Probabilistic 3D Object Recognition

Ilan Shimshoni and Jean Ponce
Department of Computer Science and Beckman Institute
University of Illinois, Urbana, IL 61801
e-mail: ilan@cs.uiuc.edu ponce@cs.uiuc.edu

Submitted to ICCV 1995

Abstract

We present a probabilistic 3D object recognition algorithm which uses the probabilistic peaking effect of measured angles and ratios of lengths to guide the recognition process. Isoratio and isoangle curves are traced on the viewing sphere and are used to compute conditional probabilities. These probabilities are incorporated into a probabilistic model which deals with various types of uncertainty associated with the recognition process, such as uncertainty due to imperfect results of edge and corner detection (e.g. detecting edges which belong to the background, not detecting edges which belong to the object, and uncertainty in edge and corner position). We perform pose estimation and compute the uncertainty in pose due to the uncertainty in measurements in the image. The uncertainty pose subspaces are used to find hypotheses which reinforce each other. We compute the probability that these sets of hypotheses are correct and use these probabilities to rank the matching hypotheses.

1 Introduction

One of the major problems in computer vision is recognizing 3D objects that appear in a scene from a database of objects (see [5, 8] for reviews). Many recognition systems extract features such as points [14, 17], or lines [3] from the image and try to match them to corresponding features in the database. They attempt to verify each candidate match using other features in the image and then rank the verified hypotheses. When constructing such a system the major problems that have to be solved are: in which order to process the initial matches, how to verify these matches using additional image features, and how to rank the verified hypotheses. In this paper we match pairs of adjacent edges (Fig 1(a)) and trihedral corners (Fig 1(b)) to corresponding feature sets in the model database. Combining these two types of feature sets into a single recognition algorithm is one of the contributions of this paper. The initial matches are ranked by computing the probability that the match is correct. In order to verify the initial match, we compute the pose of the object assuming the match is correct and estimate the pose uncertainty due to the uncertainty in the values measured in the image. Other matches whose pose uncertainty subspaces intersect the subspace of the initial match contribute to the evidence that the initial match was correct. Sets of features contributing to the same match are ranked by a probabilistic expression which is the probability that the match of the whole set is correct.

In order to rank initial matches we have to compute the probability that a certain feature set in the image matches a feature set in the model database taking into account that the feature extraction process is not perfect and the uncertainty in the values measured in the image. When viewing objects from different directions object features seem different. The value of angles and ratios of lengths measured in the image depend on the viewing direction. The probabilistic analysis of this phenomenon is known as the probabilistic peaking effect [4, 6, 7]. The analysis shows that although viewed angles and ratios change with the viewing direction, with very high probability values measured in images will be between double to one half the original value measured in the object, with the maximum of the probability density function obtained where the measured value is equal to the original value. Therefore assuming weak perspective projection and that viewing directions are uniformly distributed on the viewing sphere, we can compute the probability that a certain image feature set matches feature sets from the model database and test these initial matches in decreasing probabilities. In this paper we develop a new method to accurately compute the probabilities taking into account occlusion of the features of the object by other features using the aspect graph of the object. In previous work [4, 6, 7, 17], such probabilities were estimated by computing the ratios and angles for points uniformly distributed on the viewing sphere. In this paper we trace curves on the viewing sphere where the ratio (isoratio) or the angle (isoangle) is constant and use the areas bounded by these curves to accurately compute the desired probabilities. These probabilities are incorporated into a probabilistic model which takes into account that the feature extraction stage is suboptimal which results in certain features not being extracted and features which do not belong to the objects in the database but to the background being extracted in that stage.

In order to verify an initial match we compute the pose of the object assuming that the match is correct and try to find other matches of other features of that object that have close poses which reinforces the initial match. We compute the viewing direction component of the pose of the object by tracing the isoratio or isoangle angle curves on the viewing sphere which correspond to ratios and angles which we measure in the image, and the viewing direction is the intersection point of those curves. The other components of the pose are easily obtained where possible. This technique works for both types of feature sets we deal with where as previous work developed separate techniques for two adjacent edges [1, 14] or trihedral corners [18]. In order to deal with the uncertainty in the values measured in the image which causes uncertainty in the pose [2], we compute pose uncertainty subspaces for each match, and only matches whose subspaces intersect the subspace of the initial match

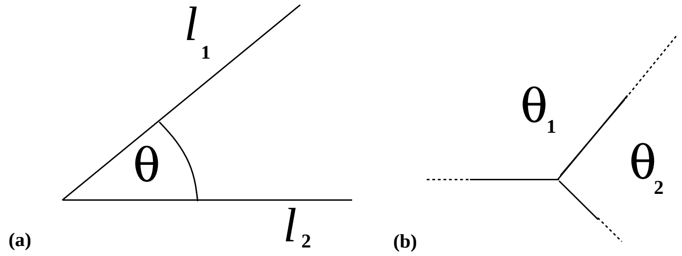


Figure 1: Image feature sets: (a)Two lines in an image l_1 and l_2. The angle between the lines is θ (b) A corner with three edges emanating from it.

reinforce the match. Our work extends [12] which dealt with the 2D model to the weak perspective 3D model.

In the final stage of the algorithm we developed a probabilistic expression to rank matching hypotheses. We take into account the number of matched features and the uncertainty in the values measured in the image to produce the probability that the hypothesis is correct.

The rest of the paper is organized as follows. In Section 2 we develop a method for computing probability density functions using equations we develop for angles and ratios. In Sect. 3 we discuss pose estimation and estimate the effects of uncertainty on the pose. In Sect. 4 we present our 3D object recognition algorithm and our probabilistic expression to rank matching hypotheses. Finally, a number of issues raised by our algorithm and future work are discussed in Sect. 5.

2 Probability Density Function Computation

2.1 Development of Curve equations

When measuring the ratio of line lengths in images or the angle between lines we want to know from what viewing directions could such an image be generated assuming a matching between the lines and certain edges in the model. Given lines l_1 and l_2 in the image (Fig. 1(a)) whose ratio of lengths is $\rho = |l_1|/|l_2|$. l_1 and l_2 match $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively in the model. For viewing directions $\mathbf{v}$ which satisfy:

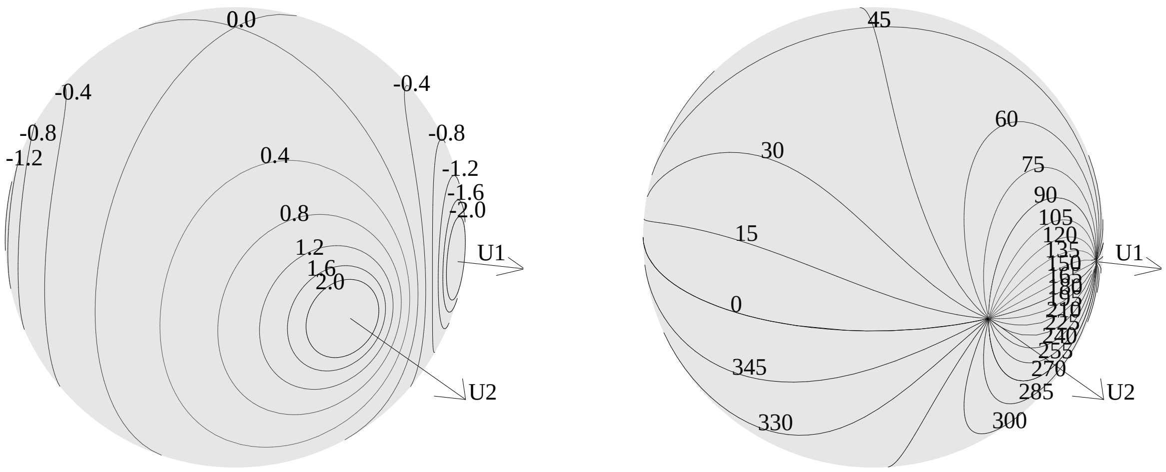
$$|\mathbf{u}_1 \times \mathbf{v}| - \rho|\mathbf{u}_2 \times \mathbf{v}| = 0, \tag{1}$$

the ratio of lengths between the projections of $\mathbf{u}_1$ and $\mathbf{u}_2$ is ρ. Squaring this equation yields a homogeneous equation of degree 2 in $\mathbf{v}$. For viewing directions which satisfy:

$$(\mathbf{u}_1 \times \mathbf{v}) \cdot (\mathbf{u}_2 \times \mathbf{v}) - \cos\theta |\mathbf{u}_1 \times \mathbf{v}||\mathbf{u}_2 \times \mathbf{v}| = 0, \tag{2}$$

the angle between the projections of $\mathbf{u}_1$ and $\mathbf{u}_2$ is θ. Squaring this equation yields a homogeneous equation of degree 4 in $\mathbf{v}$. Note that because only $\cos^2\theta$ appears in the homogeneous equation curves for $\pi - \theta, 2\pi - \theta, -(\pi - \theta)$ also result from this equation and have to be separated.

