Recognizing Volumetric Objects in the Presence of Uncertainty

Tal Arbel, Peter Whaite, and Frank Ferrie

TR-CIM-94-03 Feb. 1994

3D Vision Group
Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada

Submitted to *12th International Conf. on Pattern Recognition*, Feb. 1994

Postal Address: 3480 University Street, Montréal, Québec, Canada H3A 2A7
Telephone: (514) 398-6319 Telex: 05 268510 FAX: (514) 283-7897
Email: cim@cim.mcgill.ca

Recognizing Volumetric Objects in the Presence of Uncertainty

Tal Arbel, Peter Whaite, and Frank Ferrie

Abstract

This paper describes a new framework for parametric shape recognition based on a probabilistic model of inverse theory first introduced by Tarantola. The key result is a method for generating classifiers in the form of conditional probability densities for recognizing an unknown from a set of reference models. Our procedure is automatic. Off-line, it invokes an autonomous process to estimate reference model parameters and their statistics. On-line, during measurement, it combines these with apriori context-dependent information, as well as the parameters and statistics estimated for an unknown object, into a single description. That description, a conditional probability density function, represents the likelihood of correspondence between the unknown and a particular reference model.

The paper also describes the implementation of this procedure in a system for automatically generating and recognizing 3-D part-oriented models. Specifically we show that recognition performance is near perfect for cases in which complete surface information is accessible to the algorithm, and that it falls off gracefully (minimal false-positive response) when only partial information is available. This leads to the possibility of an *active* recognition strategy in which the belief measures associated with each classification can be used as feedback for the acquisition of further evidence as required.

1. Introduction

In this paper we describe a new framework for parametric shape recognition based on a probabilistic model of inverse theory first introduced by Tarantola in [12]. Application of this theory leads to a Bayesian recognition strategy similar to that used in other approaches [11]. However, the important distinction of our methodology is that it leads to a mechanism by which the conditional probability density functions used to classify shape models can be automatically generated. In doing so, important sources of contextual knowledge are taken into account that are less obvious in traditional approaches. Such knowledge includes i) a priori knowledge of the objects comprising the database, ii) information obtained from the process of estimating model parameters for an unknown object, and iii) information from the physical theories giving rise to the reference models themselves. The resulting density functions describe the likelihood of correspondence between an unknown model and a particular reference model. Such a measure is essential to an active recognition process which can use it as feedback in the collection of further data to resolve ambiguity.

The context of this paper is three-dimensional object recognition in which objects are represented by parametric shape descriptors such as superellipsoids [1, 2, 4, 8], deformable solids [3, 7], and algebraic curves [5, 11]. Object models are constructed through a process of *autonomous exploration* [13–17] in which a part-oriented, articulated description of an object is inferred through successive probes with a laser range-finding system. Figure 1a shows the set-up used to perform experiments — a two-axis laser range-finder mounted on the end-effector of an inverted PUMA-560 manipulator. For any particular viewpoint, such as the one shown in Figure 1b, a process of bottom-up shape analysis leads to an articulated model of the object's shape (Figure 1c) in which each part is represented by a superellipsoid primitive [4]. Associated with each primitive is a covariance matrix $\mathbf{C}$ which embeds the uncertainty of this representation and which can be used to plan subsequent gaze positions where additional data can be acquired to reduce this uncertainty further [14–16]. A system which automatically builds object models based on this principle is reported in [6, 13, 17].

Off-line, a database of object models is generated by presenting each object prototype to the model building system. Each object is in turn represented by several sets of parameters, one corresponding to each part of the object. For the experiments presented in this paper, objects are represented by a single parametric model that encompasses the entire object[1]. On-line, the recognition phase proceeds identically to model-building except for one key difference. On each iteration, i.e. gaze-point calculation → data acquisition → data merging (fusion) → parameter estimation, the conditional probability density function (CPDF) for each reference object given the current parameter estimate of the unknown object is calculated. If a clear winner stands out in terms of maximum likelihood, the process is terminated and a ranking of hypotheses along with their respective belief values returned. Otherwise the pro-

[1] The extension of our recognition strategy to multi-part objects is currently being investigated in our laboratory.

