A Computational Model of View Degeneracy

Sven J. Dickinson, *Member*, *IEEE*, David Wilkes, and John K. Tsotsos

Abstract—We quantify the observation by Kender and Freudenstein [24] that degenerate views occupy a significant fraction of the viewing sphere surrounding an object. For a perspective camera geometry, we introduce a computational model that can be used to estimate the probability that a view degeneracy will occur in a random view of a polyhedral object. For a typical recognition system parameterization, view degeneracies typically occur with probabilities of 20 percent and, depending on the parameterization, as high as 50 percent. We discuss the impact of view degeneracy on the problem of object recognition and, for a particular recognition framework, relate the cost of object disambiguation to the probability of view degeneracy. To reduce this cost, we incorporate our model of view degeneracy in an active focal length control paradigm that balances the probability of view degeneracy with the camera field of view. In order to validate both our view degeneracy model as well as our active focal length control model, a set of experiments are reported using a real recognition system operating on real images.

Index Terms—View Degeneracy, aspect graphs, object recognition.

1 Introduction

1.1 Motivation

Image segmentation is the task of transforming a signal-level image description into a symbolic description, consisting of sets of segments or features of some sort. Examples of features in common use are homogeneous regions or discontinuities described by line segments or curves. From certain viewpoints, such features may be accidentally aligned in the image, resulting in a *degenerate view*. For example, if two object edges are abutted and collinear, one longer edge is seen in their place. Or, an object may be oriented in such a way that, from a particular viewpoint, two nonparallel edges on the model appear as parallel lines in the image.

Much work in 3D object recognition from a single 2D image is based on the assumption that degenerate views rarely occur. As stated by Witkin and Tenenbaum [43], "regular relationships are so unlikely to arise by chance that, when such relationships can be consistently postulated, they almost certainly reflect some underlying causal relationship, and therefore *should* be postulated." It could be argued, therefore, that regular relations found in an image should be mapped to a set of causal relations identified on the model. For example, in Lowe's SCERPO system [27], parallel lines in the image were mapped to manually identified parallel edges in the model. Pairs of model edges that appear parallel only in a degenerate view were not tagged as parallel in the model since such degenerate views were deemed too unlikely.

Another body of work in 3D object recognition from a single 2D image [3], [2], [18], [21], [29], [30] attempts to recover complex volumetric parts, known as geons [4]. In these systems, geons are recovered by extracting the properties of a restricted generalized cylinder, including the geon's axis (straight vs. curved), its cross-section sweep (constant vs. expanding vs. expanding then contracting), its cross-section symmetry (reflective symmetry vs. rotational symmetry vs. asymmetric), and cross-section shape (straight-edged vs. curved-edge). These properties, in turn, are based on the same nonaccidental image relations proposed by Lowe [27]. Thus, a bent cylinder viewed in the plane of the curve could only be interpreted as a straight cylinder, the former being considered an unlikely degenerate view. Furthermore, those views in which one or more of the defining properties was not visible were also considered degenerate and, therefore, unclassifiable. For example, an end-on view of a cylinder, appearing as a single ellipse, encodes no information about the cross-section sweep property: It could, for example, be a rear-pointing cone, a cylinder, or a sphere.

Given arbitrarily high resolution, the various degenerate views described above are not a problem because they would occur only for a vanishingly small fraction of the possible viewpoints [5]. Unfortunately, cameras and feature extraction operators have finite resolution and, as a result, the probability of encountering a view degeneracy is no longer trivial. Kender and Freudenstein [24] were the first to observe this phenomena, but did not provide a computational model. In the first part of this paper, we introduce the first computational model for view degeneracy that can be used to compute the probability that, from a random perspective view of the object, a view degeneracy will occur. For a typical tabletop vision system parameterization, we show that this probability can be *very* significant, even as high as 50 percent.

Given that view degeneracies occur with significant probabilities, what does this mean for the designers of both viewer-centered and object-centered object recognition systems? Any correspondence between an image feature group (containing a degeneracy) and a model feature group will be incorrect. The time spent solving for model pose or

- *S. Dickinson is with the Department of Computer Science and Center for Cognitive Science, Rutgers University, New Brunswick, NJ 08903. E-mail: sven@cs.rutgers.edu*
- *J. Tsotsos and D. Wilkes are with the Department. of Computer Science, University of Toronto, 6 Kings College Rd., Toronto, Ontario, Canada M5S 1A4. E-mail: {tsotsos, wilkes}@vis.tornoto.edu.*

Manuscript received 12 May 1997; revised 16 Feb. 1999.
Recommended for acceptance by K. Bowyer.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 107716.