For tracing the curves we use an algorithm for tracing algebraic curves [15] which relies on homotopy continuation [16] to find all curve singularities and construct a discrete representation of the smooth branch curves. In order to insure that the viewing directions lie on the viewing sphere we add an additional equation $|\mathbf{v}|^2 = 1$ to the equations derived earlier. Figure 2(a) shows curves for ratios of lengths ρ such that $\log\rho$ is between -2 and 2 where the original $\log\rho$ was 0. Curves for $\log\rho$ less than 0 surround the viewing direction parallel to $\pm\mathbf{u}_1$. The curves for $\log\rho$ equal 0 intersect at the viewing direction orthogonal to the plane spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. For $\log\rho$ greater than 0 the curves now surround $\pm\mathbf{u}_2$. Figure 2(b) shows curves for angles between 0 and $360°$ where the original angle was $45°$. All the curves emanate at viewing directions parallel to $\pm\mathbf{u}_1$ and end at viewing directions parallel to $\pm\mathbf{u}_2$. This occurs because at those directions only one of the two edges is visible and therefore the angle between the two edges is not defined. For angles less than the "real" angle $45°$ the curves go between $\mathbf{u}_1$ and $\mathbf{u}_2$ or between $-\mathbf{u}_1$ and $-\mathbf{u}_2$ starting for $0°$ from directions in the plane of $\mathbf{u}_1$ and $\mathbf{u}_2$. At the "real" angle the two curves intersect like before at the viewing direction orthogonal to the $\mathbf{u}_1$ $\mathbf{u}_2$ plane. For angles greater than the "real" angle the curves go between $\mathbf{u}_1$ and $-\mathbf{u}_2$ or between $-\mathbf{u}_1$ and $\mathbf{u}_2$ ending for $180°$ with viewing directions which also lie in the $\mathbf{u}_1$ $\mathbf{u}_2$ plane. For angles θ greater than $180°$ the curve is the reflection the curve for $360° - \theta$ through the origin.



(a) (b)

Figure 2: For two lines of equal length with 45° between them: (a) the viewing sphere with isoratio curves where $\log(\rho)$ is traced; (b) the viewing sphere with isoangle curves.

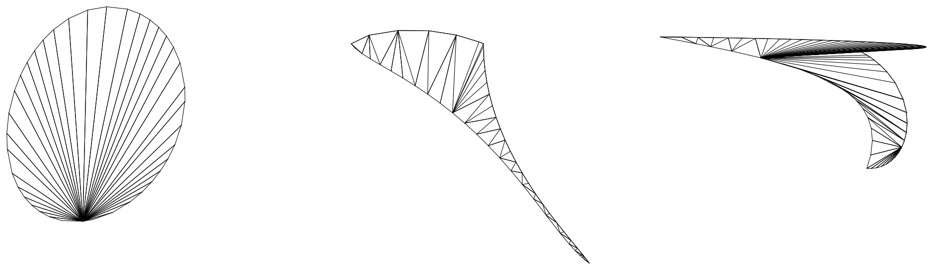


Figure 3: Computing the area bounded by a curve by triangulation.

2.2 Computing Probability Density Functions

A distribution function $F(x)$ which is the probability that $x \leq a$ is related to the probability density function $f(x)$ by:

$$F(a) = \int_{-\infty}^{a} f(x)dx, \tag{3}$$

Therefore when given one, the other can be computed. In our case we compute $F(a)$ for a ratio or an angle as follows. The curve for a is traced on the viewing sphere. This curve bounds the part of the viewing sphere for which $x \leq a$. The area bounded by the curve is triangulated using the points used to trace the curve. Examples of such triangulations are shown in Figure 3. The triangles are actually spherical triangles whose area is computed exactly, and the sum of the areas of the triangles is the area bounded by the curve.

The relationship between $F(x)$ and $f(x)$ is:

$$f(x) = \frac{dF(x)}{dx}. \tag{4}$$

Therefore, using finite differences we can estimate $f(x)$ by:

$$f(x) = \lim_{\epsilon \to 0} \frac{F(x+\epsilon) - F(x-\epsilon)}{2\epsilon}, \tag{5}$$

by computing $F(x+\epsilon)$ and $F(x-\epsilon)$.

Figure 4(a) shows the probability density function for the log of the ratio between the ratio of lengths and the original ratio of lengths for various values of the original angle α between the edges. Graphs for different original ratios need not be displayed because for a given viewing direction there is a linear relationship between the original ratio and the observed ratio.

Figure 4(b) shows the probability density function for the log of the ratio between observed angle β and the original angle α for various values of α. A logarithmic scale is used to reflect the symmetry of the probability density function for ratios and to align the results for different angles in the density function for angles. In both cases the smaller the angle α

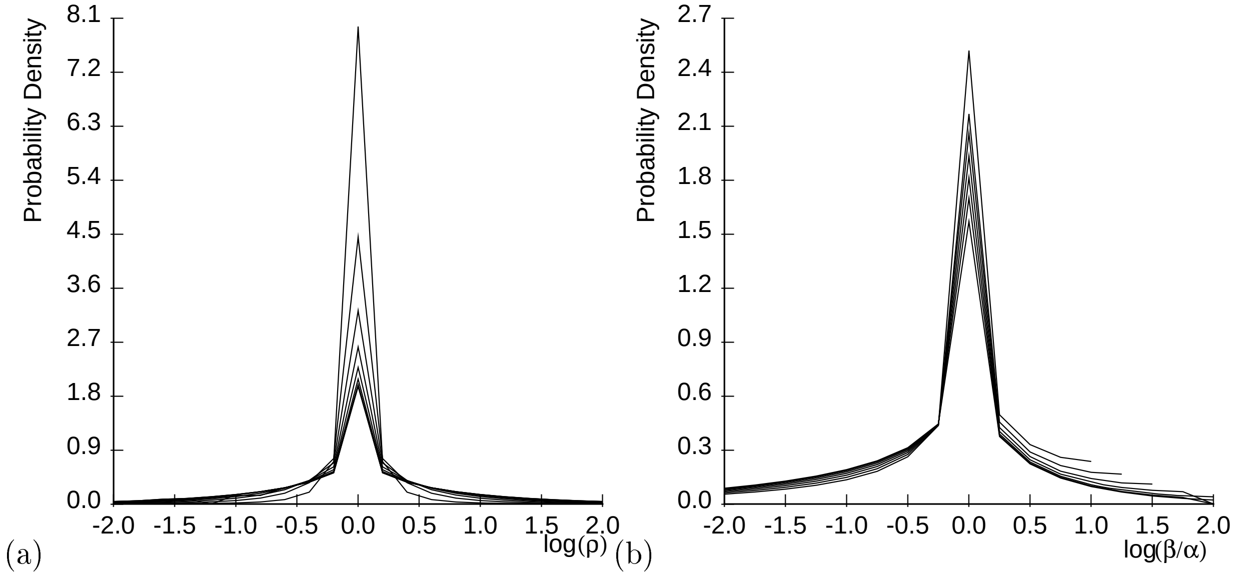


Figure 4: Probability densities for ratios and angles for various values of θ. (a) probability density for ratios. (b) probability density for angles.

the higher the peaking effect. In previous work Ben-Arie [4] and Burns et. al. [7] estimated probability density functions using ratios and angles computed of a uniform sample of the viewing sphere. However our method is more accurate.

2.3 Computing Joint Probability Density Functions

In the previous section we showed how to compute probability density functions for one value measured in the image. However, when performing a partial match of image features to model features such as the ones in Figure 1 more than one value can be measured (e.g. a ratio and an angle in Fig 1(a) or two angles in Fig 1(b)). Computing their joint probability density function can be used to further enhance our capability to correctly rank the different matching hypotheses for these image features. To demonstrate this we will examine the ratio-angle pair. The p.d.f $f_{\rho,\theta}(\rho,\theta)$ and the distribution function $F_{\rho,\theta}(\rho,\theta)$ are related in the following ways:

$$
\begin{aligned}
F_{\rho,\theta}(\rho,\theta) &= \int_0^{\rho}\int_0^{\theta} f_{\rho,\theta}(x,\alpha)\,dx\,d\alpha \\
f_{\rho,\theta}(\rho,\theta) &= \frac{\partial F}{\partial\rho\partial\theta}(\rho,\theta),
\end{aligned}
$$

where $F_{\rho,\theta}(\rho,\theta)$ is the probability that the ratio between the lines in the image is less than ρ and the angle between them is less than θ. We estimate $f_{\rho,\theta}(\rho,\theta)$ as:

$$f_{\rho,\theta}(\rho,\theta) = \lim_{\mu\to 0}\lim_{\nu\to 0} \frac{F_{\rho,\theta}(\rho+\mu,\theta+\nu)+F_{\rho,\theta}(\rho-\mu,\theta-\nu)-F_{\rho,\theta}(\rho+\mu,\theta-\nu)-F_{\rho,\theta}(\rho-\mu,\theta+\nu)}{2\mu 2\nu}, \tag{6}$$

where the numerator is the area bounded by the curves for $\rho \pm \mu$ and $\theta \pm \nu$. The area is found by tracing the four curves, finding the intersection points between them using homotopy, extracting the boundary of the region and computing its area.

