask.

We present the results of our matching scheme here. Our database consisted of range images of different facial masks. Two types of experiments were carried out with the edge maps obtained from the range images of these masks. One set of experiments dealt with the estimation of the transformation parameters that would best align an edge map of an object with the edge map of the rotated and translated object. The second set involved the pairwise comparison of the edge maps to obtain the best matched edge maps in terms of minimum Hausdorff distance.

The edge maps used in the experiments are shown in figure 6. Both the jump and crease edges present in the edge map were used in matching. The allowed group of transformations included translations in x and y directions and rotation in the xy-plane. The forward distance from the model to the image $h(g(P),Q)$, where

$g(P)$ is the transformed model edge map and Q, the scene edge map was computed allowing 95% of the transformed model points to lie near the scene edge points. The values of g where the forward distance is smaller than a given threshold were the hypothesized transformations of the model image that would best align it with the scene image. These candidate transformations were then verified by determining whether the reverse distance $h(Q, g(P))$ was also smaller than the threshold. The best transformation was chosen to be the one with the smallest maximum of its forward and backward distances, among all the candidate transformations. Table 7 shows the pairwise minimum Hausdorff distances obtained when each of the edge maps was matched against the other.

Figure 8 shows the model image, the scene image, and the image containing the transformed model with the best transformation aligned with the scene edge map. Here 80% of the model points were required to match with the image during the forward and backward distance computations. The best transformation obtained is given by a translation of 8 pixels and 20 pixels in the x and y directions respectively and a rotation of 23 degrees about the z-axis in the clockwise direction.

5 Conclusions

This paper presented a three-stage approach for detecting the surface features in a robust manner. Our approach emphasized the preprocessing of sensor data before feature detection to assist in feature detection. We applied this approach to obtaining parts from range data and to detect of salient contours on free-form surfaces.

We have proposed a method for completing boundaries obliterated by the noise in the sensed data by integrating the edge map with the local minima of negative principal curvatures. The refined boundary structure facilitates decomposition of objects into their constituent parts. It appears that the proposed *local* scheme of part segmentation offers reasonable and reliable results. Our results are comparable with those reported in [17]. However, our method of boundary completion is less expensive since it is *local* in nature. Our results show that accuracy of the sensed data obtained from a White scanner and performance of curvature consistency algorithm can afford us to use a local boundary completion algorithm without compromising the reliability of the segmentation. It is well known that for a few objects, the principal of transversality cannot be exploited to extract their part structure from the data sensed from a certain orientation [4]. For such objects, our method of segmentation is an inexpensive method for deriving a good initial hypothesis for part segmentation algorithms based on either feedback strategy [8] or catalogs derived from the multi-view representations [13].

Robust feature extraction methods also lend themselves to composition of larger features from smaller features. Results of our matching method for geons show that the large matching primitives can indeed decrease matching complexity by exploiting significant global constraints. Consequently, these matching methods are more stable and robust due to their larger degree of tolerance to the missing data.

We also described a new method to extract jump and crease edges from mean curvatures and demonstrated the effectiveness of the method on various range images of arbitrary surfaces. Our method can detect both the jump and crease edges in an integrated manner and it is computationally inexpensive. The robust estimates of principal curvatures using curvature consistency criterion were crucial in obtaining stable edge maps from the depth data. From our experiments on real range images with varying levels of degradation owing to noise and laser shadows, we observe that the salient edges on the objects are detected reliably using our method.

We also extended the matching technique based on the minimum Hausdorff distance to allow rotation about the z-axis, and we tested this method using the edge maps obtained from our feature extraction method. Results of our experiments show that this method can be successfully used for both the estimation of transformation parameters that would best align two edge maps, and to match the edge maps obtained from different objects to find the best matched edge maps. The robustness of the edge maps was found to be a vital factor in reducing the false matches of objects in our database. Our experiments show that the different edge maps obtained from the range images of the same facial mask resulted in minimum Hausdorff distances consistently, thus reducing the number of mismatches between different facial masks.

