A Probabilistic Framework for Grouping Image Features

Rebecca L. Castaño Seth Hutchinson
becky@cs.uiuc.edu seth@uiuc.edu

Dept. of Electrical and Computer Engineering and The Beckman Institute
University of Illinois at Urbana-Champaign
Urbana, IL 61801

Abstract

We present a framework for determining probability distributions over the space of possible image feature groupings. Such a framework allows higher level processes to reason over many plausible perceptual groupings in an image, rather than committing to a specific image segmentation in the early stages of processing.

We first derive an expression for the probability that a set of features should be grouped together, conditioned on the observed image data associated with those features. This probability measure formalizes the principle that features in an image should be grouped together when they participate in a common underlying geometric structure. We then present a representation scheme in which only those groupings with high probability are explicitly represented, while large sets of unlikely grouping hypotheses are implicitly represented. We present experimental results for a variety of real intensity images.

1 Introduction

The perceptual grouping problem is concerned with determining how features should be organized into more abstract structures, to be used by higher level visual processes. The Gestalt psychologists claimed that humans group features based on several principles, including: proximity, symmetry, continuation, closure, and familiarity [12]. The work of the Gestalt psychologists has inspired a number of computer vision approaches to perceptual grouping [1] [9] [8] [11] (for a review of this work see[11] and [3]). Typically, these computational approaches rely on thresholds (e.g., a threshold on the difference between orientation of line segments [9], or on the linking radius described in [1]), or on certainty measures that are derived in an ad hoc fashion (e.g., basing the certainty of grouping two line segments on the proximity of their endpoints [8], or using a decaying exponential to define a probability of termination for line segments [4]).

In this paper we present a probabilistic approach to perceptual grouping based on the grouping principle that *features in an image should be grouped together when they participate in a common underlying geometric structure* (specific instantiations of this principle are described in Sections 2.1 and 4). Several of the Gestalt laws for grouping are instances of this principle, e.g., proximity, symmetry, and continuation. The principle is also consistent with the idea of nonaccidentalness, that is, the perception of structure that has a low probability of occurring only by random chance implies a single cause for the structure [8] [10] [13].

To realize our grouping principle we have adapted the framework developed in [7] for region-based image segmentation. In Section 2 we define the models required by the framework: a parameterized model of geometric structures (*parameter space*), a characterization of how well a set of observed image features fits to a particular geometric structure (*observation space*), and a probabilistic model that describes the image formation process and its effects on the distribution of features in the image (*degradation model*).

Equipped with these models, in Section 3 we derive the probability that a set of features should be grouped together, conditioned on the observed image data associated with those features. Computation of this probability does not rely on parameter estimation, thereby avoiding problems associated with estimation-based methods, (e.g., degraded performance with small data sets, and lack of robustness due to outliers).

The number of possible feature groupings for a typical image is intractably large. Thus the formalism proposed in Section 5 for representing all possible feature groupings in an image explicitly represents groupings that have high probability, while implicitly representing large sets of groupings that have small probability.

The results of using these methods and probabilistic models on actual intensity image data are given in Section 6. In the final section, Section 7, conclusions are drawn and directions for future work are suggested.

2 Probabilistic Models

The models introduced in Section 1 will now be formalized. Researchers have used parts of these models in the context of region segmentation (e.g., [6]).

2.1 Parameter space

Here we consider geometric structures with finite parameterizations. The corresponding parame-

0-8186-7190-4/95 $4.00

ter space is defined as a vector of random variables, $\mathbf{U_k} = [U_k^1\ U_k^2\ \ldots\ U_k^r]$. A vector value that $\mathbf{U_k}$ can take on is denoted by $\mathbf{u}_k$.

As an example, suppose that we wish to find all segments in an image that are part of the same straight line. Our geometric structure, a straight line, can be parameterized by a pair (θ, d), where θ represents the normal of the line and d represents the normal distance from the origin of the reference frame to the line. Thus the parameter space is $\mathbf{U} : [0, \pi) \times \mathbf{R}$, representing all straight lines.

2.2 Observation space

The set of all features in an image is denoted by $\mathcal{S}$. Each feature, S_k, consists of a set of image data points, $S_k = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{|S_k|}\}$, each of which can be represented by $\mathbf{x}_i = \langle x_i, y_i \rangle$. Since noise is introduced into the system during the imaging, digitization and edge extraction processes, each component of $\mathbf{x}_i$ is modeled as a random variable.

