Probabilistic Methods for 3–D Object Recognition

Joachim Hornegger[1], **Heinrich Niemann**[1], and **Ilan Shimshoni**[2]

[1]Lehrstuhl für Mustererkennung,
Universität Erlangen–Nürnberg,
Martensstr. 3, D–91058 Erlangen, Germany
Email: {hornegger,niemann}@informatik.uni-erlangen.de

[2]Department of Computer Science,
Technion,
Technion City 3200, Israel
Email: ilan@csa.cs.Technion.ac.il

Abstract

This work addresses various probabilistic approaches which are suitable for classification and localization of 3–D objects in gray–level images. We first present a statistical model for appearance based vision. The statistical behavior of image features, the projection model, the rotation and translation, as well as the assignment of model and image features are represented as a single density function. The second approach ranks correspondence hypotheses between model and image features by the probability that they are correct. For each hypothesis the object's pose is recovered and the region of the pose space compatible with the image uncertainty is computed. The introduced algorithms have been fully implemented. Examples demonstrate the suggested methods with real image data. **Keywords:** statistical object modeling, statistical object recognition, pose estimation, probabilistic peaking effect

1 Introduction and Motivation

There has been considerable and still increasing interest in applications of probabilistic methods for image analysis [8, 15, 17]. In many fields of computer vision statistical principles may lead to improvements compared with standard techniques, because uncertainty manifests itself in many aspects of image processing. The interpretation process is based on sensing information. Uncertainty results from image acquisition properties, like camera distortion, quantization errors, varying illumination conditions, or from model incompleteness, e.g. camera models. In addition the output of low–level preprocessing algorithms like edge detection may be inaccurate or incomplete. High–level processes must take into account this probabilistic behavior. The development of robust computer vision algorithms requires the understanding and adequate modeling of uncertainty.

This work concerns with new ideas to combine continuous and discrete random variables for the statistical description of objects in the space of observations. The probabilistic models consider the statistical behavior of features (like points, lines, angles or ratios of lengths), and the matching of model and image features. We will give recipes for the construction of statistical models which will either use composed probability density functions or the probabilistic peaking effect by tracing iso–angle and iso–ratio curves on the viewing sphere. We both show how the matching problem of observed and model features can be circumvented within the probabilistic formalism by marginalization, and discuss the generation of statistical match hypotheses and their ranking in the other case. The suggested models account for various types of uncertainty, and for each object class and pose hypothesis a probabilistic measure can be computed. Thus, the Bayesian decision rule can be applied to classify observed objects. The presented statistical framework will be tested for 3–D object recognition and localization using 2–D gray–level images.

2 Statistical Classification

The object recognition problem is understood as the assignment of (a subset of) observed image features to a class Ω_κ ($1 \leq \kappa \leq K$), which represents a single object or a set of special types of objects. Objects are usually represented as models and one of the major problems in computer vision is the recognition of 3–D objects in a scene as instances from a database of models. The computation of a transform which maps the model to the observed image features is commonly summarized as the localization problem of objects. In most applications, observed features depend on the object's pose. Thus classification and localization processes usually influence each other.

Statistical classifiers [4, 11] known from pattern recognition require feature vectors $\boldsymbol{c}$ of fixed dimensions and a probabilistic characterization of classes $\{\Omega_\kappa | 1 \leq \kappa \leq K\}$. The conditional densities $p(\boldsymbol{c}|\Omega_\kappa)$ and the a priori probabilities $p(\Omega_\kappa)$ should be known for each class Ω_κ of objects. For an observed feature $\boldsymbol{c}$, the optimal decision rule with respect to misclassifications is

$$\begin{aligned} \lambda &= \operatorname*{argmax}_\kappa p(\Omega_\kappa|\boldsymbol{c}) \\ &= \operatorname*{argmax}_\kappa \frac{p(\Omega_\kappa)\; p(\boldsymbol{c}|\Omega_\kappa)}{p(\boldsymbol{c})} \quad , \end{aligned} \tag{1}$$

i.e. we decide for the class with maximum a-posteriori probability. A classifier which is based on this decision rule is called Bayesian classifier [11]. The definition of the required a–posteriori probabilities is a highly non–trivial task, and it is a priori not obvious how this statistical concept can be applied to solve the object recognition and localization problem.