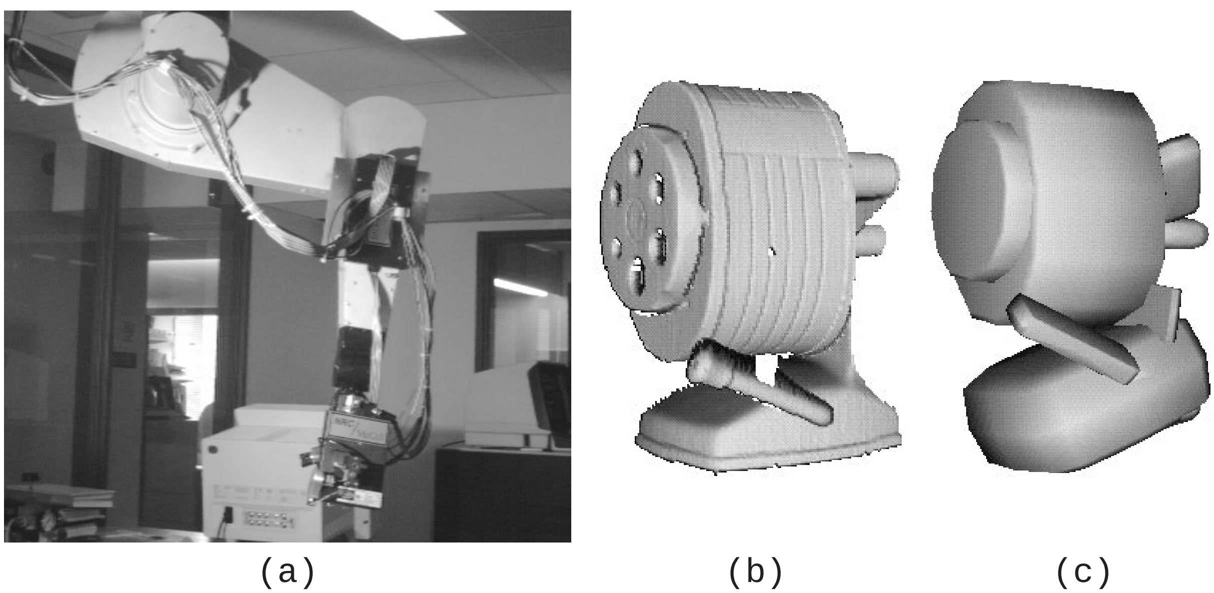


FIGURE 1. (a) Mobile laser range-finding system used to construct object models. (b) Laser range-finder image of a pencil sharpener rendered as a shaded image. (c) An articulated, part-oriented model of the sharpener using superellipsoid primitives; 8 superellipsoids are used, one corresponding to each of the parts of the object.

cess is allowed to continue and the CPDF's for each reference model updated on the basis of the newly acquired data. In this way, evidence can be incrementally gathered during the process of exploration.

The above recognition strategy raises a number of fundamental issues which are the focus of this paper. First, how is parametric uncertainty used and communicated between the processes of model building and recognition? Clearly they are not independent. Furthermore, the recognition process must take both the uncertainties in the database, as well as the measurement uncertainties of the unknown object, into account. By applying the Tarantola theory we show how the appropriate CPDF's used for recognition can be determined from such information. Second, what is the best manner in which to accumulate information? The model-building process is expensive, the merging of data from different viewpoints in particular [9, 10]. While this might be acceptable for database generation, recognition tasks must often be performed rapidly. An alternative is to consider the use of partial information obtained independently from different viewpoints. We show that the CPDF's generated from this data retain their selectivity and result in a minimum number of false-positive indications. In turn we show how the resulting ambiguities can be resolved without the need for data fusion through consensus from several different viewpoints.

Finally, one must consider the stability of model parameters. Representations based on superquadrics, for example, pose a number of problems due to degeneracies in shape and orientation. Other parametric forms, e.g. algebraic surfaces [5], are sometimes less problematic and can offer a more stable basis for recognition purposes. However it is still desirable to choose forms in which physical attributes can be ascribed to model parameters in an intuitive manner. The finite-element representations introduced by Pentland and his colleagues are a case in point [3, 7]. For our purposes, where shape is initially partitioned into part-oriented segments, superellipsoids are attractive both in the range of shapes they can represent as well as their computational simplicity. We have developed a method of avoiding degeneracies in the case of the superellipsoid, which permits the use of this convenient parametric form without incurring undue computational overhead. However, the discussion of this method is beyond the scope of this paper.

We begin in Section 2 with a brief overview of the inverse theory [12] and then formulate the problem of model recognition. This leads to a method of deriving, for each object model instance, the conditional probability of that model given the current estimated parameters of the unknown and their covariances. Two sets of experiments are presented in Section 3 which describe and compare the performance of the recognition procedure using CPDF's computed from complete and partial surface information respectively, as well as incremental recognition experiments. Finally we conclude in Section 4 with some general observations on our current work and points for future research.

2. The Inverse Problem Theory

The recognition problem requires us to infer from measurements of an unknown object that model which best represents it in a data base of known objects. Like all inverse problems, the recognition problem is ill posed in that, i) several models can give rise to identical measurements and, ii) experimental uncertainty gives rise to uncertain measurements. As a result it is not possible to identify the unknown object uniquely. There are various ways of conditioning ill posed problems, but these all require strong, and often implicit, a priori assumptions about the nature of the world. As a result a method may work well only in specific cases and because of the hidden implicit nature of the conditioning assumptions, cannot be easily modified to work elsewhere.