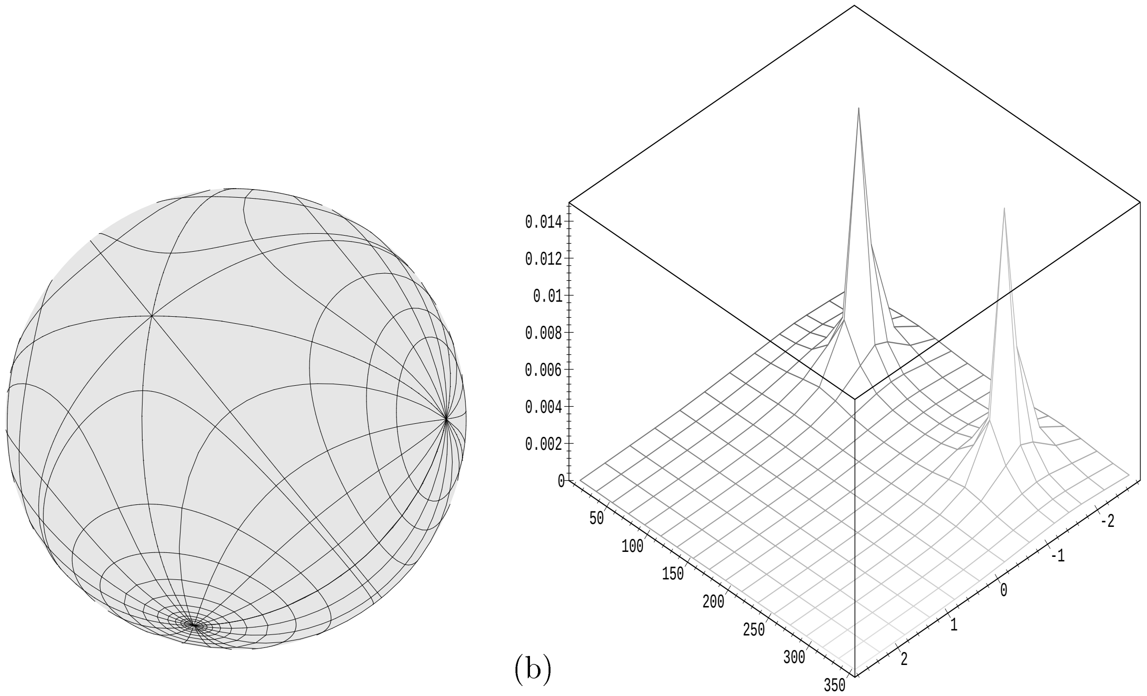
2.4 Computing Joint Distribution Functions

For one of the basic components of our recognition system we have to be able to compute for two adjacent edges $\mathbf{u}_1$ and $\mathbf{u}_2$

$$P(\rho_1 < \rho < \rho_2, \theta_1 < \theta < \theta_2), \tag{7}$$

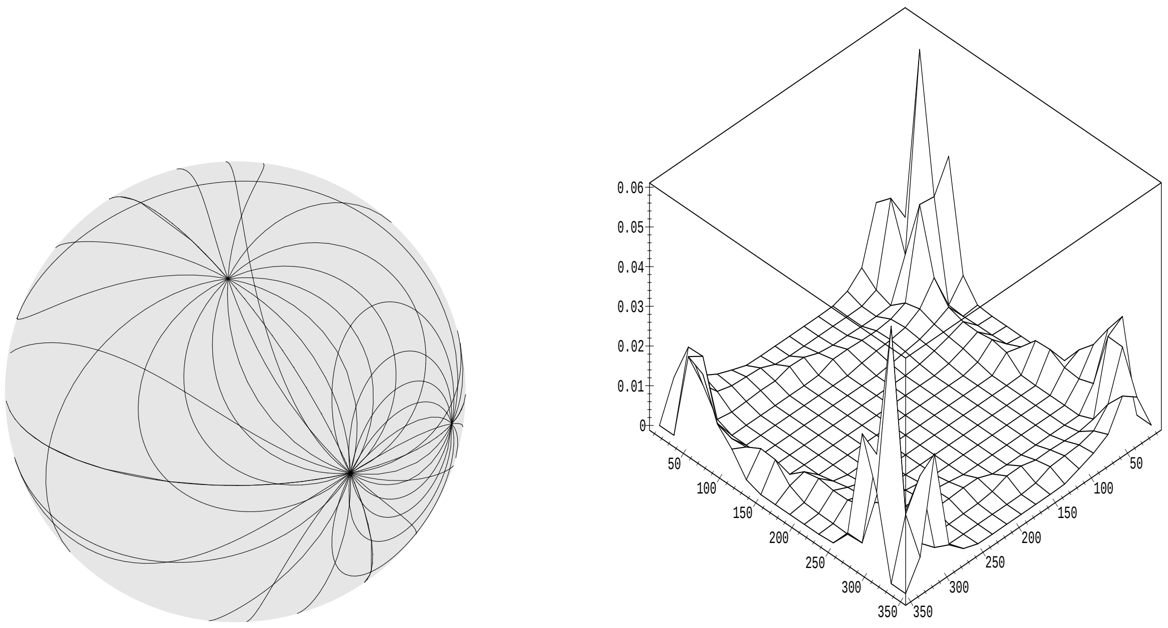
which is the probability that when viewing these edges from a randomly selected viewing direction the ratio of the lengths ρ will be between ρ_1 and ρ_2 and the angle between the edges will be between θ_1 and θ_2. We will denote this region of the (ρ,θ) space by $Reg(\rho_1,\rho_2,\theta_1,\theta_2)$. Using the technique described in the previous section we trace the curves of ρ_1,ρ_2,θ_1 and θ_2 and find the area bounded by them. Because that for every value of ρ or θ there are two curves, two such areas exist. We have divided the $(\log\rho,\theta)$ space into identical rectangles. The $\log\rho$ dimension was truncated at $-k$ and k where k is an arbitrary limit. The corresponding regions on the viewing sphere of each rectangle were traced producing a tessellation of the viewing sphere shown in Fig 5(a). The areas of the regions which represent $P((\rho,\theta)\in Reg(\rho_1,\rho_2,\theta_1,\theta_2))$ were plotted in Fig 5(b).

For corner feature sets curves for the two angles are plotted producing a tessellation of the viewing sphere shown in Fig 6(a). The areas of the regions which represent $P((\theta_1,\theta_2)\in Reg(\theta_{11},\theta_{12},\theta_{21},\theta_{22}))$ were plotted in Fig 6(b). Corners are characterized by the three angles



(a) (b)

Figure 5: (a) Tessellation of the viewing sphere into ratio/angle regions; (b) a graph of the joint probability function as estimated by the area of the regions in (a).



(a) (b)

Figure 6: (a) Tessellation of the viewing sphere into angle/angle regions; (b) a graph of the joint probability function as estimated by the area of the regions in (a).

between the three pairs of edges. Although the curves (such as the ones in Fig 2(b).) are determined by the first two angles the third angle influences how those curves will intersect and thus the area of the regions (if they exist). To demonstrate how different the results can be, we have computed the joint probability function of a number of corners, and display in Figure 7(top) the graph of the probability function. The bottom of the figure shows the non-empty regions.

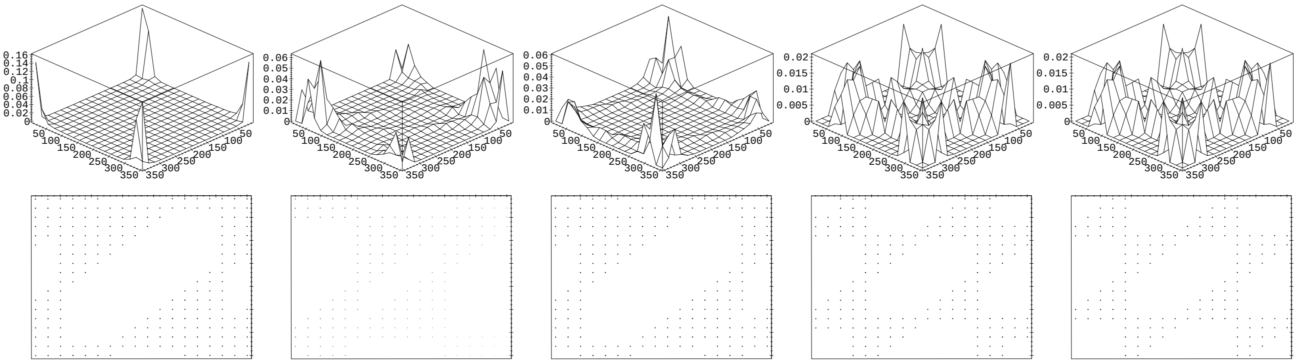


Figure 7: For corners with the following angles: $(10°, 20°, 25°)$, $(40°, 60°, 40°)$, $(40°, 60°, 80°)$, $(80°, 80°, 80°)$ and $(90°, 90°, 90°)$ graphs of joint probability functions and of non-empty regions are displayed.

