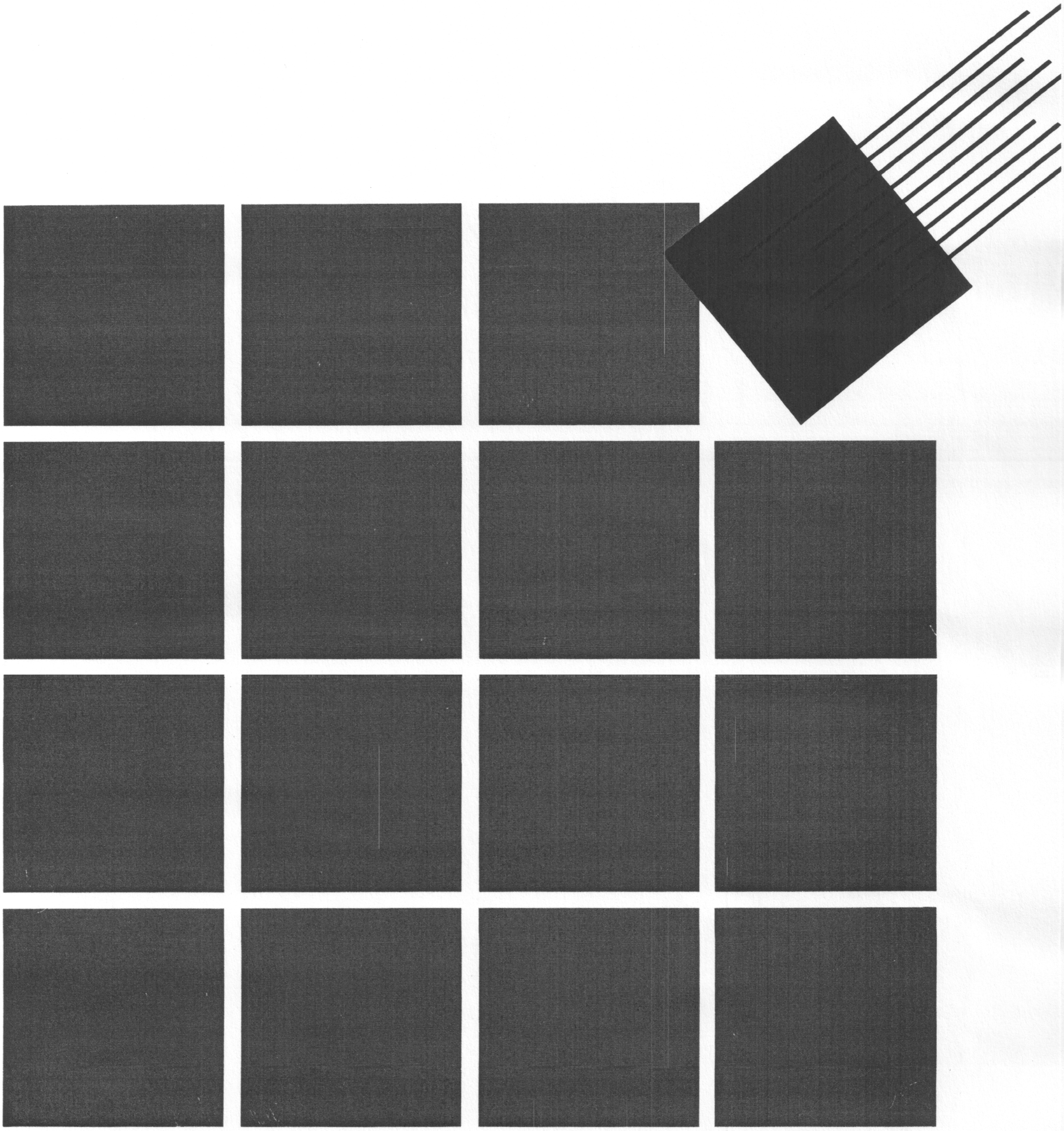
UNITED STATES POSTAL SERVICE

ADVANCED TECHNOLOGY CONFERENCE

OCTOBER 21-23, 1986



SESSION IV-A: WEDNESDAY, OCTOBER 22

Topic: Stamps, Inks and Cancellation
(Moderator: M. Sack)