We clarify necessary extensions of the basic a-posteriori probabilities for single feature vectors by a simple example. Let us consider a 3–D cube, where the considered features are the cube's 3–D vertices. An object is thus not associated with a single, but a set of features. Any 2–D view of this object appearing in a gray–level image is the result of a mapping from the 3–D model into the 2–D image space combined with a rotation and translation in the model space. This transformation includes the following parts (see Figure 1):

- 3–D rotation $\boldsymbol{R}$ and translation $\boldsymbol{t}$,
- self–occlusion,
- projection $\mathcal{P}$ from the model into the image space, and

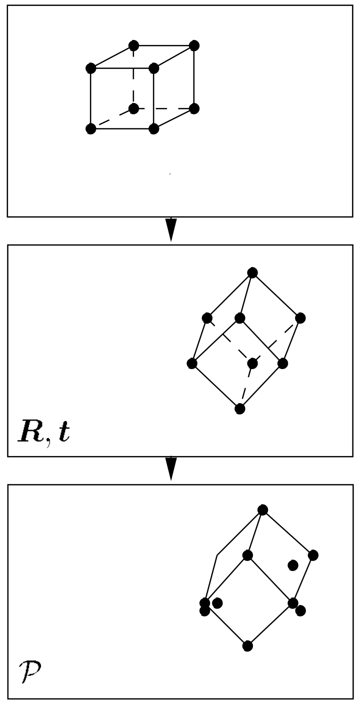

Figure 1: The transformation from the model into the image space

- segmentation errors (e.g. multiple detected point features in Figure 1).

This consideration shows that the number of observed, lower dimensional image features is not expected to be constant for arbitrary images. The cardinality of the available feature set depends on the viewing directions, on the illumination conditions, and on the used segmentation algorithms. In addition to components mentioned above, the correspondence ζ_κ between the features in a model of class Ω_κ and image features, and the range information is lost.

3 Statistical Modeling

The statistical description of 3–D objects appearing in images can be divided up into different statistical components: The uncertainty of observable features, the dependency of features on the object's pose, and the correspondence between model and image features.

3.1 Statistical Modeling of Features

Common features used for object recognition are point or line features as well as related measures like ratios or angles [2, 10]. An object in the model space is characterized by the set $\boldsymbol{c}_\kappa = \{\boldsymbol{c}_{\kappa,1}, \boldsymbol{c}_{\kappa,2}, \ldots, \boldsymbol{c}_{\kappa,n_\kappa}\}$ of model features. Of course, different features require different statistical representations. Independent of its concrete geometrical appearance, each observable feature $\boldsymbol{o}_k$ underlies either a parametric distribution given by a continuous density $p(\boldsymbol{o}_k|\boldsymbol{a}_{\kappa,l}, \boldsymbol{R}, \boldsymbol{t})$ or a discrete probability $p(\boldsymbol{o}_k|\boldsymbol{R}, \boldsymbol{t}) \in [0,1]$. Herein, $\boldsymbol{a}_{\kappa,l}$ characterizes the distribution of the corresponding model feature $\boldsymbol{c}_{\kappa,l}$ and the parameters $\boldsymbol{R}$ and $\boldsymbol{t}$ represent the degrees of freedom for the object's pose. The density $p(\boldsymbol{o}_k|\boldsymbol{a}_{\kappa,l}, \boldsymbol{R}, \boldsymbol{t})$, including model– and pose–specific parameters, results from the probability density function $p(\boldsymbol{c}_{\kappa,l}|\boldsymbol{a}_{\kappa,l})$ attached to $\boldsymbol{c}_{\kappa,l}$ by applying a standard density transform [1]. The mapping from the model into the image space is characterized by $\boldsymbol{R}$ and $\boldsymbol{t}$.

The composed density which characterizes the set of observed features $\boldsymbol{O} = \{\boldsymbol{o}_1, \boldsymbol{o}_2, \ldots, \boldsymbol{o}_m\}$ is given by

$$p(\boldsymbol{O}|\{\boldsymbol{a}_{\kappa,1}, \ldots, \boldsymbol{a}_{\kappa,n_\kappa}\}, \boldsymbol{R}, \boldsymbol{t}) = \prod_{k=1}^{m} p(\boldsymbol{o}_k|\boldsymbol{a}_{\kappa,l_k}, \boldsymbol{R}, \boldsymbol{t}) \quad , (2)$$

if the correspondence is known and statistical independence is assumed. The concrete representation of the involved probabilities depends on the used features and dependency structures.