For this reason we have adopted the very general inverse problem theory of Tarantola [12]. It makes the sources of knowledge used to obtain inverse solutions explicit, so if conditioning is required the necessary assumptions about that knowledge are apparent and can be examined to see if they are realistic. Also, and importantly, the question of whether a solution is ill-posed or not is shown correctly to be an operational issue. The theory tells us how the knowledge we have can be combined to obtain a solution, but leaves any decision about its usefulness up to the tasks that require it. For example, when attempting to recognize objects we would ideally want the unknown model be identified correctly all the time. Because of experimental uncertainties this can never happen, and there is always the possibility that an object will be identified incorrectly. Only the task can know if the likelihood of errors is acceptable.

This raises the interesting question of what we should do if the level of errors is not acceptable. Because the sources of knowledge are explicit they are not only visible to the operational tasks, but are also potentially open to manipulation by them. In principal it should be possible for the task to condition or actively acquire the a priori knowledge required to make the solution acceptable. We have already demonstrated that what we call autonomous exploration functions well at the model building level [13, 17] and we now intend, with the aid of this theory, to incorporate feedback from the recognition task as well.

2.1. The Inverse Solution. The theory postulates that our knowledge about a set of parameters is described by a probability density function over the parameter space. This requires us to devise appropriate density functions in order to represent what we know about the world.

Tarantola's theory specifies three separate sources of a priori knowledge:

1. The knowledge given by a theory which describes the physical interaction between models $\mathbf{m}$ and measurements $\mathbf{d}$, denoted $\theta(\mathbf{d}, \mathbf{m})$.
2. Information we have about the model from measurements, denoted $\rho_D(\mathbf{d})$.
3. Information from unspecified sources about the kinds of models which exist in the world, denoted $\rho_M(\mathbf{m})$. Knowledge like this is a powerful constraint and can be used to eliminate many of the unconstrained solutions.

The solution to the inverse problem is in principal quite straight forward — it is simply a matter of combining the sources of information. The logical operation of *conjunction* is appropriate, i.e. the solution to the inverse problem is given by the theory AND the measurements AND any a priori information about the models. Tarantola extends the notion of logical conjunction to define the conjunction of two states of information [12, pages 29–31]. With this definition we can therefore combine the information from the joint prior probability density function $\rho(\mathbf{d},\mathbf{m})$ and the theoretical probability density function $\theta(\mathbf{d},\mathbf{m})$ to get the a posteriori state of information

$$\sigma(\mathbf{d},\mathbf{m}) = \frac{\rho(\mathbf{d},\mathbf{m})\,\theta(\mathbf{d},\mathbf{m})}{\mu(\mathbf{d},\mathbf{m})} \tag{1}$$

where $\theta(\mathbf{d},\mathbf{m}) = \theta(\mathbf{d}|\mathbf{m})\ \mu_M(\mathbf{m})$ and $\rho(\mathbf{d},\mathbf{m}) = \rho_D(\mathbf{d})\ \rho_M(\mathbf{m})$ over the joint space $M \times D$. The so called non-informative probability density $\mu(\mathbf{d},\mathbf{m}) = \mu_D(\mathbf{d})\mu_M(\mathbf{m})$ represents the reference state of information in much the same way that noise is used when measuring information in terms of signal to noise ratios. The formulation of appropriate non-informative densities is a complex issue, but for our purposes we will assume that all the non-informative densities are uniform over their respective spaces.

Accordingly, (1) is more general that the equations obtained through traditional approaches, but degenerates to them in specific cases. Under the conditions mentioned, the solution is identical to the Bayesian solution [12, page 61] where the a posteriori information about the model parameters is given by the marginal probability density function:

$$\sigma(\mathbf{m}) = \rho_M(\mathbf{m}) \int_D \frac{\rho_D(\mathbf{d})\ \theta(\mathbf{d}|\mathbf{m})}{\mu_D(\mathbf{d})}\ d\mathbf{d}. \tag{2}$$

2.2. The Part Recognition Problem. In the system we have constructed, range measurements are taken, surfaces are reconstructed then segmented into parts, and individual models are fit to each part. We will treat *the whole system as a measuring instrument.* Given some model $\mathbf{m}$ in the scene, range measurements are taken and from these an *estimate* of the model is obtained, $\mathbf{d}$, which we call a *measurement of the model* in the scene.