Topic: Character and ZIP Code Recognition
(Moderator: T. Barnum)

Topic: Optical Processing for Address Block Location and Reading
(Moderator: F. Glickman)

SESSION IV-B: WEDNESDAY, OCTOBER 22

Topic: Range Image Interpretation
(Moderator: F. Glickman)

Modeling of Mail Pieces with Superquadrics

Franc Solina and Ruzena Bajcsy

General Robotics and Active Sensory Perception Laboratory
UNIVERSITY of PENNSYLVANIA
Department of Computer and Information Sciences
Philadelphia, PA 19104-6389

Abstract

Unlike standard 3-D object recognition, segmentation and classification of mail pieces cannot rely on a limited set of specific models. Although mail pieces can be classified by shape into parallelopipeds, cylinder-like objects etc., they do not conform exactly to these perfect geometrical shapes due to rounded edges, distorted corners, and bulging sides. With 3-D shape representations like generalized cylinders and polyhedral approximations such degradations from the ideal prototype are difficult to express. We propose to use superquadrics for modeling mail pieces because the roundness/squareness of edges can be directly expressed with two shape parameters. A recognition procedure using superquadrics is outlined.

Introduction

Postal services are currently facing the problem of automating mail piece handling. At present most of the mail pieces except letters are handled manually due to the large variability in size and shape.[1] Any automatic system for handling mail pieces has to determine location, orientation, size, and shape of mail pieces in order to manipulate them accordingly. Computer vision is a promising way to satisfy these requirements, which involve scene segmentation, shape description, and identification.

The problem of characterizing mail pieces is somewhere between scene description and object recognition. For scene description, a unique description of objects is not necessary. A bottom-up driven succession of representations that depends on the viewing direction and orientation of objects and ends up in polyhedral approximations or generalized cylinders is sufficient. On the other hand, to recognize an object in the scene as one from a set of predefined models, a computer vision system must have models of these objects in order to compare them with input data. Comparing object models with representations that depend on viewing direction and orientation of objects is difficult although they represent the same object. For recognition of 3-D objects, view point independent, 3-D models are required. The comparison of incoming data and the model must be done on the same level of abstraction. Since the data from the sensors can never give as good as or as complete information as contained in the model (i.e. due to self-occlusion), the comparison is done in an intermediate, symbolic description domain. Properties or features of objects can be used to make a hypothesis as to the which model corresponds to the object and also to verify that hypothesis.

Object recognition can be considered as a two part process; a low level or data driven part to build a description of objects, and a high level part where this description is compared to the models. Most working recognition systems rely on fixed, definitive models intended only for environments where a limited, preselected number of objects is encountered. Mail pieces, however, do not come just in a few uniform shapes and sizes. Having individual models for each mail piece is hence not feasible. This is why segmenting a scene and representing individual mail pieces is not object recognition in the strict sense which is selecting the right ready-made model from a predefined set.

Like any other objects, mail pieces can be grouped into classes or categories due to their properties or features, shape being the most important. It is important to note that object categories are not arbitrary sets of objects but reflect the structure of the world.[2, 3] The shape classification which is used for manual handling of mail pieces and which identifies parcels, flats, tubes, rolls, and irregular packages, reflects such structure. An automated mail handling system should also divide the mail pieces into appropriate classes, give the shape description of each piece by identifying the necessary parameters of the class model, and provide the position and orientation in a world coordinate system. Although category recognition is easy for humans to perform, machine vision is only beginning to work on it.[4, 5]

The next section describes modeling of mail pieces with superquadrics, followed by a section on the recognition process which uses the proposed model. Discussion at the end of the paper points out the advantages of using superquadrics for modeling mail pieces.