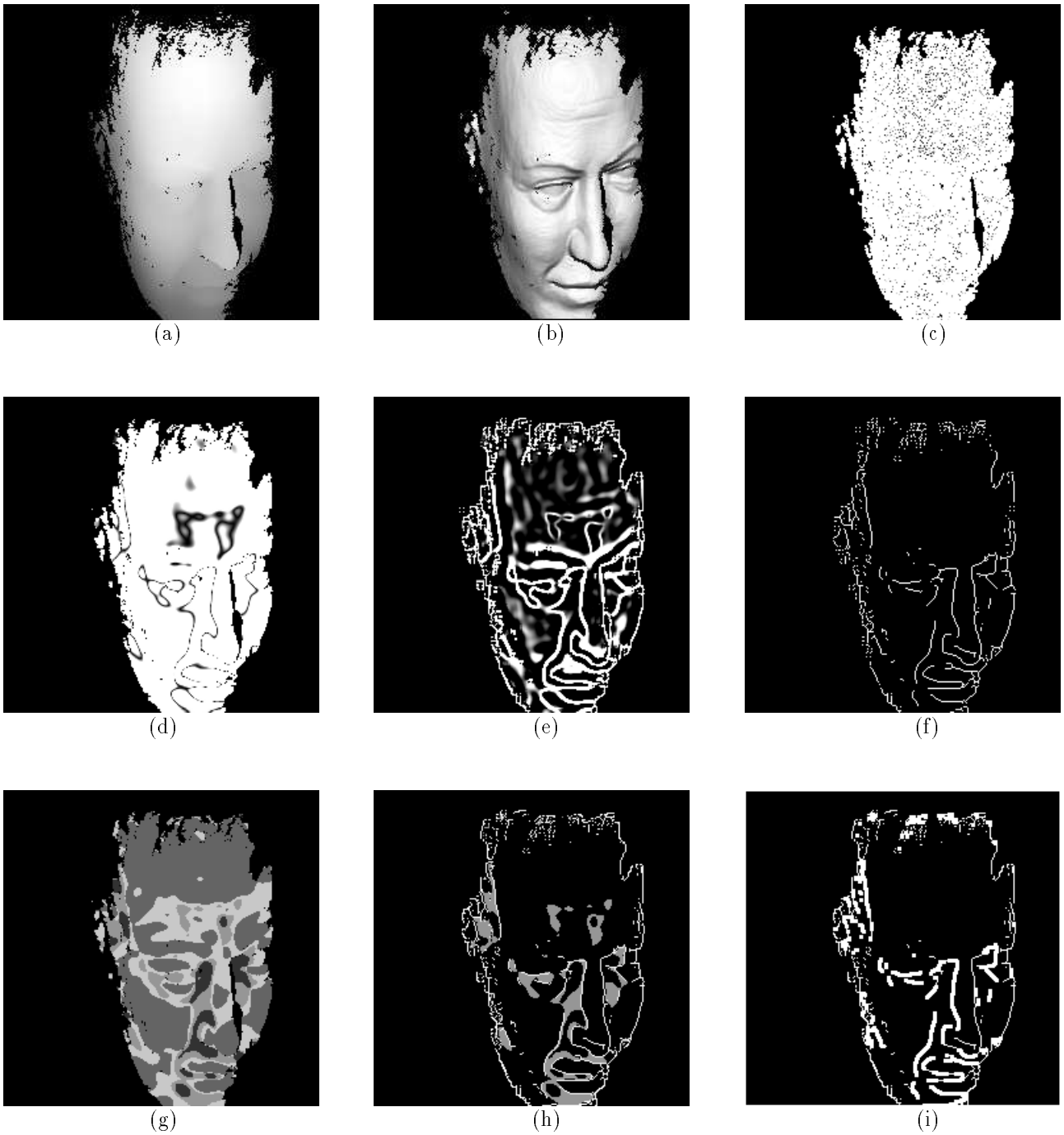

Figure 5: (a) Range image; (b) Depth as pseudo intensity; (c) Scaled mean curvature values before smoothing using curvature consistency; (d) Scaled mean curvature values after smoothing; (e) Output of an edge detector using the Laplacian mask; (f) Thresholded and thinned edge map; (g) Surface segmentation map using K-H sign; (h) Edge map and saddle valley regions; (i) The jump and crease edges, with crease edges shown as thick lines.

6 Acknowledgments

We thank Prof. Frank Ferrie, Gilbert Soucy, and Shailendra Mathur at McGill University for sharing their curvature consistency software with us. We are also grateful to Prof. Ruzena Bajcsy, Alok Gupta, Helen Anderson, Arup Mukherjee, and Luca Bogoni for access to superquadric estimation software developed at the University of Pennsylvania. This work was supported by the NSF research grants IRI-9103143 and CDA-8806599.

References

[1] M. Brady, J. Ponce, A. Yuille, and H. Asada, "Describing surfaces," *Computer Vision, Graphics and Image Processing:Image understanding*, vol. 32, pp. 1–28, 1985.

[2] F. Ferrie, S. Mathur, and G. Soucy, "Feature extraction for 3-D model building and object recognition," in *Handbook of Pattern Recognition and Image Processing* (T. Y. Young, Ed.), New York: Academic Press, 1993. To Appear.

[3] D. Cormo, *Differential geometry of curves and surfaces*. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1976.

[4] D. Hoffman and W. Richards, "Parts of recognition," *Cognition*, vol. 18, pp. 65–96, 1985.

[5] I. Biederman, "Human image understanding: Recent research and a theory," *Computer Vision, Graphics, and Image Processing*, vol. 32, pp. 29–73, 1985.

[6] N. S. Raja, *Obtaining Generic Parts From Range Images Using A Multi-View Representation*. PhD thesis, Michigan State University, E. Lansing, Michigan, 1992.