2.2.1. *Information Obtained from Physical Theories.* We first formulate an appropriate distribution to represent what is known about the physical theory that predicts estimates of the model parameters given a model in the scene. Such a theory is too difficult to formulate mathematically given the complications of our system. We therefore collect an empirical theory through a process called the *training* or learning stage of the recognition process. Here, monte carlo experiments are run on N measures of a known model exactly as in traditional statistical pattern classification methods. The conditional probability density function $\theta(\mathbf{d}|\mathbf{m})$ is calculated for each model by assuming a multivariate normal distribution. For each model class, a mean and covariance matrix are calculated. Therefore, the final equation for $\theta(\mathbf{d}|\mathbf{m})$ is:

$$\theta(\mathbf{d}|\mathbf{m}) = N(\mathbf{d} - \mathbf{g}(\mathbf{m}), \mathbf{C}_T) \quad (3)$$

where N is the multivariate normal distribution such that:

$$N(\mathbf{d} - \mathbf{g}(\mathbf{m}), \mathbf{C}_T) = \frac{1}{\sqrt{(2\pi)^n \ |\mathbf{C}_T|}} \exp\left(-\frac{1}{2}(\mathbf{d} - \mathbf{g}(\mathbf{m}))^T \mathbf{C}_T^{-1} (\mathbf{d} - \mathbf{g}(\mathbf{m}))\right) \quad (4)$$

and $\mathbf{C}_T$ is the covariance matrix describing estimated modelling errors for a model $\mathbf{m}$, and n is the dimension of the data space.

2.2.2. *Information Obtained from Measurements.* Much of the knowledge we have about a problem comes in the form of experimental measurements of observable parameters. In our system [15], we obtain an estimate of the observed model parameters $\mathbf{d}_{obs}$, and also an estimate of their uncertainty in the covariance operator $\mathbf{C}_d$. The assumption we make is that the multivariate normal distribution $N(\mathbf{d} - \mathbf{d}_{obs}, \mathbf{C}_d)$ represents our belief in the measurements. The probability density function representing this information is the conditional probability density function $\nu(\mathbf{d}_{obs}|\mathbf{d})$, such that:

$$\nu(\mathbf{d}_{obs}|\mathbf{d}) = \rho_D(\mathbf{d})/\mu_D(\mathbf{d}) = N(\mathbf{d} - \mathbf{d}_{obs}, \mathbf{C}_d) \quad (5)$$

2.2.3. *A Priori Information on Model Parameters.* In the current context, there are a discrete number of reference models, $\mathbf{m}_i, i = 1 \ldots M$. The probability density function used to convey this knowledge is

$$\rho_M(\mathbf{m}) = \sum_i P(\mathbf{m}_i)\ \delta(\mathbf{m} - \mathbf{m}_i), \quad (6)$$

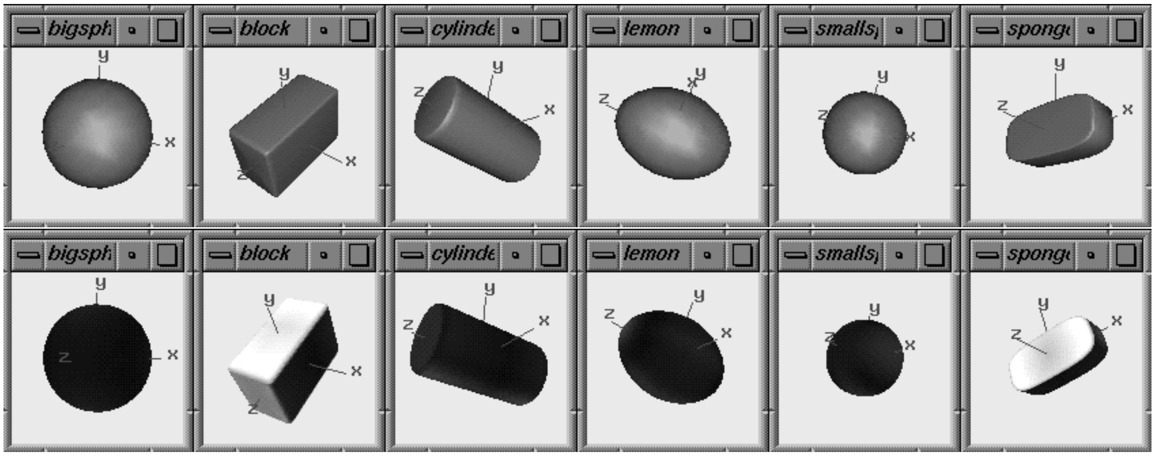
where $P(\mathbf{m}_i)$ is the a priori probability that the i^{th} model occurs.