Models for mail pieces

As already noted, the difficulty in modeling mail pieces is their nonuniform shape and size. While different sizes of identically shaped objects can be represented with generalized cylinders by defining ranges of possible values of parameters like width, length, and diameter,[6] this still assumes models with perfect edges and corners. Although mail pieces can be classified according to shape into parcels, flats, tubes, rolls etc. they do not conform to perfect geometrical shapes like parallelopipeds or cylinders because of rounded edges, distorted corners, bulging sides, and wrinkled wrapping. These degradations from ideal shapes must be taken into account during recognition. With standard 3-D shape representations like generalized cylinders and polyhedral approximations such degradations from the ideal prototype are difficult to express. Superquadrics, on the other hand, have all the advantages of generalized cylinders, so that ranges of parameter values can be used to describe objects of different sizes. Two additional shape parameters are used to directly control the roundness/squareness of edges.

Superquadrics are a family of parametric shapes that extend the basic quadric surfaces and solids.[7] They have a mathematically concise definition and can model a large variety of "standard" building blocks like spheres, cylinders and prisms. The surface that they describe is everywhere derivable and deformations are easily defined on them.

A superquadric surface is defined by the following column vector:

$$\vec{x}(\eta, \omega) = \begin{bmatrix} a_1 \cos^{\varepsilon_1}(\eta) \cos^{\varepsilon_2}(\omega) \\ a_2 \cos^{\varepsilon_1}(\eta) \sin^{\varepsilon_2}(\omega) \\ a_3 \sin^{\varepsilon_1}(\eta) \end{bmatrix} \qquad \begin{array}{c} -\frac{\pi}{2} \leq \eta \leq \frac{\pi}{2} \\ -\pi \leq \omega < \pi \end{array} \qquad [1]$$

The parameters η and ω correspond to latitude and longitude angles of the 3-D vector $\vec{x}$ in spherical coordinates. The scale parameters a_1, a_2, a_3 define the size of superquadrics in directions x, y and z respectively. ε_1 is the squareness parameter along the z axis and ε_2 is the squareness parameter in the x-y plane. When ε_1 and ε_2 are both 1, the surface vector defines a sphere, a_1, a_2, and a_3 being equal. When ε_1 is < 1 and $\varepsilon_2 = 1$, cylinders result. Cuboid shapes are produced when both ε_1 and ε_2 are < 1. In the limit, for very small ε_1 and ε_2, a cube can be described. Flat beveled shapes are produced when either ε_1 or ε_2 is $= 2$. Equation 1 describes superquadrics in their standard position and orientation. They can be translated and rotated into a desired position in the world coordinate system with a homogeneous transform.[8]

Normal vector is available continuously over the surface of superquadrics, even across edges that are very square. Superquadrics define a continuous set of shapes. Nevertheless, a limited set of basic primitives can be defined by fixing the two shape parameters ε_1 and ε_2. All primitive solids with three-fold symmetry like spheres, cylinders and cuboids can be selected. A class of similar shapes can be defined by allowing the parameters to change in allowable limits. The solids can be furthermore easily modified by deformations like tapering, twisting and bending.[9]

Superquadrics are hence suitable as generic models for mail pieces. They can easily model rounded edges and corners just by setting different values of shape parameters ε_1 and ε_2. Since superquadrics are symmetrical, the parts of objects that are not in the field of view, or are not visible due to occlusion, are assumed to continue so that they form a symmetrical object.