If, for example, normally distributed 3–D features are given, then $\boldsymbol{a}_{\kappa,l}$ includes the 3–D mean vector and the (3×3)–covariance matrix. If an orthogonal projection from the 3–D model space into the 2–D image space is assumed, an affine transform is given by

$$\boldsymbol{o}_k = \boldsymbol{R}\boldsymbol{c}_{\kappa,l} + \boldsymbol{t} \quad , \qquad (3)$$

where $\boldsymbol{o}_k, \boldsymbol{t} \in \mathbb{R}^2$, $\boldsymbol{c}_{\kappa,l} \in \mathbb{R}^3$, and $\boldsymbol{R} \in \mathbb{R}^{2\times 3}$. The observable image features are also Gaussian distributed. The mean vector of the transformed feature is $\boldsymbol{R}\boldsymbol{\mu}_{\kappa,l} + \boldsymbol{t}$ and the covariance matrix is $\boldsymbol{R}\boldsymbol{\Sigma}_{\kappa,l}\boldsymbol{R}^T$ [1].

3.1.1 Point Features

Point features are quite often used for object recognition and localization. Statistical tests in [13, 16] show that point features in the image space are normally distributed. Let $\boldsymbol{o}_k \in \mathbb{R}^2$ be the 2–D point feature and $\boldsymbol{c}_{\kappa,l} \in \mathbb{R}^3$ the corresponding also normally distributed 3–D point feature. With respect to these constraints, we get

$$p(\boldsymbol{o}_k|\boldsymbol{a}_{\kappa,l}\boldsymbol{R}, \boldsymbol{t}) = \mathcal{N}(\boldsymbol{o}_k|\boldsymbol{R}\boldsymbol{\mu}_\kappa + \boldsymbol{t}, \boldsymbol{R}\boldsymbol{\Sigma}_{\kappa,l}\boldsymbol{R}^T) \quad , (4)$$

i.e. uncertainty of features is characterized by Gaussian density functions including the mean vector, the covariance matrix, and the parameters $\boldsymbol{R}, \boldsymbol{t}$ for the feature transform. For each feature, the densities differ in the mean vectors and covariance matrices, but they all share the pose–specific parameters $\boldsymbol{R}$ and $\boldsymbol{t}$.

3.1.2 Straight Line Segments

The statistical modeling of straight line segments is similar to point features. A 3–D straight line segment $\boldsymbol{c}_{\kappa,l}$ is characterized by an initial and an end point, i.e. $\boldsymbol{c}_{\kappa,l} = (\boldsymbol{c}_{\kappa,l,1}, \boldsymbol{c}_{\kappa,l,2}) \in \mathbb{R}^3 \times \mathbb{R}^3$. If statistical independence of these normally distributed points is assumed for simplicity, the statistical behavior of the straight line feature $\boldsymbol{c}_{\kappa,l}$ is given by

$$p(\boldsymbol{c}_{\kappa,l}|\boldsymbol{a}_{\kappa,l}) = \prod_{s=1}^{2} \mathcal{N}(\boldsymbol{c}_{\kappa,l,s}|\boldsymbol{\mu}_{\kappa,l,s}, \boldsymbol{\Sigma}_{\kappa,l,s}). \quad (5)$$

Due to the projection from the 3–D model space into the 2–D image space, the depth information and the identification of initial and end points get lost. In (3) we have seen, how the rotation, translation, and projection affects the densities for single point features. The lost identification of initial and end points is substituted by a random process. Possible orders of point pairs are assumed to be uniformly distributed. If two pairs of points are feasible, then we have

$$p(\boldsymbol{o}_k|\boldsymbol{a}_{\kappa,l}, \boldsymbol{R}, \boldsymbol{t}) = \frac{1}{2}\sum_{\tau}\prod_{s=1}^{2} \mathcal{N}(\boldsymbol{o}_{k,s}|\boldsymbol{R}\boldsymbol{\mu}_{\kappa,l,\tau(s)} + \boldsymbol{t}, \boldsymbol{R}\boldsymbol{\Sigma}_{\kappa,l,\tau(s)}\boldsymbol{R}^T), (6)$$

where τ covers all permutations. For two observable 2–D straight line features $\boldsymbol{o}_1$ and $\boldsymbol{o}_2$ the ratio $r \in \mathbb{R}$ between these straight line segments' length and the angle $\theta \in [0°; 360°]$ between them can be computed, and can be used for recognition and localization. The statistical modeling of these features can either be derived from (6) applying a complicated density transform, or from the computation of discrete probabilities for finite intervals, as suggested in [14].