Until now we have not taken occlusion into account. For the pair of edges both edges must be totally visible. In the case of the corner the corner must be visible but the edges can be partially occluded. Using the aspect graph of the object as computed by [11, 10], we determine in which of the noncritical regions the features are visible and which critical curves bound those regions. The area of the intersection between those regions and regions computed earlier produce the probability distribution. To demonstrate the effect of occlusion we chose two edges of an L shaped object (Fig 8(a)) computed its aspect graph (Fig 8(b)) and the probability distribution Fig 8(c),(d) compared to Fig 5 which disregards occlusion. Similar probability distributions have been computed for the case of the trihedral corner.

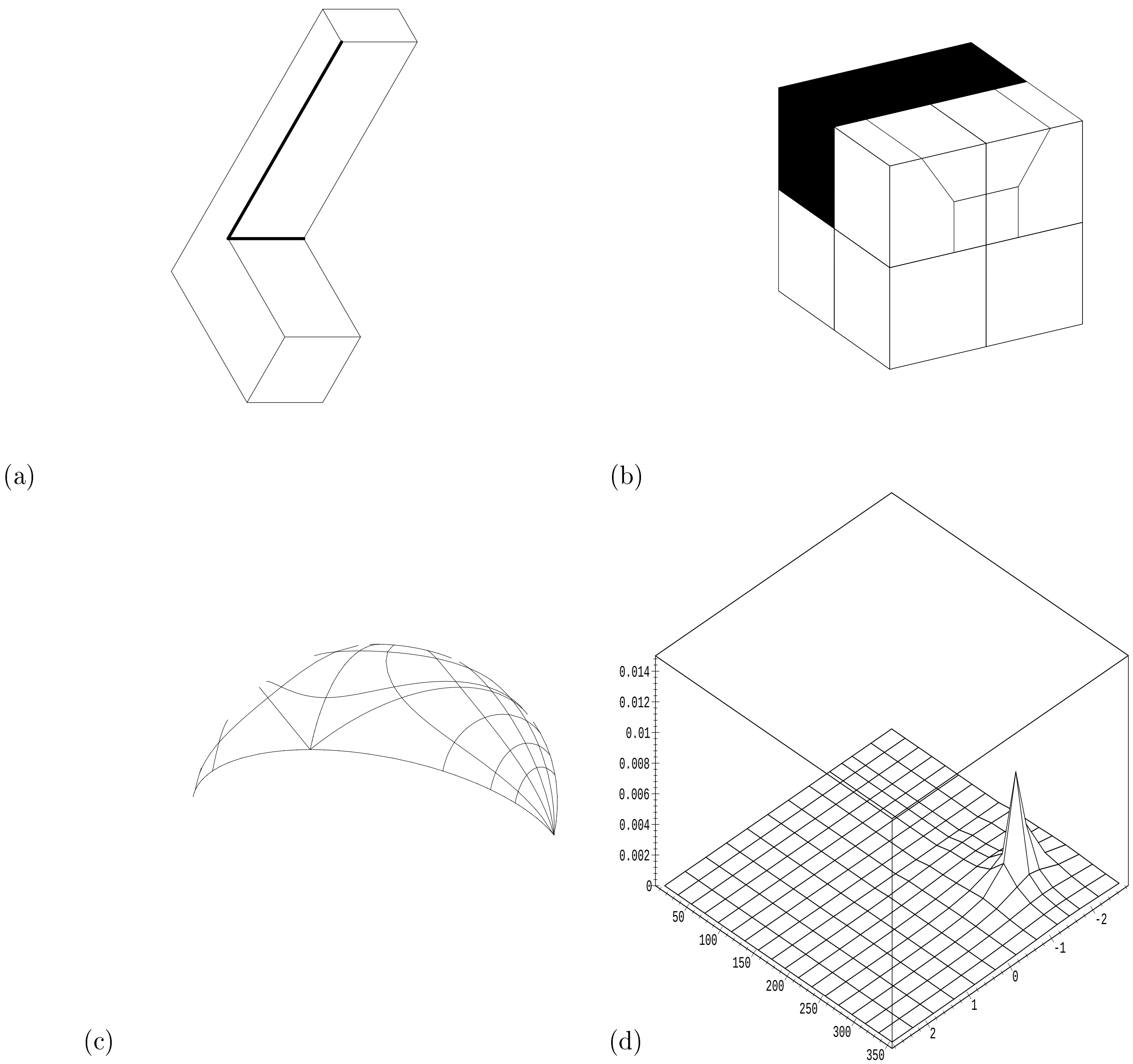


Figure 8: (a) The L shaped object with the selected edges highlighted. (b) the aspect graph of the object. The shaded region is the region in which the two edges are visible. (c) Tessellation of the viewing sphere into regions. (d) a graph of the joint probability function as estimated by the area of the regions in (c).

3 Pose Estimation

3.1 Pose Estimation

Estimating the pose from three points [1, 14] or a trihedral corner [18] have been extensively studied. We present here a single approach which deals with both types of feature sets. We regard the pose of an object in weak perspective projection as a combination of the following four components:

- a viewing direction $\mathbf{v}$ which is a point on the viewing sphere,
- a rotation of the image by δ degrees around the viewing direction,
- a scale s,
- a translation t in the image.

Therefore the pose is a point in the six dimensional space $\{S^2 \times [0 \cdots 2\pi] \times [s_{min} \cdots s_{max}] \times \mathbb{R}^2\}$ where S^2 is the unit viewing sphere, s_{min} and s_{max} are the minimum and maximum assumed scales respectively, and the translation t must be a point in the image. The pose transformation P for a point p is:

$$P(p) = sR(\delta)(p \cdot \mathbf{v}_2, p \cdot \mathbf{v}_3) + t, \tag{8}$$

where $\mathbf{v}, \mathbf{v}_2, \mathbf{v}_3$ is an orthonormal basis of $\mathbb{R}^3$ and $R(\delta)$ is a rotation by δ degrees.

We will now show what components of the pose can be recovered from angles and lengths measured in the image given an hypothesis which matches a minimal number of features in the image to features in the model. Each measured angle or ratio imposes a one-dimensional constraint on possible viewing directions. In order to determine $\mathbf{v}$ two such constraints are needed. Two pairs of types of curves were considered: the ratio/angle pair and the angle/angle pair. The ratio/ratio pair is not considered because in order to measure two ratios we would need an hypothesis which contains additional matches which we are trying