Classification scheme. Classification of mail pieces is necessary because different shaped mail pieces require different handling. A classification scheme has to reflect the shape of mail pieces (like the one used for manual handling of mail) but can depend in part also on the nature of the automated manipulation (robot arms equipped with grippers or suction pumps, fixed automation). With superquadrics it is possible to represent distinct shape categories like parallelopipeds and cylinders of different sizes, and by

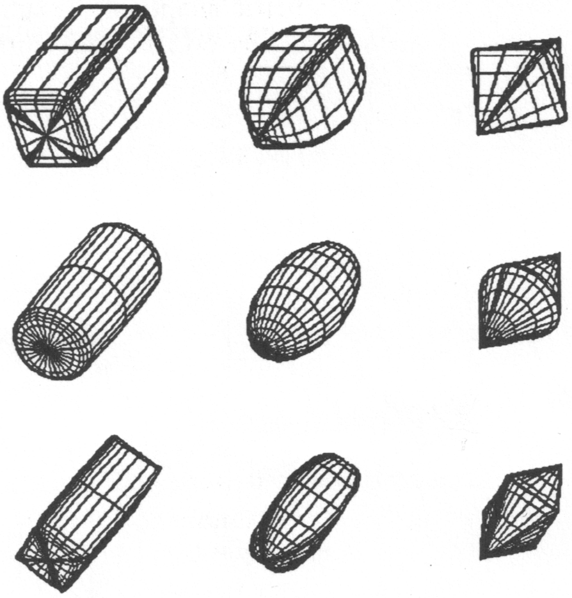


Figure 1 A set of superquadrics. The control parameters a_1 and a_2 are $= \frac{1}{2} a_3$. The shape parameters ε_1 and ε_2 change from 0.2 to 1 and 2 from left to right and from top to bottom of the figure respectively.

adjusting ε_1 and ε_2 the whole spectrum of shapes in between to model rounded edges and corners, or bulging sides. Global deformations of superquadrics like bending, twisting, and tapering are also possible, also these are more difficult to reconstruct from images since more parameters must be computed.

By presetting some superquadric parameters we propose the following geometric shape classification of mail pieces:

- thin rectangular parallelopipeds, where the thickness is very small or negligent compared to width and height for modeling letters and flats ($a_1 \ll a_2$, a_3 and ε_1, $\varepsilon_2 \ll 1$),
- rectangular parallelopipeds for modeling rectangular box-like packages (ε_1, $\varepsilon_2 \ll 1$),
- cylinders with circular or elliptical cross section for modeling tubes and rolls ($\varepsilon_1 \ll 1$ and $\varepsilon_2 = 1$),
- irregular objects as the class of objects that do not belong to any of the above categories ($\varepsilon_1 \approx 0.5$ and $\varepsilon_2 \approx 0.5$).

The goal of the vision system is to classify each mail piece into one of the above classes on the basis of specific features which will be discussed in the next section, and to compute the size parameters a_1, a_2 and a_3, together with the position and orientation of the mail piece in the world coordinate system. The initial values of shape parameters ε_1 and ε_2 which are preset for each mail piece class, can subsequently be fine tuned, if necessary, to better represent the shape of the mail piece. This process follows the so called *prototype and deformation* paradigm which is common in human perception for describing objects.[10]

The mail piece model has two levels of representation; a *prototype* that is recognizable due to specific features, and *parameters* that can change in a predefined range and have to be computed to conform the prototype to the specific mail piece. The following three figures show a letter (Figure 2), a box-like package (Figure 3), and a tube (Figure 4), each modeled by a superquadric. These figures were generated from reflectance images by hand picking the right model. Then size parameters (a_1, a_2, and a_3) and the position of the mail piece in the picture coordinates were computed.

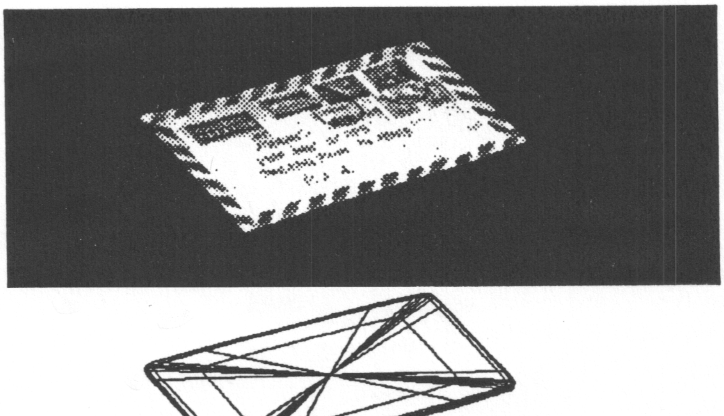


Figure 2 A reflectance image of a letter and the corresponding superquadric model.