2.2.4. *Solution to the Inverse Problem.* Substituting the probability density functions (3), (5), and (6) into (2) gives us the final equation for the a posteriori probability density function

$$\sigma(\mathbf{m}) = \sum_i P(\mathbf{m}_i) N(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_i), \mathbf{C}_D)\ \delta(\mathbf{m} - \mathbf{m}_i). \quad (7)$$

where $\mathbf{C}_D = \mathbf{C}_d + \mathbf{C}_T$. This density function is comprised of one delta function for each model in the database. Each delta function is weighted by the *belief* $P(\mathbf{m}_i)N(\mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_i), \mathbf{C}_D)$ in the model $\mathbf{m}_i$, and is essentially calculated by convolving the normal distributions in (3) and (5). The advantage of the method is that rather than establish a final decision as to the exact identity of the unidentified object, it communicates the degree of confidence in assigning the object to each of the model classes. It is then up to the interpreter to decide what may be inferred from the resulting distribution.

Representations based on superquadrics pose a number of problems due to degeneracies in shape and orientation. In solving this problem, work has being done in



Displayed above are the reference objects that result from training on complete surface data: a big sphere, a block, a cylinder, a lemon, a smaller sphere, and a rounded block. Below, the same models are shaded according to the projection of parameter uncertainties into 3-D space. White reflects large uncertainties, and black indicates parameters that are tightly constrained. For example, the light face of the block shows that the y size parameter is more uncertain than the x.

FIGURE 2. Six representatives that result from training

representing objects by multi-modal distributions, where each mode contains information about a possible equivalent form. Discussion of this process is beyond the scope of this paper.

3. EXPERIMENTS

Six objects were chosen for the purposes of testing the recognition procedure. These objects included a small wooden sphere (rad = $20mm$), a slightly larger wooden sphere (rad = $25mm$), a wooden block, a wooden cylinder, a plastic lemon, and a rounded block made of a sponge-like material. The objects were selected because they consisted of single parts that conformed well to superellipsoids. They were relatively symmetric in all three planes, did not bend or taper, and were man-made. They varied in size and shape, so as not to be clustered together too tightly in five-dimensional feature space. However, their distributions overlapped sufficiently enough in several dimensions so that the recognition procedure was challenged in its discrimination task.

Training (Section 2.2.1) automatically produced object class representatives, by measuring the object numerous times. Each individual model was created by scanning the object from several views using a laser range-finder[2], then a superellipsoid model was fit to the data, and the resulting parameters stored. For the purposes of creating a stable database for recognition, it was established that three views of each object, 120° apart were sufficient to constrain the fitting procedure. Each sample was scanned

[2]The density of scanning was such that each pixel of an $85 \times 85 pixel^2$ image represented $3mm^2$.

from a random scanning position, producing 24 samples of each object. Figure 2 illustrates the six representative models of each object that result from training.

3.1. Matching Using Complete Information. In the first experiment, recognition was performed using an unknown model computed from a sequence of views covering the visible surfaces of an unknown object. The intent of this experiment was to validate the recognition procedure against models produced by the autonomous exploration process on running to completion[14]. Twenty-four samples of each object, each scanned from three different viewpoints, were presented to test the invariance of recognition against variations in sampling and viewpoint. Using maximum likelihood as the basis for recognition, i.e. choosing the model with the highest confidence value, the results shown in Table 1 are obtained.

Models	Big sphere	Block	Cylinder	Lemon	Small sphere	Rnd.Block
# correct	24	24	24	24	24	24
# incorrect	0	0	0	0	0	0
# undeterm.	0	0	0	0	0	0

Table 1. Matching samples taken from multiple viewpoints

The results indicate that the system can successfully recognize an instance of any object in the database with perfect results, provided that its surfaces are accessible, independently of viewpoint and sampling order. This is to be expected given that the probability density functions of each of the unidentified objects exhibit small variations in parameter space due to the relatively complete information available. Training produces reference models that also have narrow distributions, and which are well separated from each other. The distribution of the unidentified object would necessarily overlap that of the correct reference model more than the others.

Examination of the resultant beliefs shows that the recognition system is certain about the reference model it chooses in all cases. Examples of the non-normalized belief distributions of the lemon and block can be found in Table 2. These results indicate that complete information allows the system to correctly identify objects with a high degree of certainty. The high beliefs reflect the fact that both the measurement distributions and the reference model distributions are "delta-like" and close together.

3.2. Matching Using Partial Information. Since complete information is not always available (and potentially expensive to acquire), a more realistic test would be to determine the parameters of an unknown model from partial information. In the limit this would consist of attempting to base recognition on data acquired from a single viewpoint and would clearly violate the assumptions implicit in the training process. Furthermore, it has been shown elsewhere that the resulting model parameters would be inherently less stable [14]. However, should the procedure still retain some of its earlier selectivity — as evidenced by a low degree of false positive matches — then an incremental procedure becomes a possibility. In this second set of experiments, recognition was performed on thirty-six single-view samples of each object.