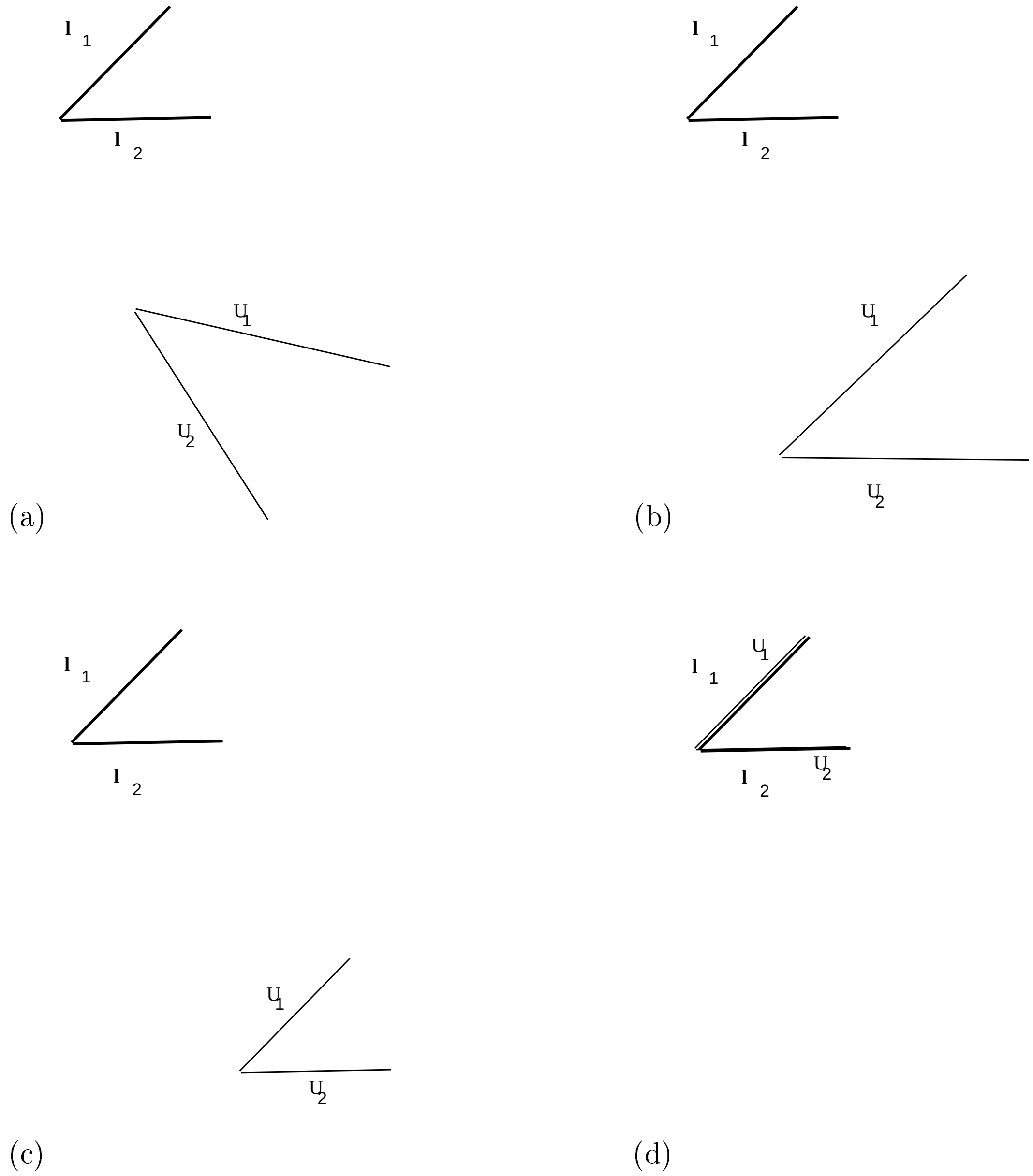


Figure 9: Stages in pose estimation from lines l_1 and l_2 in the image which match edges $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively in the model. (a) the viewing direction $\mathbf{v}$ estimated as the intersection between the curve for the ratio and the angle; (b) the angle δ is computed; (c) the scale s is computed. (d) the translation t is computed.

to avoid. The ratio/angle pair occurs when two adjacent edges in the image (Fig 1(a)) are matched with edges in the model. The angle/angle pair occurs when three edges which are adjacent to a vertex (Fig 1(b)) are matched to corresponding edges in the model. In this case the length of the edges does not have to be reliably measured in the image. $\mathbf{v}$ is obtained as the intersection points between two curves. The degrees of the ratio and angle curves are 2 and 4 respectively. Therefore the number of intersection points found for the ratio/angle and angle/angle pairs will be at most 8 and 16 respectively. However as shown in Section 2 the angle equations generate curves for $\theta, \pi-\theta, 2\pi-\theta$ and $-(\pi-\theta)$. As most of the solutions are for the other angles the number of solutions is much less. In some cases some of the remaining solutions can be eliminated by visibility considerations (i.e., the features are occluded from that viewing direction). The rotation angle δ is determined for both pairs by rotating the results of applying the projection of the object in direction $\mathbf{v}$ (Fig 9(a)) until the angle between the corresponding edges is $0°$ (Fig 9(b)). The scale and translation can not be obtained for the angle/angle case except if we assume that the vertex in the model is at the origin. Therefore only in the case of the ratio/angle pair they are recovered (Fig 9(c) and Fig 9(d)).

3.2 Pose Uncertainty Estimation

Until now we have not taken into account the uncertainty in the values we measured in the image. We measure three coordinate pairs in the ratio/angle case and four coordinate pairs in the angle/angle case. Assuming that each of these values has a normal distribution with the measured value as its mean and with a certain variance σ^2 (which can be found empirically) this induces a probability density function on the pose of the object due to these features. Thus, a multivariate normal distribution of the values measured in the image induces a distribution on the pose space.

When trying to determine whether two feature sets are part of one instance of a certain model we use the poses recovered from them. Had there been no uncertainty in the measurements of the feature sets, both poses would be equal. When taking uncertainty into consideration we have to calculate the probability that the two poses are close enough. How to calculate the exact probability will be discussed in Sect. 4. In this section we describe how to eliminate most pairs of poses whose matching probability is negligible [12, 13].

By assuming that the uncertainty in the values measured in the image is bounded by $k\sigma$ where k is an arbitrary constant, we compute the size of the subspace of the pose space which could generate the feature set. If the subspaces for both hypotheses have an empty intersection the hypothesis pair is discarded. The purpose of this section is to show how to compute the pose subspace associated with an hypothesis.

We divide the computation into two stages. In the first stage we compute the uncertainty in the values measured in the image such as ratios, angles and lengths due to the uncertainty in vertex position. This stage does not depend on the model features matched to the image feature set. In the second stage we combine the results of the first stage with the uncertainty due to interrelations between the various components of the pose to produce the estimated pose uncertainty subspace.

In the first stage we find the maximum and minimum values for the various values measured in the image assuming the uncertainty is bounded by $\epsilon = k\sigma$. For the ratio/angle case (Fig. 10(a)) given vertices P_0,P_1 and P_2 we find P_0',P_1' and P_2' within their respective uncertainty regions to get extreme values for ratios angles or lengths. As we assume that the extreme values happen on the boundary of the uncertainty region we only need to find three angles ϕ_i such that $P_i' = P_i + (\epsilon\cos(\phi_i), \epsilon\sin(\phi_i))$. To demonstrate the general approach we shall deal with the case of finding extreme ratios between the length of the lines l_1' and l_2'. We define $F(\phi_0, \phi_1, \phi_2) = (||l_1'||/||l_2'||)$. The maximum (minimum) is obtained when

$$\frac{\partial F}{\partial \phi_0}(\phi_0, \phi_1, \phi_2) = \frac{\partial F}{\partial \phi_1}(\phi_0, \phi_1, \phi_2) = \frac{\partial F}{\partial \phi_2}(\phi_0, \phi_1, \phi_2) = 0.$$

At first it seems that a three parameter optimization must be performed to obtain the

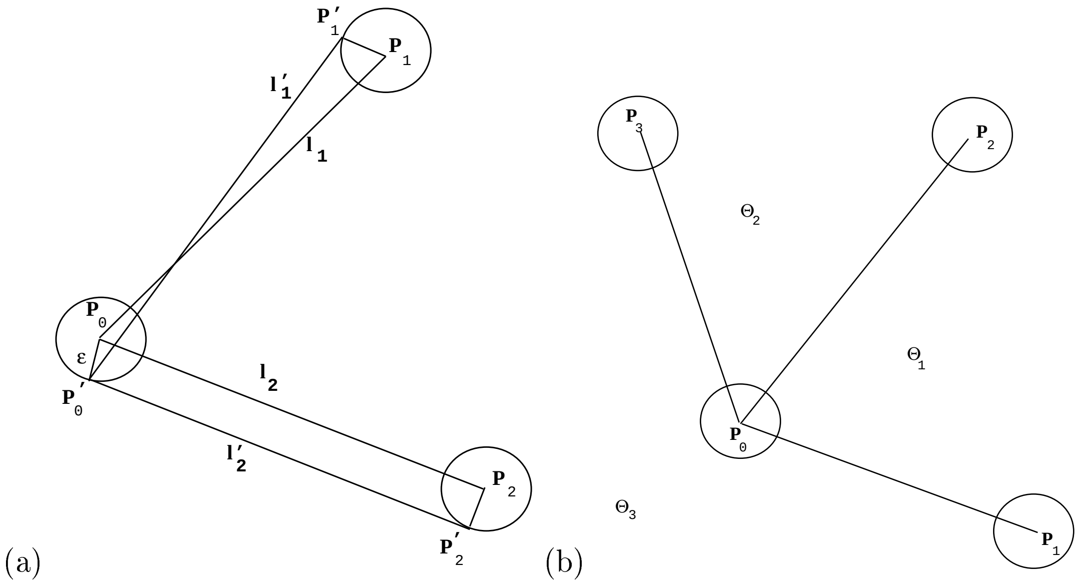


Figure 10: Given vertices P_i, find P_i' within their respective uncertainty regions to get extreme values for ratios angles or lengths: (a) the two-edge case; (b) the trihedral corner case.

Value	Function	Computing ϕ_1 and ϕ_2
Ratio	$\|\|l_1'\|\|/\|\|l_2'\|\|$	$(P_i - P_0') \parallel (P_i' - P_i)$
Angle	$cos^{-1}((l_1' \cdot l_2')/(\|\|l_1'\|\|\|\|l_2'\|\|))$	$(P_i' - P_i) \perp (P_i' - P_0')$
Scale	$(\|\|l_1'\|\|/\|\|l_1\|\| + \|\|l_2'\|\|/\|\|l_2\|\|)/2.0$	$(P_i - P_0') \parallel (P_i' - P_i)$
Rotation	$Bis^{\perp} \cdot Bis'$ where Bis is the normalized bisector	$(P_i' - P_i) \perp (P_i' - P_0')$
Translation	Maximal value is ϵ	$\phi_i = \phi_0$

Table 1: A table describing the different values which we would like to maximize (minimize), the function which is maximized (minimized) and how to compute ϕ_1 and ϕ_2 given ϕ_0.

maximum (minimum) ratio. However when given a value for ϕ_0, the values for ϕ_1 and ϕ_2 can be computed directly. In this case ϕ_i is chosen such that $P_i - P_0'$ is parallel to $P_i' - P_i$. The direction of $P_i' - P_i$ is chosen according to which type of extreme value is required. Thus we are left with a one-dimensional optimization which can be easily computed. In Table 1 we summarize similar techniques for finding the extreme values for all the values which we are looking for. Examples of the extreme values are shown in Figure 11. For the case of a corner (Fig. 10(b)) we measure the minimum and maximum values for $\theta_1, \theta_2, \theta_3 = 2\pi - (\theta_1 + \theta_2)$ and the rotation.