To evaluate the probability that a set of features should be grouped together, it is necessary to evaluate how well the features fit to a single geometric structure. For this purpose we define an observation space.

Let $\mathbf{U}_k$ denote the parameter space for a particular grouping of features that contains the feature S_k. Let $\mathbf{S}_k$ denote the set of all random variables associated with the data points in feature S_k. Let $\psi_k^i(\mathbf{S}_k, \mathbf{U}_k)$ be a function of the random variables in $\mathbf{S}_k$ and the parameter space, $\mathbf{U}_k$. The observation space of the set S_k is defined as the vector of random variables, $\mathbf{Y}_k = [\psi_k^1, \psi_k^2, \ldots, \psi_k^q]$. A vector value that $\mathbf{Y}_k$ may take on is denoted $\mathbf{y}_k$.

Continuing the example using straight lines as the geometric structures, a point, $\mathbf{u}_k = (\theta_k, d_k)$, in the parameter space defines a line in the image plane. Let $\phi(\mathbf{u}_k) = 0$ be the implicit mapping of points in the parameter space to a line in the image plane. Let $\delta(\mathbf{x}_i, \phi(\mathbf{u}_k))$ be the signed distance of the data point $\mathbf{x}_i$ to the line described by the zero set of $\phi(\mathbf{u}_k)$,

$$\delta(\mathbf{x}_i, \phi(\mathbf{u}_k)) = x_i cos\theta_k + y_i sin\theta_k - d_k. \quad (1)$$

We define the observation $\mathbf{y}_k(S_k, \mathbf{u}_k)$, to be a vector in which each component is the distance of a data point to the line $\phi(\mathbf{u}_k)$,

$$\mathbf{y}_k(S_k, \mathbf{u}_k) = [\delta(\mathbf{x}_1, \phi(\mathbf{u}_k)), \cdots \delta(\mathbf{x}_{|S_k|}, \phi(\mathbf{u}_k))]. \quad (2)$$

2.3 Degradation model

The imaging process introduces uncertainty into the measured data. We model this uncertainty by a conditional probability density function, $f(\mathbf{y}_k|\mathbf{u}_k)$, called the degradation model, which represents the conditional density of an observation vector, $\mathbf{y}_k$, given the parameter value, $\mathbf{u}_k$, of the underlying geometric structure. Since each data point $\mathbf{x}_i$ is assumed conditionally independent of every other data point given the parameter vector $\mathbf{u}_k$,

$$f(\mathbf{y}_k|\mathbf{u}_k) = \prod_{\mathbf{x}_i \in S_k} f(\delta(\mathbf{x}_i, \phi(\mathbf{u}_k))|\mathbf{u}_k). \quad (3)$$

The degradation model that we have selected is i.i.d. Gaussian with zero mean and variance σ^2. Each observed point is assumed to be displaced along a line perpendicular to the ideal line by an amount characterized by the Gaussian distribution. Using the Gaussian distribution in (3), we can write the degradation model for lines as

$$f(\mathbf{y}_k|\mathbf{u}_k) = \prod_{\mathbf{x}_i \in S_k} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{\delta(\mathbf{x}_i, \phi(\mathbf{u}_k))}{\sigma}\right)^2} \quad (4)$$

Equation (4) can be used to quantify the deviation of a set of image data points from a line with given parameter $\mathbf{u}_k$.

2.4 Prior model

The joint pdf of $\mathbf{u}_k$, $f(\mathbf{u}_k)$, is called the prior model. The prior model is a density over the parameter space that represents the expected distribution of features over the space. We will assume that all feature parameter values are equally likely. Thus, $f(\mathbf{u}_k)$ will be a uniform distribution over the parameter space.

3 Probability of Feature Groupings

Using the models defined in Section 2, we now develop an expression for the probability that a set of features, $\{S_{\alpha_i}, .., S_{\alpha_n}\}$, should be grouped together. This hypothesis is represented by $H_{\alpha_1,..,\alpha_n}$ where $S_{\alpha_i} \in \mathcal{S}$. Note that the hypothesis H_{12} is equivalent to asserting that S_1 and S_2 fit the same geometric structure and therefore that $\mathbf{u}_1 = \mathbf{u}_2$. We denote by $P(H_{ij}|\mathbf{y}_i, \mathbf{y}_j)$ the probability that the features S_i and S_j associated with a pair of observations $\mathbf{y}_i$ and $\mathbf{y}_j$, should be grouped together.