Trial	Big sphere	Block	Cylinder	Lemon	Small sphere	Rnd Block
1	0	0	0	5.11	0	0
2	0	0	0	6.53	0	0
3	0	0	0	12.66	0	0
4	0	0	0	70.70	0	0
5	0	0	0	42.32	0	0
6	0	0	0	27.13	0	0

a) Belief distributions of the lemon

Trial	Big sphere	Block	Cylinder	Lemon	Small sphere	Rnd Block
1	0	6.09	0	0	0	0
2	0	6.24	0	0	0	0
3	0	9.87	0	0	0	0
4	0	1.58	0	0	0	0
5	0	15.21	0	0	0	0
6	0	11.67	0	0	0	0

b) Belief distributions of the block

TABLE 2. Results of several iterations of recognition of a)lemon and b)block viewed from multiple viewpoints

Here, data was collected at 40° intervals in 4 different equatorial planes. The same methodology as in the first experiment was applied in the recognition of the unknown model parameters. The results obtained are shown in Table 3.

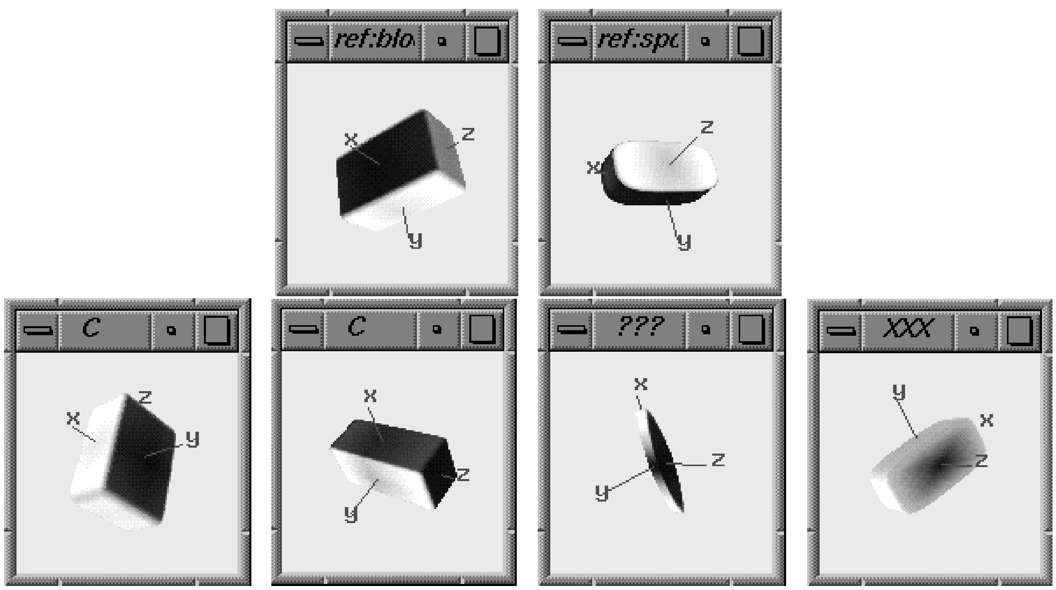
Models	Big sphere	Block	Cylinder	Lemon	Small sphere	Rnd.Block
# correct	36	26	33	36	36	19
# incorrect	0	3	0	0	0	0
# undeterm	0	7	3	0	0	17

TABLE 3. Matching samples taken from a single viewpoint

As expected, recognition based on partial information is less certain than in the previous case where the complete surfaces of the unknown object were accessible. Here, undetermined states exist in situations where the values of the a posteriori probability density functions are extremely low (on the order of 10^{-30}). Due to numerical underflow, the procedure produces beliefs of zero for each of the reference models. This situation occurs when the parameters determined for a particular viewpoint differ significantly from any of the models in the database. We call such viewpoints *uninformative*. Here, the large covariances that result produce wide distributions that do not sufficiently overlap any of the reference distributions. In three cases (the wooden block) false-positive recognition did occur. This happened when the block was scanned face-on, leaving it unconstrained in several directions such that the resulting parameters were closer to the rounded block than to any other reference object.

Figure 3 shows some specific examples of recognition attempts of the block from different viewing positions. In the first two cases, the procedure correctly identified the objects as corresponding to the block despite wide fluctuations in their size parameters. This is due to the fact that the models encompass the uncertainties corresponding to these parameters in their representations. The reference model also learned of these possible variations during training, incorporating them it its representation. Therefore, the distributions were close enough to that of the reference block to make a correct identification. This reinforces the hypothesis that objects need not be represented by extremely accurate descriptions. Rough size and shape representations are sufficient as long as the reference object has learned about these possible fluctuations in the training stage. In the third case, the system could not identify the object as being any of the known models. One can see that this model does not visually resemble any of the references in size or shape. In the final case, the system incorrectly identified the block as being the rounded block. Here, one can see that the model is visually closer to the rounded block. Despite permitting fluctuations in size, the reference block is quite certain about its shape parameters, and does not permit rounded edges. This is indicated by the black shading around the block reference model's edges.