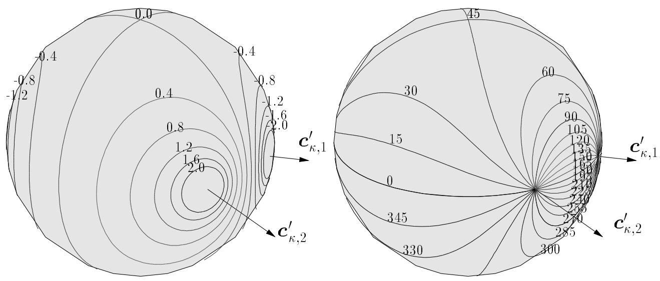

Figure 2: Iso–ratio (left, where $\log_2(r) \in [-2; 2]$) and iso–anlge (right) curve for two lines of equal length and with 45° between them.

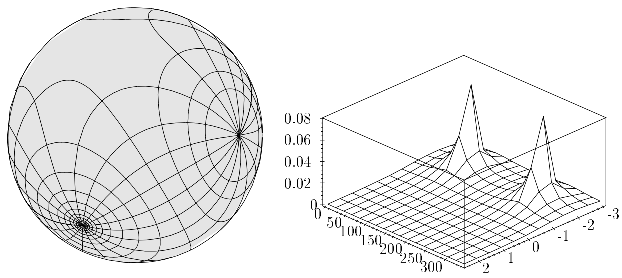

Figure 3: Tessellation of the viewing sphere (left) and the joint probability function (right)

3.1.3 Ratios and Angles

If $(\boldsymbol{c}_{\kappa,1}, \boldsymbol{c}_{\kappa,2}) \in \mathbb{R}^3 \times \mathbb{R}^3$ are the corresponding model lines of $\boldsymbol{o}_1$ and $\boldsymbol{o}_2$, and $\boldsymbol{v} \in \mathbb{R}^3$ denotes the viewing direction, we have the following constraints for r and θ:

$$|\boldsymbol{c}'_{\kappa,1} \times \boldsymbol{v}| - r|\boldsymbol{c}'_{\kappa,2} \times \boldsymbol{v}| = 0 \qquad (7)$$

and

$$(\boldsymbol{c}'_{\kappa,1} \times \boldsymbol{v})(\boldsymbol{c}'_{\kappa,2} \times \boldsymbol{v}) - \cos\theta |\boldsymbol{c}'_{\kappa,1} \times \boldsymbol{v}||\boldsymbol{c}'_{\kappa,2} \times \boldsymbol{v}| = 0, \quad (8)$$

where $\boldsymbol{c}'_{\kappa,s} = \boldsymbol{c}_{\kappa,s,1} - \boldsymbol{c}_{\kappa,s,2}$, $s = 1, 2$. For the normalized viewing direction $\boldsymbol{v} \in \mathbb{R}^3$ with $\|\boldsymbol{v}\|^2 = 1$, iso–ratio and iso–angle curves can be defined. Examples for these curves are shown in Figure 2, and their computation is done by an algorithm for tracing algebraic curves which relies on homotopy continuation [14].

In order to compute the probability for observing a set $\boldsymbol{O} = \{\boldsymbol{o}_1, \ldots, \boldsymbol{o}_n\}$ of features, the density function for each measured value $\boldsymbol{o}_k$ in the image is required. For a pair $(\log_2(r), \theta)$ of ratio and angle this probability depends on the viewing direction $\boldsymbol{v}$ or on the object's pose in the scene. Instead of using parametric densities including parameters of the feature's geometry and its pose, discrete probabilities will be computed for each viewing direction. The $(\log_2(r), \theta)$–space can be quantized on a rectangular grid. For each rectangle, the average value of the discrete probabilities is computed and stored in a table. Figure 3 shows the tessellation of the sphere (left) and the derived 2–D probability function (right). Each region in the parameter space is limited by upper and lower bounds for ratios and angles. Because of (7) and (8) the mentioned boundary values define polygons on the sphere, whose area A results from

$$A = \sum_{i=1}^{z} \alpha_i - (z-2)\pi \quad , \qquad (9)$$

where z is the number of vertices of the polygon, and $\alpha_i \in [0; 2\pi]$ is the spherical angle between two adjacent angles. To obtain a probability, the area A is divided by 4π, the area of the complete sphere. Such a table can be computed for each ratio–angle–pair. Thus for each observed feature pair and a given viewing direction the discrete probability can be computed.