If the parameter value associated with a grouping is given, the degradation model completely describes the density of the corresponding observations, which implies

$$f(\mathbf{y}_1, \mathbf{y}_2|\mathbf{u}_1, \mathbf{u}_2) = f(\mathbf{y}_1|\mathbf{u}_1) f(\mathbf{y}_2|\mathbf{u}_2). \quad (5)$$

Thus, if the parameter value of the grouping to which a feature belongs is known, no other observations will affect the density of the observation of that feature.

We assume that $\mathbf{Y}_1$ is marginally independent of $\mathbf{Y}_2$ given $\neg H_{12}$, i.e.,

$$f(\mathbf{y}_1, \mathbf{y}_2|\neg H_{12}) = f(\mathbf{y}_1) f(\mathbf{y}_2). \quad (6)$$

For instance segments that are not on the same line are unrelated (independent). Other models of the relationship between observations that do not belong to the same geometric structure could also be used.

By applying Bayes' rule and (6) the probability that two features participate in the same geometric structure can be written as

$$P(H_{12}|\mathbf{y}_1,\mathbf{y}_2) = \left(1+\frac{P(\neg H_{12})}{P(H_{12})}\frac{f(\mathbf{y}_1)f(\mathbf{y}_2)}{f(\mathbf{y}_1,\mathbf{y}_2|H_{12})}\right)^{-1} \quad (7)$$

The first ratio in (7), $\frac{P(\neg H_{12})}{P(H_{12})}$ represents the prior probability that two features should be grouped together while the other ratio in the denominator of (7),

$$\frac{f(\mathbf{y}_1)f(\mathbf{y}_2)}{f(\mathbf{y}_1,\mathbf{y}_2|H_{12})}, \quad (8)$$

represents the effect that the observed data have on the probability of a grouping hypothesis.

The marginal pdf $f(\mathbf{y}_1)$ can be written as

$$f(\mathbf{y}_1) = \int f(\mathbf{y}_1|\mathbf{u}_1)f(\mathbf{u}_1)d\mathbf{u}_1. \quad (9)$$

This expression for $f(\mathbf{y}_1)$ involves the degradation and prior models defined in Section 2.3.

If H_{12} is *true*, then $\mathbf{y}_1$ and $\mathbf{y}_2$ participate in the same geometric structure, i.e., $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{u}_{12}$. Using (5) we may write

$$f(\mathbf{y}_1,\mathbf{y}_2|H_{12}) = \int f(\mathbf{y}_1|\mathbf{u}_{12})f(\mathbf{y}_2|\mathbf{u}_{12})f(\mathbf{u}_{12})d\mathbf{u}_{12}.$$

We may now rewrite (8) in terms of the degradation and prior models as

$$\frac{\int f(\mathbf{y}_1|\mathbf{u}_1)f(\mathbf{u}_1)d\mathbf{u}_1 \int f(\mathbf{y}_2|\mathbf{u}_2)f(\mathbf{u}_2)d\mathbf{u}_2}{\int f(\mathbf{y}_1|\mathbf{u}_{12})f(\mathbf{y}_2|\mathbf{u}_{12})f(\mathbf{u}_{12})d\mathbf{u}_{12}}. \quad (10)$$

For the line grouping application substitute the degradation model of (4) into (9) to obtain an expression for $f(\mathbf{y}_k)$ as follows:

$$f(\mathbf{y}_k) = \int \frac{1}{(2\pi\sigma^2)^{\frac{|S_k|}{2}}} e^{-\frac{1}{2}\sum_{\mathbf{x}_i\in S_k}\left(\frac{\delta(\mathbf{x}_i,\phi(\mathbf{u}_k))}{\sigma}\right)^2} f(\mathbf{u}_k)d\mathbf{u}_k. \quad (11)$$

Using (11) in (7) we can determine the probability that any two sets of points are samples of the same line.

Our experiments require computing the probability of the hypothesis that n features are consistent. We address these computations, and the numerical computations of the integrals, in [3].