[7] S. J. Dickinson, A. P. Pentland, and A. Rosenfeld, "3D shape recovery using distributed aspect matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 174–198, February 1992.

[8] A. Gupta and R. Bazcsy, "Surface and volumetric segmentation of range images using biquadrics and superquadrics," in *The 11th IAPR International Conference on Pattern Recognition*, (The Hague), pp. 158–162, 1992.

[9] G. Stockman, G. Lee, and S. W. Chen, "Reconstructing line drawings from wings: The polygonal case," in *Proceedings of the Third IEEE International Conference On Computer Vision*, (Osaka, Japan), pp. 526–529, 1990.

[10] W. E. L. Grimson, *Object Recognition by Computer: The Role of Geometric Constraints*. Cambridge: MIT Press, 1990.

[11] A. K. Jain and S. G. Nadabar, "MRF Model-Based Segmentation of Range Images," in *Third IEEE International Conference on Computer Vision*, vol. 3, (Osaka, Japan), pp. 667–671, 1990.

[12] F. Solina and R. Bajcsy, "Recovery of parametric models from range images: The case for superquadrics with global deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 131–147, February 1990.

[13] N. S. Raja and A. K. Jain, "Recognizing geons from superquadrics fitted to range data," *Int. J. Image and Vision Computing, special issue on Range Image Understanding*, vol. 10, no. 3, pp. 179–190, 1992.

[14] P. J. Besl, "The free-form surface matching problem," in *Machine vision for three-dimensional scenes* (H. Freeman, Ed.), New York: Academic Press, Inc., 1990.

[15] A. Jain and P. Flynn, Eds., *3D Object Recognition*. New York: Elsevier, 1993.

[16] D. P. Huttenlocher and W. J. Rucklidge, "A multi-resolution technique for comparing images using the Hausdorff distance," tech. rep., Department of Computer Science, Cornell University, December 1992.