In the second stage components of the pose subspace are computed. When computing

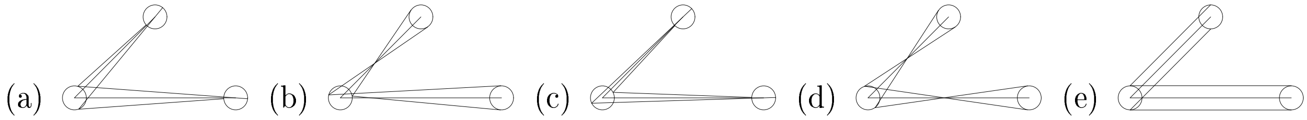
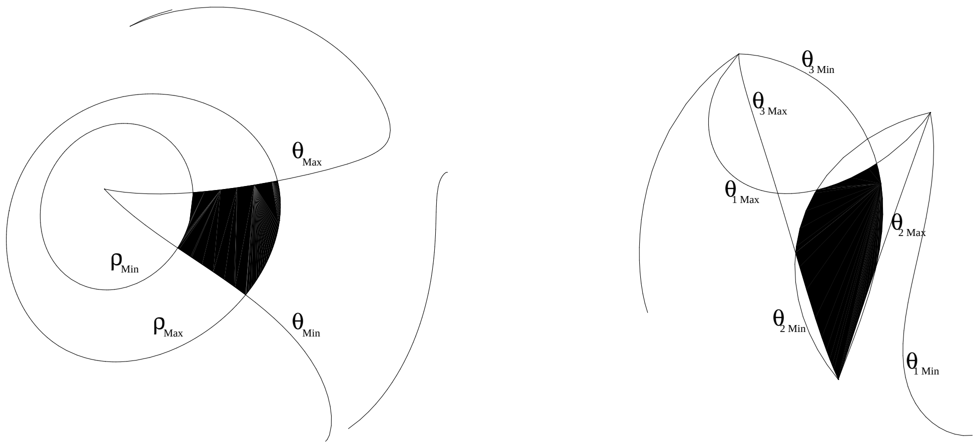


Figure 11: The cases where minimal and maximal values are measured. Bounded uncertainty in vertex position is assumed. (a) ratio between the lengths of edges; (b) angle between the edges; (c) length of edges; (d) rotation of the bisector of the two edges; (e) translation of the feature set in the y direction.



(a) (b)

Figure 12: The uncertainty region on the viewing sphere bounded by curves for the possible minimal and maximal values measured in the image. (a) the two-edge case; (b) the trihedral corner case.

the projection of a model point, transformations due to $\mathbf{v}, \alpha, s$ and t are applied to the point in that order. Therefore, uncertainty in the initial components of the projection can increase the uncertainty of the other components of the pose. The uncertainty of α and s is affected by the uncertainty in $\mathbf{v}$, and the uncertainty in t is affected the uncertainty in $\mathbf{v}, \alpha$ and s.

For the two-edge case the region of the viewing sphere which belongs to the subspace is the region bounded by the curves for $\rho_{max}, \rho_{min}, \theta_{max}$ and θ_{min} which were computed in the previous step (Fig. 12(a)). Using the boundary of the region we find the minimal and maximal values for the position of one of the vertices, the rotation, and the scale of the projected features. Using the extreme values for the the rotation α and scale s from both stages, we compute a a worst case estimate for them. In order to compute the translation component we have to take into account the effects of all previous parts of the pose. We start with the region denoted by the extreme polar coordinates $(\beta_{min} - \beta_{max}, r_{min} - r_{max})$ of the projection of a model feature point due to the viewing direction. Rotation affects the angle part of the region and scale affects the length part of the region. To that we add ϵ the translation uncertainty in the image and get the translation part of the subspace.

For the case of the trihedral corner we compute the boundary of the region on the viewing sphere bounded by curves for the minimum and maximum values of θ_1, θ_2 and θ_3 (Fig 12 (b)). As in the previous case we estimate the uncertainty for the rotation. Due to lack of information the components of scale and translation can not be estimated independently. Therefore for each possible value of the scale we estimate the translation uncertainty.

In Table 2 we tabulate the pose uncertainty for model feature sets with various ratios and angles given an image feature set with ratio 1.4 and angle 315°. Nearly all the components of the pose subspace do not change much between the different examples. Only the size of the viewing sphere component changes dramatically for the different examples. The closer the ratio and angle measured in the model are to the values measured in the image the larger the viewing sphere component is.

This phenomenon, which is another aspect of the probabilistic peaking effect can be exploited for solving the following problem. Given a number of probable correspondences between points in the image and points in a model choose a subset of the correspondences to compute the best pose estimation [9]. For each triple the smaller the size of the subspace the smaller the uncertainty in the pose. For larger point subsets the intersection of subspaces can be computed and the subset which yields the smallest subspace is chosen.

4 Object Recognition

In this section we use the techniques presented in the previous sections for recognition of 3D objects. This method tackles two important problems which recognition algorithms come

Ratio	Angle	Viewing Sphere	Rotation	Scale	Translation
1.0	280.0°	0.011	0.092	0.139	0.0080
0.3	90.0°	0.001	0.101	0.147	0.0329
1.4	315.0°	0.085	0.184	0.158	0.0459
0.3	315.0°	0.001	0.109	0.173	0.0099
1.4	90.0°	0.009	0.124	0.114	0.0142

Table 2: Tabulation of the fractions of components of the uncertainty pose subspace for a given image feature set whose ratio and angle are 1.4 315.0° respectively for several model feature sets with different ratios and angles.

across. First, we deal with the fact that the results of edge and corner detection are noisy causing many features not to be detected or only partially detected. Second, we developed a probabilistic expression to rank recognition results taking into account the number of image features matched to model features and the uncertainty in the values measured in the image.

4.1 Preprocessing

In the preprocessing stage lookup tables $RA(\rho, \theta)$ and $AA(\theta_1, \theta_2)$ are built for the ratio/angle and angle/angle pairs respectively. The tables are built by combining joint probability distribution tables which were described in Sect. 2.4 for all adjacent edges or corners in the model database.

Entries for hypotheses whose probability is below a preset threshold ϕ are not entered into the tables. The threshold is set such that:

$$\sum_{p(x)>\phi} p(x)dx = \Phi, \tag{9}$$

where Φ is a probability threshold. In setting Φ there is a tradeoff between the speed of the algorithm and the success rate of the algorithm.

The probabilities computed are used to determine the induced probabilities on the identity of the object. We denote by E a feature set measured in the image with values within a certain region of the (ρ, θ) or (θ_1, θ_2) set, by M_i the event in which the i^{th} model from the database appears in the image, by $H_j^{(i)}$ the hypothesis that the j^{th} model feature set of the i^{th} model matches E and by $S_j(E)$ the area on the viewing sphere from which E will be observed if hypothesis $H_j^{(i)}$ is true. Using Bayes' law

$$P(M_i|E) = \frac{P(E|M_i)Pr(M_i)}{P(E)},$$

where

$$P(E) = \sum_i P(E|M_i)P(M_i),$$

and

$$P(E|M_i) = \sum_j |S_j(E)| = \sum_j P(E, H_j^{(i)}|M_i),$$

and

$$P(M_i|E) = \sum_j P(H_j^{(i)}|E) = \frac{\sum_j P(E, H_j^{(i)}|M_i)P(M_i)}{\sum_i \sum_j P(E, H_j^{(i)}|M_i)P(M_i)}.$$

From that we deduce that

$$P(H_j^{(i)}|E) = \frac{P(E, H_j^{(i)}|M_i)P(M_i)}{\sum_m \sum_j P(E, H_j^{(m)}|M_m)P(M_m)}. \tag{10}$$

$P(M_i)$ can be determined using any prior knowledge we have about the likelihood of a model to appear in the image. $P(H_j^{(i)}|E)$ is computed for the hypotheses in the list for each entry in the tables and the lists are sorted by probability.

It is important to note that the regions in the (ρ, θ) or (θ_1, θ_2) spaces do not have to be equal. Moreover, the recognition task will perform better if regions with multiple probabilistic peaks are divided into smaller regions. In general, the (ρ, θ) or (θ_1, θ_2) space can be partitioned using a quadtree approach where a region is divided into four regions only if the order of the hypotheses of the subregions is different than the order of the hypotheses of the parent region.