4 Bilateral Symmetries

While features that have exactly the same geometric structure can be grouped, the framework also allows for a more general application - that of grouping features that participate in geometric structures with common characteristics. In this situation some, but not all of the parameters of the features are the same. We demonstrate such an application by determining groups of segments that participate in bilateral symmetries with parallel axes.

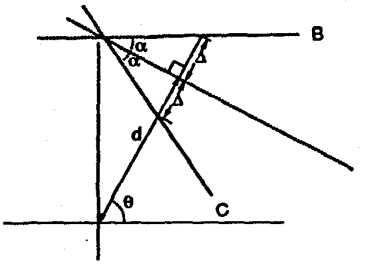


Figure 1: Symmetry parameterization.

The parameter space used for the symmetries is $\{(\theta,d,\alpha,\Delta) : \theta \in [0,2\pi), d,\Delta > 0, \alpha \in [0,\pi)\}$, where symmetry contours have been restricted to straight lines. Two parameters, θ and d, define the axis of symmetry while α and Δ define the sweeping rule, as shown in Figure 1.

To establish an observation model corresponding to bilateral symmetries we first note that a point (θ,d,α,Δ) in the parameter space defines two lines in the image plane labeled as lines B and C in Figure 1. These lines can be expressed implicitly using $\phi_1(\mathbf{u}_k)$ and $\phi_2(\mathbf{u}_k)$. We have

$$\phi_{1,2}(\mathbf{u}_k) = \{(x,y) : x\,cos(\theta_k \pm \alpha_k) + y\,sin(\theta_k \pm \alpha_k) - (d_k \pm \Delta_k)\,cos(\alpha_k) = 0\}. \quad (12)$$

Let $\delta(\mathbf{x}_i,\phi_j(\mathbf{u}_k))$ be the displacement of the point $\mathbf{x}_i$ from the line $\phi_j(\mathbf{u}_k)$ where $j \in \{1,2\}$.

The observation for two line segments S_{k_1} and S_{k_2} that form a symmetric relationship is defined as

$$\mathbf{y}_k(S_{k_1},S_{k_2},\mathbf{u}_k) = [\ \delta(\mathbf{x}_{1_1},\phi_1(\mathbf{u}_k)),\ \cdots,\delta(\mathbf{x}_{1_{|S_{k_1}|}},\phi_1(\mathbf{u}_k)), \delta(\mathbf{x}_{2_1},\phi_2(\mathbf{u}_k)),\ \cdots,\delta(\mathbf{x}_{2_{|S_{k_2}|}},\phi_2(\mathbf{u}_k))\ ] \quad (13)$$

To compare symmetries with the same axis, but different sweeping rules, it is necessary to generalize the notion of the hypothesis. Recall that H_{12} represented the hypothesis that $\mathbf{u}_1 = \mathbf{u}_2$, i.e. $\theta_1 = \theta_2$, $\alpha_1 = \alpha_2$, $d_1 = d_2$, and $\Delta_1 = \Delta_2$. We now define H'_{12} to be the hypothesis that $\theta_1 = \theta_2$ and $d_1 = d_2$. Since we are interested only in collinear axes, no restrictions will be placed on the sweeping rules, that is, the relationship of α_1 to α_2 or the relationship of Δ_1 to Δ_2.

The expression for $f(\mathbf{y}_1,\mathbf{y}_2|H'_{12})$ is

$$\int f(\mathbf{y}_1|(\theta_{12},d_{12},\alpha_1,\Delta_1))f(\mathbf{y}_2|(\theta_{12},d_{12},\alpha_2,\Delta_2)) f(\theta_{12},d_{12},\alpha_1,\Delta_1,\alpha_2,\Delta_2)d\mathbf{u}_1d\mathbf{u}_2.$$

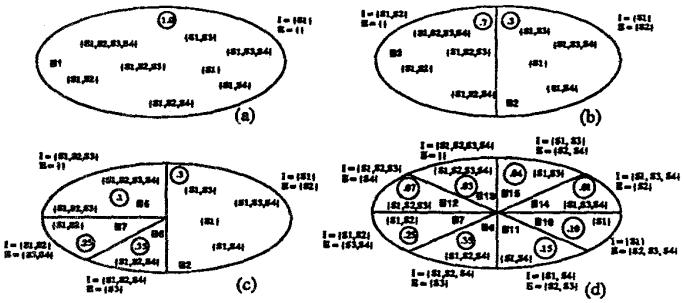


Figure 2: Covers of π_1.