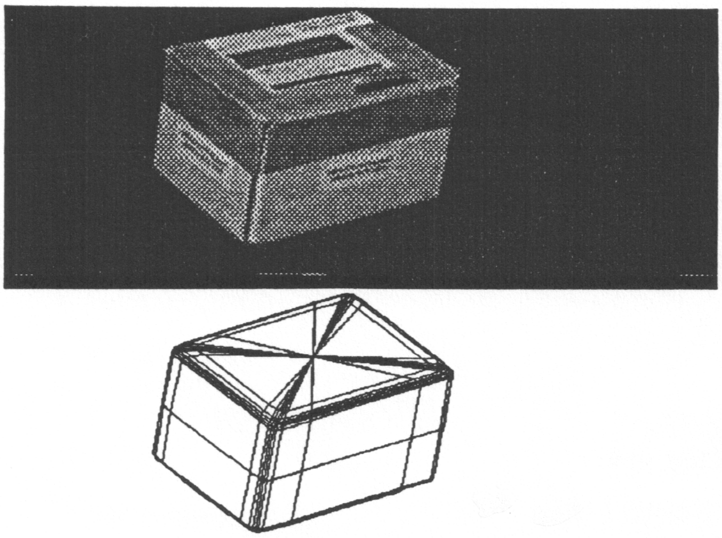


Figure 3 A reflectance image of a box-like package and the corresponding superquadric model.

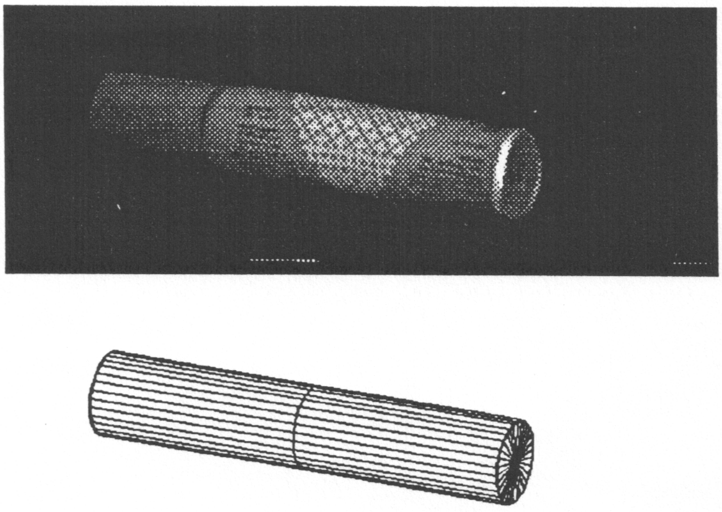


Figure 4 A reflectance image of a tube and the corresponding superquadric model.

Recognition procedure

The whole process of segmenting and describing mail pieces can be conveniently divided into a low level data driven, bottom-up part, that produces surface regions and edges; a model reasoning part that involves selecting the right model; and verification of the model.

The input to the segmentation and recognition procedure of unsorted mail pieces are calibrated range images which give accurate depth values for each point in the image. If the depth resolution of range images is not sufficient to segment overlapping thin mail pieces, additional gray-level reflectance images are required. Segmentation of reflectance images will start on the result of range segmentation. Regions in the reflectance image, corresponding to the ones identified in the range image, will be further segmented. Reflectance segmentation, on the other hand, extracts also labels, stamps and other texture on the surface of the mail which should not confuse object segmentation. The problem is ambiguous since even a human observer cannot always differentiate between a large label on a package or a post-card lying on top of it, without actually examining the content, i.e. identifying the address label. Reflectance segmentation can be done using standard techniques of region growing and edge finding. Reflectance segmentation will be used only as a supplement to close the gaps in range segmentation.