Table 4 shows the belief distributions resulting from incremental attempts at recognizing the lemon and the block. Here, data is collected from single views at 40°



In the top boxes are the square block and rounded block reference models. Below these are four different attempts at recognizing the square block from different viewing positions. In each case the model is compared to the each of the six references in turn, and beliefs in each are computed. Above each model one can see the result of running a maximum likelihood algorithm on the results. *C* indicates a correct recognition, *???* indicates an undetermined state, and *XXX* refers to a false recognition. Here, the system identifies the square block as being the rounded one. The objects are shaded according to their uncertainties (see figure 2).

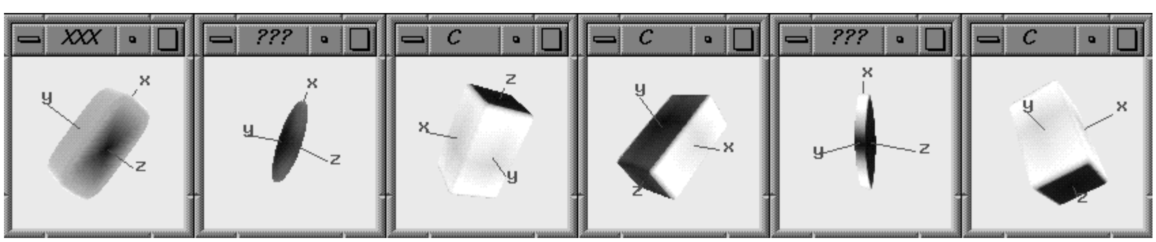
FIGURE 3. Examples of recognition of the block from single views

Viewpoint	Big sphere	Block	Cylinder	Lemon	Small sphere	Rnd Block
0°	0	0	0	2.97×10^{-21}	0	0
40°	0	0	0	6.93×10^{-15}	0	0
80°	0	0	0	0.18	0	0
120°	0	0	0	2.44×10^{-5}	0	0
160°	0	0	0	8.07×10^{-3}	0	0
200°	0	0	0	3.38×10^{-4}	0	0
240°	0	0	0	1.10×10^{-16}	0	0
280°	0	0	0	0.31	0	0

a) Belief distributions of the lemon

Viewpoint	Big sphere	Block	Cylinder	Lemon	Small sphere	Rnd Block
0°	0	4.00×10^{-13}	0	0	0	1.16×10^{-5}
40°	0	0	0	0	0	0
80°	0	0.33	0	0	0	0
120°	0	0.05	0	0	0	0
160°	0	0	0	0	0	0
200°	0	0.21	0	0	0	0
240°	0	0	0	0	0	0
280°	0	0.05	0	0	0	0

b) Belief distributions of the block



Displayed above are the first six attempts at successively recognizing the block at 40° increments. Shading is in accordance with parameter uncertainties (see figure 2). The results of running a maximum likelihood algorithm are found above each box (see figure 3).

TABLE 4. Results of incremental recognition of a)lemon and b)block viewed from 40° single viewpoints

intervals in an equatorial plane. One can see that the beliefs are considerably weaker than in the previous case consisting of complete information. The first iteration in the recognition of the block produced a false-positive identification. In this case, the resulting distribution did overlap with the distribution of the reference block. The belief in the rounded block is quite low, indicating that the system is quite uncertain about the identification. In fact, whenever a false-positive identification occurs, the system produces very low beliefs. This suggests that if the threshold for undetermined states were raised, the incorrect identifications would become undetermined states.

The preceding suggests the possibility of an incremental recognition procedure. It is based on the following observations obtained empirically over successive trials:

i) Uninformative views generally result in undetermined states. Detection of such events is a clear indicator that further sampling is required.
ii) Informative views are generally accompanied by high confidences (beliefs), but with the possibility of a false-positive indication.
iii) The likelihood of successive false-positive indications is very small. First, this is a consequence of the high selectivity of the reference distributions which result in low frequencies of false-positive indications in the first place (e.g. Table 3). Second, it is unusual for observer motion to result in similar viewpoints in two successive views (general position assumption).

In the longer version of this paper we argue that as a general rule, three consistent identifications suffice for a stable scene. An example is shown in Table 4 which shows a sequence corresponding to the first 6 entries in the second half of Table 4. Iteration 1 is inconclusive, the object is either a square or rounded block (However the results of running a maximum likelihood algorithm indicate that the object is a rounded block). In iterations 2 and 5 the object is undetermined. Iterations 3, 4, and 6, on the other hand, consistently support the correct classification of the unknown object as the square block.

4. Conclusions

In this paper we have presented a new framework for parametric shape recognition based on a probabilistic model of inverse theory introduced by Tarantola [12]. We have shown how a Bayesian recognition strategy can be derived automatically by applying the theory and have demonstrated its implementation in a system for recognizing 3-D objects based on superellipsoid parameters.