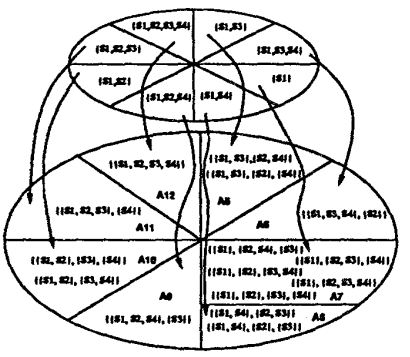


Figure 3: Π cover induced by π_1 cover.

The sweeping rules are independent, so we can separate the integrals over α and Δ to obtain

$$= \int_0^{2\pi} \int_{-\infty}^{\infty} \Big\{ \int_{-\pi}^{\pi} \int_{-\infty}^{\infty} f(\mathbf{y}_1|(\theta_{12}, d_{12}, \alpha_1, \Delta_1)) f(\theta_{12}, d_{12}, \alpha_1, \Delta_1) d\alpha_1 \, d\Delta_1 \int_{-\pi}^{\pi} \int_{-\infty}^{\infty} f(\mathbf{y}_2|(\theta_{12}, d_{12}, \alpha_2, \Delta_2)) f(\theta_{12}, d_{12}, \alpha_2, \Delta_2) d\alpha_2 \, d\Delta_2 \Big\} dd_{12} \, \theta_{12}. \quad (14)$$

Equation (14) can be used in (7) to determine $P(H'_{12}|\mathbf{y}_1, \mathbf{y}_2)$, the probability that a pair of bilateral symmetries have collinear axes. Our results use a further generalization to group bilateral symmetries with parallel axes.

5 Computing Distributions

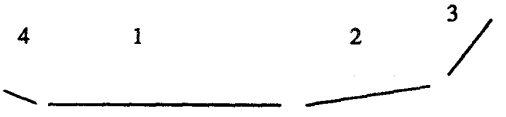
In Section 5.1, we develop a framework for obtaining distributions over feature groupings. Then, in Section 5.2, we use the framework for feature groupings in developing a framework for obtaining distributions over partitions of the feature data sets.

The idea behind the framework is that the probability density will be highly concentrated in a small portion of the space of possible results (partitions or feature groupings). We begin with a very coarse representation of the probability distribution over the entire space and successively refine the representation in the area of highest density until a sufficient approximation to the probability distribution has been obtained. As we shall see, the condition for a sufficient approximation is well-defined.

5.1 Distributions over Feature Groupings

We will demonstrate the incremental refinement procedure used to determine the distribution of probability via an example; the concepts presented in the example are formally defined in [3].

Suppose that the following line segments have been extracted from an image; thus, $\mathcal{S} = \{S_1, S_2, S_3, S_4\}$.



Let π_i be the set of all feature groupings containing S_i. A subset of π_i is called an event. The space π_1 is represented in Figure 2(a) and is labeled event B_1.

The space of π_1 can be divided into two events (sets) as shown in Figure 2(b). One event, B_2, consists of all feature groupings that contain S_1, but do not contain S_2, i.e. all groupings for which $H_{1,2}$ is *false*. In our example, the probability of B_2 is .7. The other event, B_3, consists of all feature groupings that contain both S_1 and S_2, i.e. all groupings for which $H_{1,2}$ is *true*. The events B_2 and B_3 form a partition of π_1, referred to as a cover of π_1. Thus, $P(B_3) = 1 - P(B_2) = .3$.

Each event can be completely specified by two sets of features rather than enumerating every grouping in the event. The first set, I, contains the features that are included in every grouping in the event while the second set, E, contains those features that are not in any grouping in the event.

Individual feature groupings correspond to pairwise disjoint events; thus the probability of each event in the cover is the sum of the probabilities of the individual feature groupings contained in the event. In Figure 2(c), the feature grouping B_6 has a probability greater than the other three events in the cover. It is not necessary to refine events B_2 and B_5 further in order to know that no feature grouping in either of these events has a greater probability than that of B_6. Thus, those sets with high probability can be refined to individual feature groupings without ever spending time exploring parts of the cover with very low probability.