So far, we have seen that different types of features result in different statistical representations. Both parametric, and non–parametric densities can be used for object recognition purposes. If the correspondence of model and image features is known, the introduced probability density functions can be used for the computation of the probability for observing a set of features.

3.2 Statistical Modeling of Feature Correspondences

If no background features exist, each observable feature $\boldsymbol{o}_k$, $1 \leq k \leq m$, corresponds to a model feature $\boldsymbol{c}_{\kappa,l}$, where κ is the object's class number. This matching can be denoted by a m–dimensional random vector $\boldsymbol{\zeta}_\kappa$, whose k–th component is the index l of the corresponding feature $\boldsymbol{c}_{\kappa,l}$ of $\boldsymbol{o}_k$. If point features are used, for example, $\boldsymbol{o}_k$ represents a 2–D image point and $\boldsymbol{c}_{\kappa,l}$ denotes the corresponding 3–D model point, then the correspondence is given by $\zeta_\kappa(\boldsymbol{o}_k) = l$, and $p(\zeta_\kappa(\boldsymbol{o}_k) = l)$ is the probability that $\boldsymbol{o}_k$ corresponds to $\boldsymbol{c}_{\kappa,l}$. For each correspondence ζ_κ, which assigns a set $\boldsymbol{O} = \{\boldsymbol{o}_1, \ldots, \boldsymbol{o}_m\}$ of observed features to corresponding model features $\boldsymbol{C}_\kappa = \{\boldsymbol{c}_{\kappa,1}, \ldots, \boldsymbol{c}_{\kappa,n}\}$, a joint discrete probability

$$p(\boldsymbol{\zeta}_\kappa) = \prod_{k=1}^{m} p(\zeta_\kappa(\boldsymbol{o}_k)) \qquad (10)$$

can be computed, and each hypothesized correspondence can be weighted by a statistical measure.

3.3 Construction of Model Densities

Combining the probability density function for the observed features $\boldsymbol{O}$ and the correspondence ζ_κ between model and image features, we get the density function

$$p(\boldsymbol{O}, \zeta_\kappa | \boldsymbol{B}_\kappa, \boldsymbol{R}, \boldsymbol{t}) = p(\zeta_\kappa) \prod_{k=1}^{m} p(\boldsymbol{o}_k | \boldsymbol{a}_{\kappa,\zeta_\kappa(\boldsymbol{o}_k)}, \boldsymbol{R}, \boldsymbol{t}) \quad , \qquad (11)$$

i.e. a probability for the appearance of $\boldsymbol{O}$ and ζ_κ for a given set of pose parameters. Here $\boldsymbol{B}_\kappa$ summarizes the discrete probabilities for correspondences and the parameters $\boldsymbol{a}_{\kappa,l}$, $1 \leq l \leq n_\kappa$. Usually, the correspondences between model and image features are latent. In [7] the elimination of the missing correspondence by marginalization is suggested. This is possible due to the probabilistic modeling of the correspondence between involved features. Other approaches, however, prefer the ranking of match hypotheses based on probability measures [15].

Assuming that the correspondence is not part of the observation, the joint probability density function for observing $\boldsymbol{O}$ is given by the marginal density

$$p(\boldsymbol{O} | \boldsymbol{B}_\kappa, \boldsymbol{R}, \boldsymbol{t}) = \sum_{\zeta_\kappa} p(\zeta_\kappa) \prod_{k=1}^{m} p(\boldsymbol{o}_k | \boldsymbol{a}_{\kappa,\zeta_\kappa(\boldsymbol{o}_k)}, \boldsymbol{R}, \boldsymbol{t}) \quad . \qquad (12)$$

Thus available image features are judged without knowing the correspondences. This is consequently at least democratic in that all correspondences are considered, and the density function is not tied to a special ζ_κ.