The results indicate that the strategy is quite robust, not only in situations where complete surface information is available but also in those cases where it is only partially accessible. This leads to the possibility of an incremental recognition strategy where the beliefs associated with each different hypothesis can be used to control further acquisition of data until a prescribed degree of certainty is met. Our results demonstrate that it is indeed possible to differentiate between informative and uninformative viewpoints, a key requirement in such a mechanism. We have shown that the case of false-positive indications can be effectively dealt with by insisting on a consensus between three different viewpoints in affirming a particular hypothesis. Simply put, the same maximum likelihood match would have to be present from at least three different viewpoints before a label is assigned. An incremental, *active* recognition strategy based on these ideas has proved successful in our laboratory.

Some observations are in order regarding the autonomous explorer, the system used to automatically generate the database models used for recognition. In the numerous trials performed during the course of this research we were able to consistently obtain stable parametric descriptions of the model database. These were largely independent of viewpoint, variations in sampling, and the trajectory chosen by the mobile laser scanner. The generation of stable, salient object models is clearly an essential ingredient in the implementation of a successful object recognition system. Future work will involve exploration guided by feedback from the recognition system. This is possible because all sources of knowledge are made explicit within the framework described. Therefore, the system could actively acquire information needed to correctly classify the objects.

The system described exhibits a high degree of selectivity in matching object primitives, paving the way for recognition of more complex objects. We are currently developing a scheme for multiple-part object recognition that involves a graph-matching procedure. We believe that this paper outlines a sound, statistical method for comparing the nodes. Given its success in discriminating based on partial information, the search-space for the graph-matching problem should be considerably reduced.

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. T. Darrell, S. Sclaroff, and A. Pentland. Segmentation by minimal description. In *Proceedings, 3RD International Conference on Computer Vision*, pages 112–116, Osaka,Japan, Dec. 1990. Computer Society of the IEEE, IEEE Computer Society Press.
4. F. P. Ferrie, J. Lagarde, and P. Whaite. Darboux frames, snakes, and super-quadrics: Geometry from the bottom up. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(8):771–784, Aug. 1993.
5. D. Keren, D. Cooper, and J. Subrahmonia. Describing complicated objects by implicit polynomials. Technical Report 102, Brown University LEMS, Laboratory for Engineering Man/Macine Systems, Division of Engineering, Brown University, Providence RI 021912 USA, 1992.
6. A. Lejeune and F. P. Ferrie. Partitioning range images using curvature and scale. In *PROC. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York City, New York, June 15-17 1993. to appear.
7. A. Pentland and S. Sclaroff. Closed form solutions for physically based shape modelling and recognition. volume 13(7), pages 715–729, July 1991.

8. N. S. Raja and A. K. Jain. Recognizing geons from superquadrics fitted to range data. *Image and Vision Computing*, April 1992.
9. G. Soucy. View correspondence using curvature and motion consistency. Master's thesis, Dept. of E.E., McGill Univ., 1992.
10. G. Soucy and F. P. Ferrie. Motion and surface recovery using curvature and motion consistency. In *PROC. Second European Conference on Computer Vision*, pages 222–226, Santa Margheri-ta Ligure, Italy, May 18-23 1992.
11. J. Subrahmonia, D. B. Cooper, and D. Keren. Practical reliable bayesian recognition of 2D and 3D objects using implicit polynomials and algebraic invariants. LEMS 107, Brown University LEMS, Laboratory fo Engineering Man/Machine systems, Division of Engineering, Brown University, Providence, RI 02912, USA, 1992.
12. A. Tarantola. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation.* Elsevier Science Publishing Company Inc., 52, Vanderbuilt Avenue, NewYork, NY 10017, U.S.A., 1987.
13. P. Whaite and F. Ferrie. Model building and autonomous exploration. In *SPIE - Intelligent Robots and Computer Vision XII: Active Vision and 3D Methods*, pages 73–85, Boston, Massachusetts, Sept.8–9 1993.
14. P. Whaite and F. P. Ferrie. From uncertainty to visual exploration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10):1038–1049, Oct. 1991.
15. P. Whaite and F. P. Ferrie. Uncertain views. In *PROC. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 3–9, Champaign, Illinois, June 15-18 1992.
16. P. Whaite and F. P. Ferrie. Active exploration: Knowing when we're wrong. In *PROC. Fourth International Conference on Computer Vision*, pages 41–48, Berlin, Germany, May 11-14 1993. Computer Society of the IEEE, IEEE Computer Society Press.
17. P. Whaite and F. P. Ferrie. Autonomous exploration: Driven by uncertainty. In *Proceedings, Conference on Computer Vision and Pattern Recognition*, Seattle, June 1994. Computer Society of the IEEE, IEEE Computer Society Press. To appear.