4.2 Dealing With Uncertainty

The input to the recognition phase of the algorithm is the result of edge and corner detection performed on the image. We identify three types of uncertainty which have to be dealt with: certain feature sets which appear in the image will not be recovered by edge and corner detectors, the image features might not match any model features, and the uncertainty in the values measured for image features.

We denote by P_d the probability that a feature set which appears in the image will be detected by the edge and corner detectors. For each hypothesis evidence pair P_d can be obtained. In general, feature sets appearing on the silhouette of the object will have a higher P_d than internal features. The values for P_d for various types of feature sets is set empirically by testing the performance of the edge detection and feature selection stages of the algorithm on typical scenes.

Incorporating P_d into (10) yields:

$$P(H_j^{(i)}|E_k) = \frac{P(E, H_j^{(i)}|M_i)P(M_i)P_d^{(j,k)}}{\sum_m \sum_l P(E, H_l^{(m)}|M_m)P(M_m)P_d^{(l,k)}}.$$

The second type of uncertainty deals with feature sets which are not "legal" feature sets. Image features may not match model features for various reasons: the features were not correctly recovered (e.g. only part of an edge was recovered), the features belong to several occluding objects, or the features belong to objects not in the database or to the background. Computing the average number of "illegal" feature sets in a scene P_{il} depends on the types of scenes which the recognition process analyzes and the quality of the edge and corner detectors. However, P_{il} can be estimated empirically by testing the system on images of typical scenes. Different types of environments will produce different values for P_{il}. In order to incorporate this information into $RA(\rho, \theta)$ and $AA(\theta_1, \theta_2)$ tables we have to be able to compute the probability that we measured values within a region in the value space due to random features. Assuming that the length of the lines and the angles between them are uniformly distributed, the distribution of $\log \rho$ which is the log of the ratio of the lengths of two lines with uniformly distributed lengths is $\frac{1}{2}e^{-|x|}$. The probability for a pair of edges to be in a rectangle in the (ρ, θ) space is:

$$P((\log \rho, \theta) \in Reg(\rho_1, \rho_2, \theta_1, \theta_2)) = \frac{1}{2}|e^{-|\rho_1|} - e^{-|\rho_2|}|\frac{(\theta_2 - \theta_1)}{2\pi}.$$

The probability for a corner to be in a rectangle in the (θ_1, θ_2) space is:

$$P((\theta_1, \theta_2) \in Reg(\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22})) = \frac{(\theta_{12} - \theta_{11})}{2\pi} \frac{(\theta_{22} - \theta_{21})}{2\pi}.$$

In each entry in the tables a new hypothesis which represents the "illegal" feature set is added to the list multiplied by P_{il}.

The last type of uncertainty is uncertainty in the values measured in the image. For the two-edge case we assume that there is uncertainty in vertex position of the three vertices. We assume the six coordinates $a_i = \{x_1, y_1, x_2, y_2, x_3, y_3\}$ have a normal distribution with $\mu = a_i$ and σ^2 which is determined empirically. This induces a probability distribution on ρ and θ which enable us to compute the probability that the actual measured values are in a region of the (ρ, θ) space. Thus

$$P(H|E) = \sum_k P(H|E_k)P(E_k|E) = \sum_k P(H|E_k) \frac{P(E|E_k)P(E_k)}{\sum_j P(E|E_j)P(E_j)}$$

where $P(E|E_k)$ is the probability that E is in the k^{th} region of the (ρ, θ) space. For most regions E_k except the ones close to E, $P(E|E_k)$ is negligible.

4.3 Recognition

Edge detection, line fitting and corner detection are performed on an image. Corners of out-degree 3 and pairs of edges assumed to be fully detected are collected. Lengths, ratios and angles are measured. For each recovered feature set lists from the appropriate entry in the table are extracted. The lists are merged and sorted by probabilities. The merged list of hypotheses is traversed in decreasing probability. For each hypothesis we try to find other hypotheses which match other image feature sets to other features in the same model, to reinforce the hypothesis. For two hypotheses to reinforce each other they have to fulfill the following conditions:

- They have to belong to the same model.

- Their partial matches must be compatible.
- Their uncertainty pose subspaces must have a non-empty intersection.

The probability for a random false recognition goes down considerably as the size of the compatible match gets bigger. However the major problem is not random matches but when several models have similar features. For example consider two models with (not necessarily similar) rectangular faces. The false recognition would find features of that face but no others where as a correct recognition would be able to match more features. In order to deal with such problems we need a measure to rank hypothesis sets of various sizes. We therefore derived a probabilistic expression which takes into account the size of the set and the uncertainty pose subsets obtained from the different hypotheses.

Given a set of image features which participate in a basic hypothesis h (ratio/angle or angle/angle) the uncertainty in the values a_i measured in the image associated with h induces uncertainty in the pose suggested by h. The p.d.f's of the measured values in the image induce a p.d.f in pose space.

Given two feature sets in the image e_1 and e_2 and two respective hypotheses h_1 and h_2, we define H as the hypothesis that h_1 and h_2 are true and both match image feature sets to the same instance of a certain model M. We would like to compute $P(h_1, h_2, H|e_1, e_2)$. Using Bayes' rule

$$P(h_1, h_2, H|e_1, e_2) = \frac{P(e_1, e_2|h_1, h_2, H)P(h_1, h_2, H)}{P(e_1, e_2)}. \tag{11}$$

We can write $P(e_1, e_2|h_1, h_2, H)$ as a marginal distribution with respect to the pose p. Therefore

$$P(e_1, e_2|h_1, h_2, H) = \int_p P(e_1|h_1, H, p)P(e_2|h_2, H, p)f_p(p)dp,$$

leading to

$$P(e_1, e_2|h_1, h_2, H) = \int_p f_1(B_1(p))f_2(B_2(p))f_p(p)dp, \tag{12}$$

where f_1 and f_2 are the p.d.f's for values measured for h_1 and h_2 respectively, B_1 and B_2 are functions which backproject the feature sets associated with h_1 and h_2 respectively in the image, and $f_p(p)$ is the p.d.f of poses in the pose space which we will assume is uniformly distributed.

$$P(h_1, h_2, H) = |S(h_1) \cap S(h_2)| P(M) P_d^{(h_1)} P_d^{(h_2)},$$

where $|S(h_1) \cap S(h_2)|$ is the area of viewing sphere in which both model feature sets are visible, and $P_d^{(h_1)}$ and $P_d^{(h_2)}$ are the probabilities that the hypotheses are detected by the edge/corner detector.

In computing $P(e_1, e_2)$ we marginalize over every pair of hypotheses h_i, h_j which could generate e_1 and e_2 respectively and whether e_1 and e_2 belong to the same object (H).

$$\begin{aligned} P(e_1, e_2) = & \textstyle\sum_i \sum_j (P(e_1, e_2, h_i, h_j, H^{(i,j)}) + P(e_1, e_2, h_i, h_j, \neg H^{(i,j)})) = \\ & \textstyle\sum_i \sum_j (P(e_1, e_2 | h_i, h_j, H^{(i,j)}) P(h_i, h_j, H^{(i,j)}) + \\ & P(e_1, e_2 | h_i, h_j, \neg H^{(i,j)}) P(h_i, h_j, \neg H^{(i,j)})). \end{aligned}$$

When e_1 and e_2 are feature sets of the same instance of a certain model the probabilities are computed in the same way that we computed the numerator. When e_1 and e_2 do not belong to the same object Then

$$P(e_1, e_2, h_i, h_j, \neg H^{(i,j)}) = P(e_1, h_i) P(e_2, h_j) P(\neg H^{(i,j)}).$$

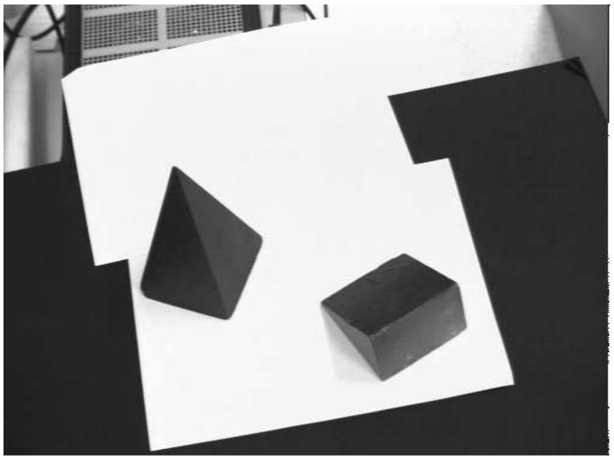
When h_i and h_j represent features belonging to different models, $P(\neg H^{(i,j)}) = 1$. If not $P(\neg H^{(i,j)})$ is determined empirically. Combining all these results we are able to compute $P(h_1, h_2, H | e_1, e_2)$. This derivation can be easily extended to more than two hypotheses.