The formalized statistical description of objects can be applied for solving the object recognition problem, if the available image features and their statistical behavior is known. A central problem thus is the estimation of model parameters or of discrete probabilities, respectively.

4 Model Acquisition

Since the statistical representation of features depends on the geometric structure of image primitives, model generation methods differ for different types of features, and are related to used projection models.

4.1 Normally Distributed Features

If normally distributed 3–D features are assumed, 3–D mean vectors and (3×3)–covariance matrices as well as discrete probabilities for correspondences have to be estimated. The sample data consist of a representative set of 2–D views (see Figure 4 for example views). For each sample view the pose parameters are known, because a calibrated camera mounted on a robot's hand is used. Due to the missing depth and the non–observable assignment between model and image features, we have to deal with an incomplete data estimation problem in a fairly high–dimensional search space. An established method for solving estimation problems based on incomplete observations is the Expectation Maximization algorithm (EM algorithm) [3]. In [7, 8] a detailed derivation and discussion of iterative parameter estimation algorithms for normally distributed point and line features are discussed. These algorithms require no heuristics for computing correspondences of features in different views. Model generation works unsupervised with respect to the missing correspondence. The model parameters are automatically computed using training images and segmentation results from different views.

4.2 Estimation of Ratio–Angle Probabilities

The estimation of discrete probabilities for ratio and angle features requires that we take into consideration each pair of edges in the model space. These 3–D edges have to be known in advance, i.e. 3–D models of involved objects are necessary for computing the statistical behavior of image features. In contrast to normally distributed image features, the components of model densities cannot be estimated from a set of 2–D views. The 3–D structure of considered objects has to be known in advance. In practice, an aspect graph and a suitable tessellation of the viewing sphere

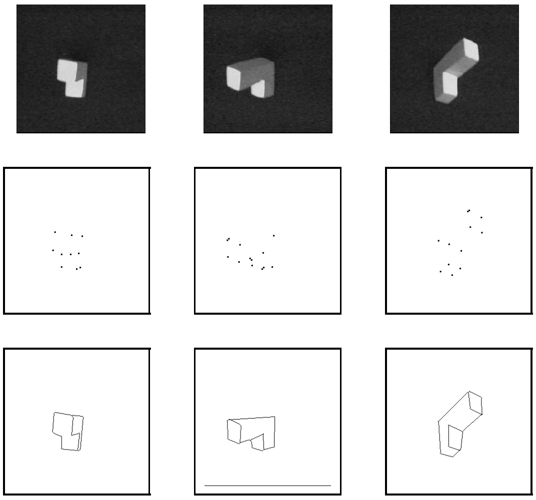

Figure 4: Examples for 2–D training views and the computed point and line features.

is used to generate the table of discrete probabilities for each pair of edges off–line [15]. An example is shown in Figure 5.

5 Statistical Object Recognition and Localization

5.1 Classification

The classification is generally based on the Bayesian decision rule (1). The observation here is restricted to the set of image features, i.e. points, lines, angles, or ratios. When the object's pose is known, for each object class the a priori probability is given, and the marginal density (12) is available, (1) can be applied for the class decision.

Due to the statistical modeling of correspondences, it is also possible to rank match hypotheses. Hypotheses can be listed and tested during the recognition stage. The classification is related to a search in the correspondence space, and the judgement of hypothesized correspondences is done using $p(\zeta_\kappa(\boldsymbol{o}_k))$ for observed features, the probability for feature correspondences.

5.2 Localization

The computation of pose parameters from three points or a trihedral corner has been extensively studied [5, 6]. The pose parameters consist of three rotation angles and three components of translation vectors. If the representation of observed features in the model is done by parametric densities, the localization of objects corresponds to a parameter estimation problem based on a smooth density function. Independently of feature correspondences, the pose, characterized by the parameters $\boldsymbol{R}$ and $\boldsymbol{t}$, can be determined by the maximum likelihood estimation

$$\{\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{t}}\} = \operatorname*{argmax}_{\boldsymbol{R},\boldsymbol{t}} p(\boldsymbol{O}|\boldsymbol{B}_\kappa, \boldsymbol{R}, \boldsymbol{t}) \qquad (13)$$

for a given set $\boldsymbol{O}$ of observed features. The maximization can be done by standard optimization techniques [12].