Bottom-up or data driven process starts with a range image and produces some intermediate description. It uses standard image formation models like edges, corners, regions, normals, and surfaces. These primitive models are combined where appropriate into features like parallel edges, mutually perpendicular planes, etc. A review of low level range image processing research[11] reveals that there are two principal approaches. One extracts edges, the other segments surfaces into planar surfaces or cylindrical surfaces. The ''edges first'' approach is successful when the objects have nice, clear edges. Mail pieces, however, have crumpled edges and beaten corners and this shape noise degrades the performance of edge finders. Crumpled paper of mail pieces can also mislead a region growing algorithm to subdivide a single face into a number of small surfaces. A combination of edge and surface extraction that works hand in hand seems to be the most reliable. Range segmentation is done by a combination of histograming and thresholding, and normal vector growing. First the supporting surface is removed. The rest of the image is segmented into planar and cylindrical surfaces by a region growing algorithm based on surface normals, while roof and occluding edges are extracted.

Model selection and verification. This section describes the selection of the model, computation of the model parameters that fit the model to the data, and verification of the model by directly comparing it to the data.

While it was not important for the data driven part, if one or more mail pieces is present in the image, the interpretation part would be more difficult. On the assumption that a single mail piece is present in the field of view, all extracted features are used for selecting the single required model. Features used here are edges and surfaces that result from the low level segmentation and their combinations.

If heaps of mail pieces can be present in the field of view then their presence must be detected first. If several isolated groups of features are detected, each group must be considered a potential heap. Features in each group are checked if they are consistent with a single object. Features of an isolated group that indicate a heap are: more than one local range maxima in the group image, concave corners in the reflectance image, concave edges in the range image, presence of parallel planes, and presence of cylindrical surfaces whose axis is not parallel. If any of the feature regions corresponds to a single object, the rule based indexing system selects the best model for it and tries to verify it. In the case of a heap, the strategy is to identify at least the top most mail pieces, these are also the easiest to pick up. The features in the group must be subdivided into regions that correspond to a single object. This is done mostly by merging those neighboring surface regions, resulting from the low level segmentation, which form convex edges and are topmost. For each object region a model will be selected. Although some object regions will correspond to partially occluded objects (by others in the heap or if they do not lay completely in the field of view) enough features might be present to select the right model. If a heap can not be segmented, its overall size and location can be represented by giving its maximum height and circumference. Such heaps can be broken up by mechanical means.

For each identified object region one of the four models is selected, based on the presence of features that result directly from low level segmentation and their combination. Features that point to a parallelopiped or box like object class are: two or three mutually perpendicular planes, plane raised from and parallel to the supporting surface, and a parallel pair of edges. Features that point to the cylinder like object class are: isolated cylindrical surface and/or circular or ellipsoidal edges. Features that indicate thin parallelopipeds are: a planar surface, pairs of parallel edges, and small differences in range from the supporting surface. Objects that do not correspond to any of the above are assigned the irregular mail piece model.

After the model selection, parameters of each object model must be computed. For this computation rough estimates of each model parameter (i.e. a_1, a_2, and a_3, while the first estimate of ε_1 and ε_2 are set for each model) and the position of the object in world coordinates (orientation and translation) are required. The spatial extent of the object can be estimated from the size of the object region. For estimating the position in world coordinates, the center of gravity and moments of inertia of the visible surfaces are computed. These initial values for position (orientation and translation) and model parameters are used as the input to a numerical procedure for modeling nonlinear systems (i.e. Levenberg-Marquardt method or one of its enhancements).[12] The numerical modeling procedure requires also the model (in this case the equation for superquadrics), and measurements of the system (these are x, y and z values of points on the surface of the mail piece). The results are parameter values of the selected model which best fits the data.

Verification of the model can be done by comparing the customized model to the range data. A measure of the goodness of fit is, for example, the sum of squares of differences between the range points and the model surface. This measure of fit among the model and the data also indicates whether the selection of the model is correct, or whether more accurate modeling is required by changing the shape parameters or employing global deformations.