When computing the various terms in (11), all the terms except (12) can be precomputed. We evaluate (12) using a Monte-Carlo technique. As the value of the integrand is negligible outside the intersection of the pose uncertainty subspaces, we only have to evaluate it for poses belonging to the intersection of the subspaces. This reduces the number of evaluations which have to be made, which speeds up the computation. As terms similar to (12) appear in the denominator of (11) whenever a new set of hypotheses with a non-empty pose subspace is found, (12) is computed and added to previous results of type (11) where it belongs. As more hypothesis sets are found the values for the ranks of the other sets are updated and converge to (11) when all hypothesis sets have been checked. As the hypotheses are traveresed in decreasing probabilities there is a high probability that the best recognition results will have the highest rank even before all hypotheses have been examined. Therefore, if the recognition process has to be interuppted due to lack of time still we can assume that the correct recognition results will be ranked high on the list.

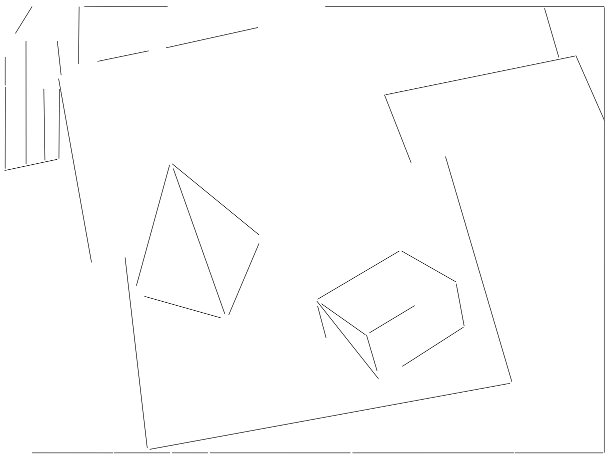
Another important characteristic of our ranking process is the way “popular” feature set sets such as rectangular faces are dealt with. In that case the same feature sets will participate in many hypothesis sets. Therefore their values of (12) will contribute to the denominators of all their ranks. Thus reducing their ranks. Where as a less “popular” but probably correct set of feature sets will have a higher rank.

4.4 Results

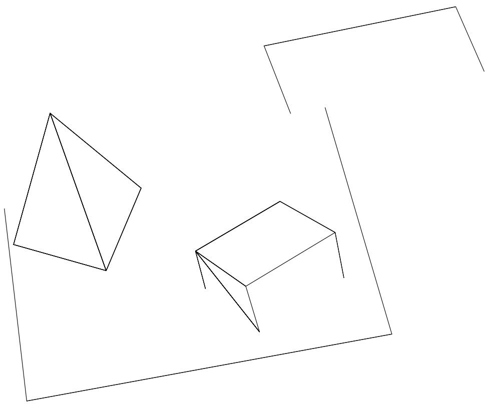
In this section we present preliminary recognition results performed on a real image. Figure 13(a),(b) shows an image and the results of edge detection performed on it. Fig. 13(c) shows the features belonging to feature sets extracted from the detected edges. Note that not all feature sets that appear in the image were detected and that some of the feature sets extracted are due to the background and shadows. As the mechanisms used for edge, corner, and line detection, and feature set extraction are not the subject of this paper, standard techniques were used. An important feature of our algorithm is that works well even with suboptimal results from the initial stages. In Fig. 13(d) the objects recognized by the algorithm are shown. Many false recognition candidates with rectangular faces were recovered using features belonging to the rectangular face of the prism. However, these



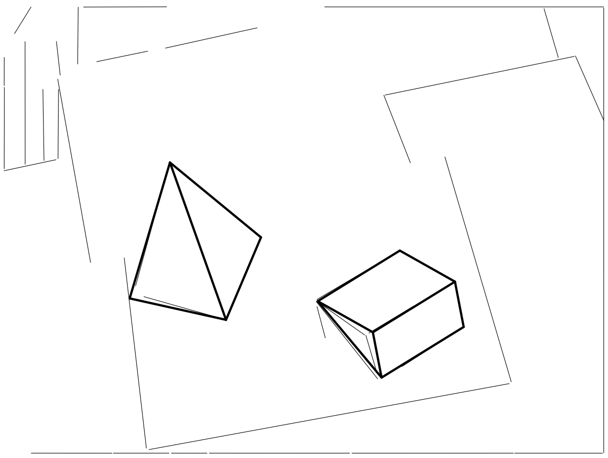
(a)



(b)



(c)



(d)

Figure 13: (a) An image of a number of objects; (b) the results of edge-detection on one of the objects; (c) feature sets recovered from the object; (d) recognized objects.

results were discarded because no other reinforcing features could be added to the set of feature sets and because all these results contributed to the denominator of (11), reducing their rank.

5 Discussion

In this paper we have presented a probabilistic 3D object recognition algorithm. We have studied the nature of the isoratio and isoangle curves and traced them on the viewing sphere. We used these curves to accurately compute conditional probabilities that image features matched model features. The probabilities were incorporated into a probabilistic model which takes into account the uncertainties inherent in the input to the recognition algorithm. These probabilities are used to decide the order in which to process the initial hypotheses. We developed a method to compute the pose of the object using a minimal feature set in the image which matches a feature set in the model database. Taking the uncertainty in values measured in the image into account we computed the uncertainty pose subspace for each hypothesis. By finding pose subspaces which have a non-empty intersection we were able to find a set of feature sets which reinforce the recognition hypothesis. We ranked these hypotheses by computing the probability that all the feature sets in the set came from the same instance of the suggested object.

Future work will be dedicated to extending this work to more complicated objects such as curved objects. We also plan to use components of this algorithm for robust feature selection for a pose estimation algorithm.

References

[1] T. D. Alter. 3D pose from 3 points using weak-perspective. *IEEE Trans. Patt. Anal. Mach. Intell.*, 16(8):802–808, August 1994.

[2] T. D. Alter and D. W. Jacobs. Error propagation in full 3D-from-2D object recognition. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, pages 892–898, Seattle, Washington, June 1994.

[3] N. Ayache and O. D. Faugeras. HYPER: a new approach for the recognition and positioning of 2D objects. *IEEE Trans. Patt. Anal. Mach. Intell.*, 8(1):44–54, January 1986.

[4] J. Ben-Arie. The probabilistic peaking effect of viewed angles and distances with application to 3-D object recognition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(8):760–774, August 1990.

[5] P.J. Besl and R.C. Jain. Three-dimensional object recognition. *ACM Computing Surveys*, 17(1):75–145, march 1985.

[6] T.O. Binford, T. Levitt, and W. Mann. Bayesian inference in model-based machine vision. In *Workshop on Uncertainty in Artificial Intelligence*, 1987.

[7] J. B. Burns, R. S. Weiss, and E. M. Riseman. View variation of point-set and line-segment features. *IEEE Trans. Patt. Anal. Mach. Intell.*, 15(1):51–68, January 1993.

[8] R.T Chin and C.R. Dyer. Model based recognition in robot vision. *ACM Computing Surveys*, 18(1):67–108, January 1986.

[9] T. L. Gandhi and O. Camps. Robust feature selection for object recognition using uncertain 2D image data. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, pages 281–287, Seattle, Washington, June 1994.

[10] Z. Gigus, J. Canny, and R. Seidel. Efficiently computing and representing aspect graphs of polyhedral objects. In *Proc. Int. Conf. Comp. Vision*, pages 30–39, Tampa, FL, December 1988.

[11] Z. Gigus and J. Malik. Computing the aspect graph for line drawings of polyhedral objects. *IEEE Trans. Patt. Anal. Mach. Intell.*, 12(2):113–122, February 1990.

[12] W. E. L. Grimson and D. P. Huttenlocher. On the verification of hypothesized matches in model-based recognition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13(12):1201–1213, December 1991.

[13] W. E. L. Grimson, D. P. Huttenlocher, and T. D. Alter. Recognizing 3D objects from 2D images; an error analysis. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, pages 316–321, Champaign, Illinois, June 1992.

[14] D. Huttenlocher and S. Ullman. Recognizing 3D solid objects by alignment with an image. *Int. J. of Comp. Vision*, 5(2):195–212, 1990.

[15] D.J. Kriegman and J. Ponce. A new curve tracing algorithm and some applications. In P.J. Laurent, A. Le Méhauté, and L.L. Schumaker, editors, *Curves and Surfaces*, pages 267–270. Academic Press, New York, 1991.

[16] A.P. Morgan. *Solving Polynomial Systems using Continuation for Engineering and Scientific Problems.* Prentice Hall, Englewood Cliffs, NJ, 1987.

[17] C. F. Olsen. Fast alignment using probabilistic indexing. In *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, pages 387–392, New York, New York, June 1993.

[18] Y. Wu, S. S. Iyenger, R. Jain, and S. Bose. A new generalized computational framework for finding object orientation using perspective trihedral angle constraint. *IEEE Trans. Patt. Anal. Mach. Intell.*, 16(10):961–975, October 1994.