In contrast to the reduction of pose estimation to a continuous optimization problem, the use of ratio–angle pairs and their discrete probabilities enforce the solution of discrete optimization tasks controlled by hypothesized correspondences of model and image features. Since the rotation angles and the translation vector are not explicitly represented as discrete probabilities, pose computation for the weak perspective projection model is divided into the following parts: the computation of the viewing direction, the rotation angle of the image about the viewing direction, the scale factor, and the (two–dimensional) translation vector.

With (7) and (8) it follows that each measured angle or ratio imposes a one–dimensional constraint on possible viewing directions. In order

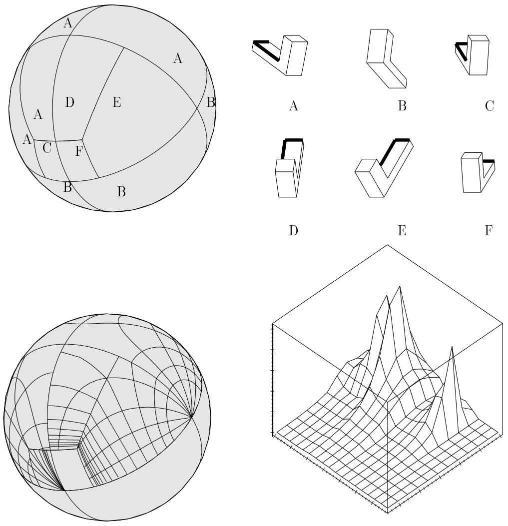

Figure 5: Tessellation, aspect graph, iso–ratio/iso–angle curves, and the computed probability for the marked pair of edges

3–D	Recognition [%]		Run Time [sec]	
	points	lines	points	lines
Ω_1	47	44	466	1882
Ω_2	78	82	485	2101
Ω_3	58	36	465	1933
Ω_4	89	76	471	1520
average	68	59	472	1859

Table 1: Recognition rates and run time for a test set of 1600 images including objects of Figure 6

to determine the viewing direction at least two constraints are needed. Once the viewing direction is known the remaining components can be recovered using standard 2–D pose estimation [2].

6 Experimental Results

The algorithms were developed and tested on different platforms and different objects.

6.1 Recognition Results using Point and Line Features

The experiments using normally distributed point and line features apply the statistical approach to learning, localization, and classification of objects in gray–level images. Figure 8 and Figure 9 show some example images. These results prove that the algorithms work with the presence of occlusion and uncertainty (cmp. Figure 7 for the occlusion of a 2–D object). Table 1 summarizes recognition results and run times on a HP 7000/735.

Ω_1 Ω_2 Ω_3 Ω_4

Figure 6: Polyedral 3–D objects

6.2 Recognition Results using Ratio and Angle Features

The model database consists of five polyhedral 3–D objects. We extract edges from the gray–level images using the Canny edge detector and

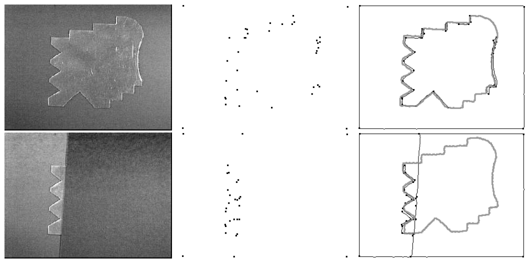

Figure 7: Localization and occlusion

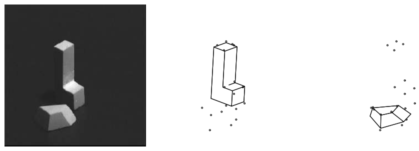

Figure 8: Scene including two polyhedral objects

detect lines. Based on these line features, ratio and angle feature sets are computed. For each feature set match hypotheses and their probabilities were derived from pre–computed look–up tables. Table 2 summarizes the run time (Sun Sparc 10) for the examples shown in Figure 10, 11, and 12.