Discussion

Automatic handling of mail pieces requires modeling of their shape. The use of superquadrics for this representation has the following advantages: the same basic representation can be used for all models just by taking different shape parameters ε_1 and ε_2, which enables uniform programming and handling of all processing. Several levels of modeling precision are possible. By fine tuning the epsilons, closer modeling of corrupt edges and bulging sides can be achieved. Global deformations like tapering and bending (this also involves estimating and computing the deformation parameters; i.e. linear tapering along an axis requires two additional parameters) can be employed if desirable. The cost of nonlinear modeling procedure used for computing the model parameters depends on their number and the accuracy of their first estimates. The number of parameters that have to be adjusted is a compromise between speed and modeling accuracy. For rough modeling, only the size parameters a_1, a_2, and a_3 have to be adjusted, while the shape parameters ε_1 and ε_2 are fixed for each mail piece category. This model can be too rigid, especially for irregular mail pieces. However, due to the exponential nature of the two shape parameters, the convergence of the modeling procedure when epsilons are involved as a variable, is slower. In the extreme, we could use just one superquadric model with variable shape parameters and perform the shape classification after the shape parameters are computed. For very accurate modeling of some mail pieces, global deformations would be a bonus if longer computation time for additional parameters is acceptable.

Superquadrics model the whole object, including hidden sides by assuming global symmetry. This enables easy computation of grasping positions and verification of consistency of object recognition in heaps by checking for intersections of superquadrics. One of the most important aspects of computer vision is whether the results of our interpretation can be verified. With superquadrics, the goodness of fit between the model and data can be easily checked.

The recognition process as outlined above can be parallelized in the following stages: first the initial low level processing can be done on a parallel architecture, where a processing element corresponds to each pixel. Later processing of all mail pieces (regions) in one image can be done in parallel.

Acknowledgements

This work was supported in part by NSF-CER/DCR82-19196 A02, NSF/DCR-8410771, Air Force/F49620-85-K-0018, and ARMY DAAG-29-84-K-0061.

References

1. Task Statement RFI-BL-0004, BOA Task 7, , U.S. Postal Service, Washington, DC (March 1986).
2. E. Rosch, ''Principles of Categorization,'' in *Cognition and categorization*, ed. E. Rosch and B. Lloyd,Erlbaum, Hillsdale, NJ (1978).
3. B. Tversky and K. Hemenway, ''Objects, Parts, and Categories,'' *Journal of Experimental Psychology: General* **113**(2) pp. 169-193 (June 1984).
4. S. A. Stansfield, ''Representation and control within an intelligent, active, multisensor system for object recognition,'' *Proc. 1985 Int. Conf. IEEE Systems, Man, and Cybernetics*, pp. 215-219 (1985).
5. F. Solina and R. Bajcsy , ''Shape and function,'' *Proceedings of SPIE's Cambridge Symposium on Optical and Optoelectronical Engineering*, (October 1986). To appear.
6. R. A. Brooks, ''Symbolic reasoning among 3-D models and 2-D images,'' Report No. STAN-CS-81-861, Department of computer science, Stanford university (1981).
7. A. H. Barr, ''Superquadrics and angle-preserving transformations,'' *IEEE Computer Graphics and Applications* **1** pp. 11-23 (1981).
8. R. Paul, *Robot Manipulators,* MIT Press, Cambridge, MA (1981).
9. A. H. Barr, ''Global and local deformations of solid primitives,'' *Computer Graphics* **18**(3) pp. 21-30. (1984).
10. A. P. Pentland, ''Perceptual Organization and the Representation of Natural Form,'' *Artificial Intelligence* **28**(3) pp. 293-331 (1986).
11. P. Besl and R. Jain, ''Range Image Understanding,'' *IEEE Proceedings on Computer Vision and Pattern Recognition*, (June 1985).
12. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vettering, *Numerical Recipies,* Cambridge University Press, Cambridge, England (1986).