5.2 Partitions of Feature Groupings

The method of determining distributions on partitions of $\mathcal{S}$ builds on the refinement procedure just described. Let Π be the set of all possible partitions of $\mathcal{S}$. A subset of Π is called a partition event. A partition event that has only one element represents a single partition of $\mathcal{S}$ and is called a ground event.

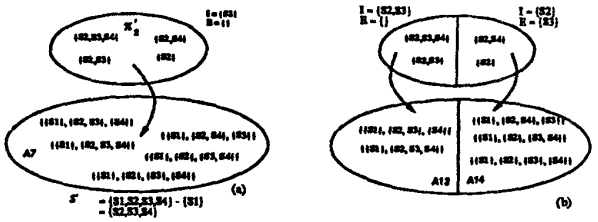


Figure 4: Refinement of π_2'.

Each cover on π_i induces a cover on Π since every partition in Π contains exactly one feature grouping that includes S_i. For example, the cover shown in Figure 2(d), divides Π into eight sets, as shown in Figure 3.

A ground event in π_1 represents one feature grouping in a partition, but there may be many partitions that include a particular feature grouping. For example, in Figure 3, all partitions in A_7 contain the feature grouping $\{S_1\}$.

The Π cover can be refined further by focusing on the feature grouping containing one of the features not in the π_1 ground event feature grouping. We can refine A_7 by looking at the feature grouping containing S_2. Let π_2', shown in Figure 4(a), represent all possible feature groupings which contain S_2, (but not S_1). The sample space π_2' is refined in the same way that π_1 was refined.

The set π_2' is formed using the elements of $\mathcal{S}' = \mathcal{S} - \{S_1\}$. When $\mathcal{S}^{(i)} = \emptyset$, all the features are in a grouping and a Π ground event has been reached.

6 Experimental Results

Experimental results are shown in Figures 5 - 13. The initial input to our algorithms was obtained by processing the image with a Canny edge detector [2], yielding edge contours represented as sets of connected points. These edges were divided into straight segments using an algorithm similar to the iterative endpoint fitting algorithm in [5]. Deviations of the points in a segment from linearity are treated as noise.

The distribution of most probable partitions gives an indication of how rapidly the probability of alternate partitions decreases. The similarities of the partitions may be noticed by a high level process and used in conjunction with additional information to make distinctions between the most probable partitions.

The most probable partition of line segments was used as input for the symmetry axis partitioning. Partitions of parallel axes are shown as a series of figures. All axes in a figure belong to the same group. Each figure shows the axes in black. For reference, the input lines are shown in gray. The algorithm groups sets of data, but does not estimate any symmetry axes. We have estimated the projection of the lines of symmetry onto the axis for visualization purposes.

This example, while very simple, demonstrates the information that can be obtained from the probabilities of the top several partitions. From the distribution of the top five partitions, it can be seen that the most likely partition is significantly more likely than any other partition. This is due to the fact that every input line carries significant information and leads us to conclude that no other partitions need be considered as viable alternative explanations of the data.

7 Conclusions and Future Research

We have presented a framework for determining the most probable partitions of data sets, representing image features, into feature groupings. The framework was demonstrated to perform well in two examples.

Other models of features and noise can be used in the framework. For example, different contour models such as B-splines could be used. For more complex models and high dimension parameter spaces, the required integrals become difficult to calculate, so simplifications and approximations may be necessary.

REFERENCES

[1] Michael Boldt, Richard Weiss, and Edward Riseman. Token-based extraction of straight lines. *IEEE Trans. Syst. Man Cybern.*, 19(6):1581–1594, Nov./Dec. 1989.

[2] John Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Machine Intell.*, 8(6):679–698, November 1986.

[3] Rebecca L. Castano. A probabilistic framework for grouping image features. Master's thesis, University of Illinois at Urbana-Champaign, 1994.

[4] I. J. Cox, J. M. Rehg, and S. Hingorani. A Bayesian multiple hypothesis approach to edge grouping and contour segmentation. Technical report, NEC, 1993.

[5] Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis.* John Wiley & Sons, Inc., New York, 1973.

[6] S. Geman. Experiments in Bayesian image analysis. *Bayesian Statistics*, 3:159–172, 1988.