7 Summary and Conclusion

This paper presented a probabilistic approach to solve the 3–D object recognition problem. The statistical modeling of objects is based on composed density functions including discrete probabilities for feature correspondences, and discrete or continuous probability density functions for modeling the statistical behavior of image features dependent on the viewing direction. Considered features were point and line features with parametric densities and scalar measures derived from line features, which were characterized by discrete probabilities. It is shown, how differ-

Figure 9: Object scene including background features

Figure	Feature Sets	Match Hypotheses	Total Run Time
10	25	3134	68 min
11	11	1516	12 min
12	16	15672	26 min

Table 2: Experimental results using ratio and angle features

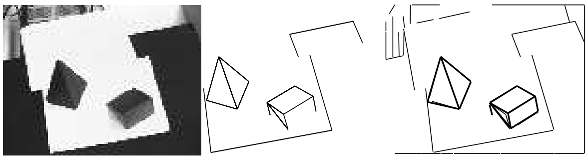

Figure 10: Classification and localization results using ratio, and angle features

ent types of model densities can be applied for object recognition and pose estimation. Experimental results demonstrate the correctness of the chosen statistical framework. Future research should concentrate on more efficient localization algorithms and extensions for the recognition of more complex objects in natural scenes instead of simple polyhedral objects. Furthermore, an objective comparison of different recognition algorithms should not only discuss theoretical relations, but should also be tested on common sample data. The development of a suitable test set is planned.

8 References

1. T. W. Anderson. *An Introduction to Multivariate Statistical Analysis.* Wiley Publications in Statistics. John Wiley & Sons, Inc., New York, 1958.
2. N. Ayache and H. Faugeras. HYPER: a new approach for the recognition and positioning of two–dimensional objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):44–55, January 1986.
3. A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, 39(1):1–38, 1977.

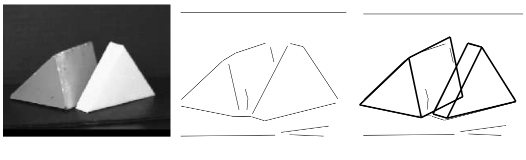

Figure 11: Localization and classification

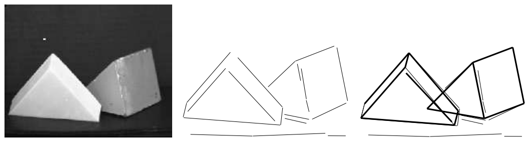

Figure 12: Localization and classification

4. L. Devroye, L. Györfi, and G. Lugosi. *A Probabilistic Theory in Pattern Recognition*, volume 31 of *Applications of Mathematics, Stochastic Modelling and Applied Probability.* Springer, Heidelberg, 1996.
5. M. Dhome, M. Richetin, J.-T. Lapresté, and G. Rives. Determination of the attitude of 3–D objects from a single perspective view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(12):1265–1278, Dezember 1989.
6. R. M. Haralick, C. Lee, K. Ottenberg, and M. Nölle. Review and analysis of solutions of the three point perspective pose estimation problem. *International Journal of Computer Vision*, 13(3):331–356, December 1994.
7. J. Hornegger. *Statistische Modellierung, Klassifikation und Lokalisation von Objekten.* Shaker, Aachen, 1996.
8. J. Hornegger and H. Niemann. Statistical learning, localization, and identification of objects. In ICCV 95 [9], pages 914–919.
9. *Proceedings of the* 5^{th} *International Conference on Computer Vision (ICCV)*, Boston, June 1995. IEEE Computer Society Press.
10. A. K. Jain and P. J. Flynn, editors. *Three–Dimensional Object Recognition Systems*, Amsterdam, 1993. Elsevier.
11. H. Niemann. *Pattern Analysis and Understanding.* Springer, Heidelberg, 1990.
12. W.H. Press, B.P. Flannery, S. Teukolsky, and W.T. Vetterling. Numerical recipes - the art of numerical computing, C version. Technical report, 35465-X, 1988.
13. K. B. Sarachik. An analysis of the effect of Gaussian error in object recognition. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, AI Lab., Massachusetts, 1994.
14. I. Shimshoni. Interpreting images of polyhedral objects in the presence of uncertainty. PhD thesis, Department of Computer Science, University of Illinois, Urbana–Champaign, 1995.
15. I. Shimshoni and J. Ponce. Probabilistic 3D object recognition. In ICCV 95 [9], pages 488–493.
16. W. M. Wells III. Statistical Object Recognition. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, AI Lab., Massachusetts, February 1993.
17. G. Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Methoden*, volume 27 of *Applications of Mathematics.* Springer, Heidelberg, 1995.