[17] F. Ferrie, J. Lagarde, and P. Whaite, "Darboux frames, snakes, and super-quadrics: Geometry from the bottom-up," in *IEEE Workshop on Interpretation of 3-D scenes*, (Austin, Texas), pp. 170–176, IEEE, 1989.

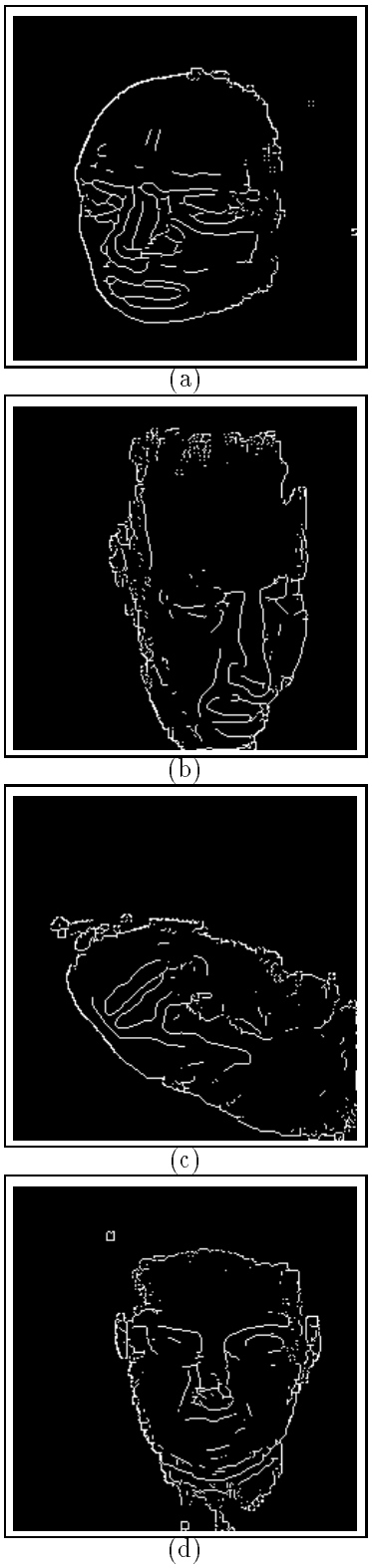

Figure 6: Test Images: (a) Image1; (b) Image2; (c) Image3; (d) Image4.

	Image1	Image2	Image3	Image4
Image1	2	20	20	16
Image2	-	3	8	14
Image3	-	-	3	14
Image4	-	-	-	1

Figure 7: Minimum Hausdorff distance between images.

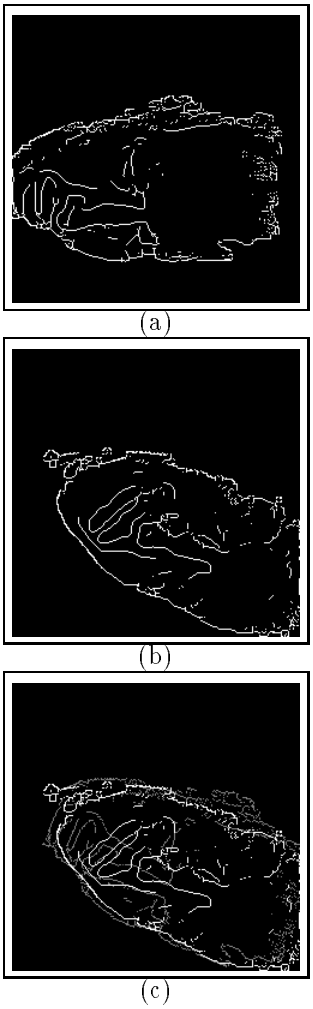

Figure 8: (a) Model edge map; (b) Scene edge map; (c) The aligned model edge map with the scene.