[7] Steven M. LaValle and Seth Hutchinson. A framework for constructing probability distributions on the space of image segmentations. *Comput. Vision and Image Understanding*, 61(2):203–230, March 1995.

[8] D.G. Lowe. Three-dimensional object recognition from single two-dimensional images. *Artif. Intell.*, 31:355–395, 1987.

[9] R. Mohan and R. Nevatia. Perceptual organization for scene segmentation and description. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(6):616–635, June 1992.

[10] I. Rock. *The Logic of Perception.* MIT Press, Cambridge, MA, 1983.

[11] Sudeep Sarkar and Kim L. Boyer. *Computing Perceptual Organization in Computer Vision.* World Scientific Publishing Co., 1994.

[12] Max Wertheimer. Principles of perceptual organization. In D. Beardslee and M. Wertheimer, editors, *Readings in Perception*, pages 115–135. VanNostrand, Princeton, NJ, 1958.

[13] A. Witkin and J. Tenenbaum. On the role of structure in vision. In J. Beck, B. Hope, and A. Rosenfeld, editors, *Human and Machine Vision*, pages 481–543. Academic Press, 1983.

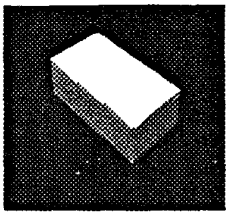


Figure 5: 512 x 512 image.

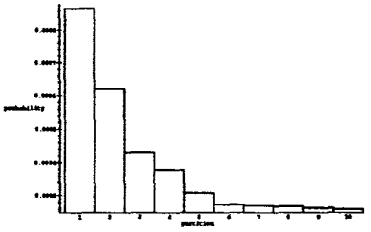


Figure 6: Probability of 10 most likely partitions.

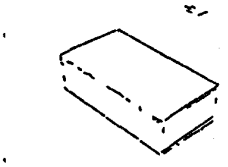


Figure 7: Edge detector output fitted with straight lines.

(a) (b) (c) (d)

Figure 8: Four most probable partitions. Figure (a) is the most probable.

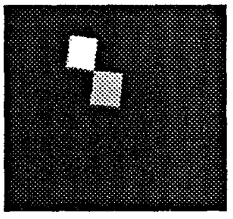


Figure 9: 512 x 512 image.

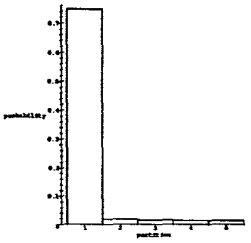


Figure 10: Probability of five most likely partitions.

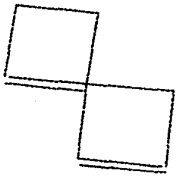
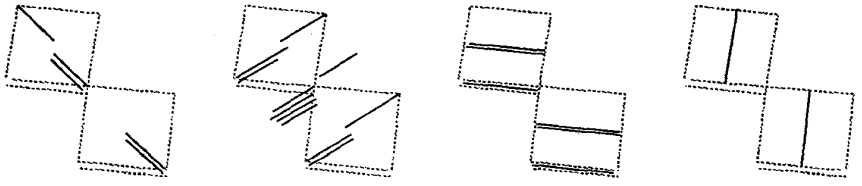
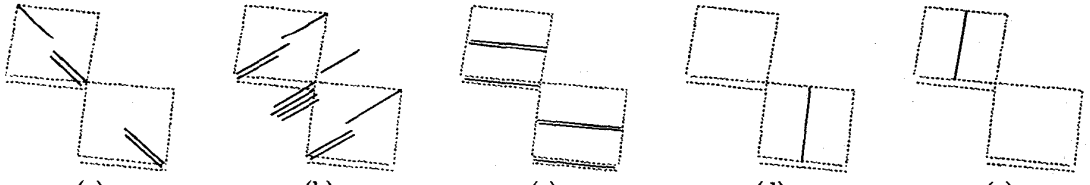


Figure 11: Straight edge segments. There are 9 lines.



(a) (b) (c) (d)

Figure 12: Parallel symmetry axes. This is the most probable partition of the 24 pairs of segments that have a significant overlap into sets of parallel axes. Each figure shows all the axes in one group of parallel axes.



(a) (b) (c) (d) (e)

Figure 13: Parallel symmetry axes. This is the second most probable partition. Each figure shows all the axes in one group of parallel axes.