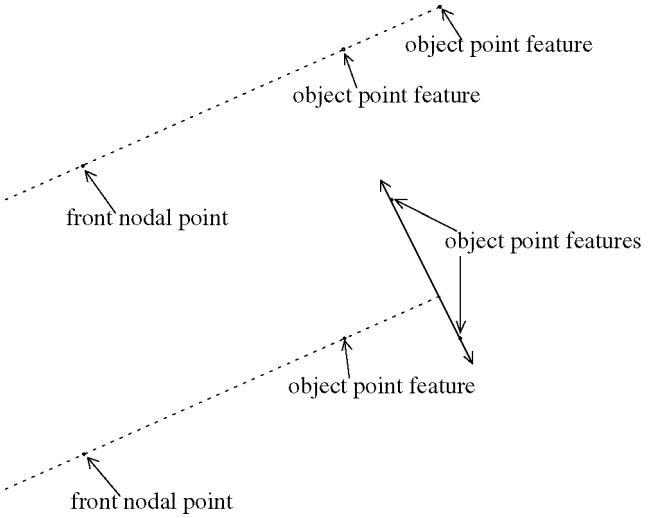


Fig. 1. Our definition of view degeneracy. The two cases are the collinearity of the front nodal point of the lens with two object point features (top), or with one object point feature and a point on one line or line segment (bottom).

verifying the correspondence will only serve to reduce recognition efficiency and performance. In other cases, recognition may not even be possible when degenerate views prevent critical 3D inferences. Conversely, in view-based recognition systems using aspect graphs or collections of views, care must be taken to explicitly include the degenerate views in the representation or else single-view recognition may not be possible.

Given that view degeneracies occur with significant probabilites, what does this mean for the designers of both viewer-centered and object-centered object recognition sytems? Any correspondence between an image feature group (containing a degeneracy) and a model feature group will be incorrect. The time spent solving for model pose or verifying the correspondence will only serve to reduce recognition effciency and performance. In other cases, recognition may not even be possible when degenerate views prevent critical 3D inferences. Conversely, in view-based recognition systems using aspect graphs or collections of views, care must be taken to explicitly include the degenerate views in the representation or else single-view recognition may not be possible.

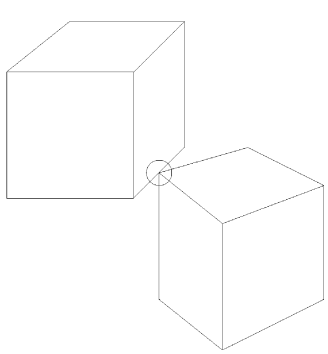


Fig. 2. An example of view degeneracy in which a vertex is imaged in the plane of a scene edge.

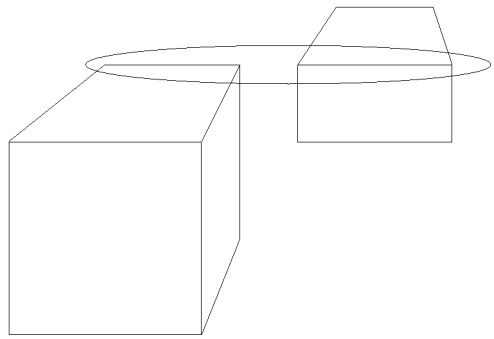


Fig. 3. An example of view degeneracy in which parallel scene lines are imaged in their own plane.

In the second part of this paper, we explore the impact of our model on the design of object recognition systems. For a particular recognition system, we show that there exists a high correlation between degenerate views and ambiguous views of an object, linking view degeneracy with the cost of object disambiguation. One way to reduce this cost is by lowering the probability of encountering a degenerate view. Having already established that the probability of view degeneracy is sensitive to camera focal length, we then introduce a quantitative prescription for active focal length control that allows a trade-off to be achieved between the camera field of view and the probability of view degeneracy. We conduct a series of experiments applying a real object recognition system to a set of real images to validate both our view degeneracy model as well as our active focal length control model.

1.2 Related Work

The notion that finite-resolution cameras and finite-width feature operators could significantly increase the probability of view degeneracy was first articulated by Kender and Freudenstein [24]. However, no computational model was provided for determining exactly how high this probability could be for a given object or how sensitive this probability was to various system parameters, e.g., focal length, operator width, etc. Dickinson et al. [12], [15], [14] analyzed the set of views of simple volumetric parts and found that the empirically derived view degeneracy probabilities were

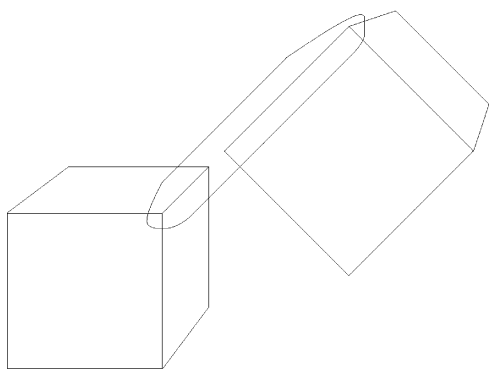


Fig. 4. An example of view degeneracy in which coplanar scene lines are imaged in their own plane.

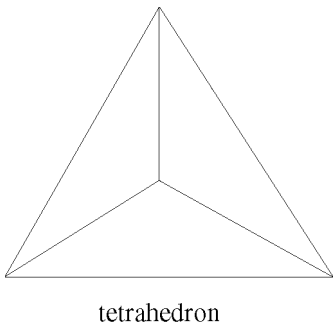


Fig. 5. An example of view degeneracy in which perfect symmetry gives preference to a flat, 2D interpretation of the lines.

significant, leading to the inclusion of "degenerate views" in their aspect database.

In 1993, Wilkes, et al. [40] introduced the first computational model of view degeneracy based on a full perspective camera model. Shortly after that, Shimshoni and Ponce [32] also noted that "accidental views" will appear over a finite region of the viewing sphere when considering finite resolution. Assuming orthographic projection, they presented an algorithm for constructing the finite resolution aspect graph of a polyhedron. Preceding the work of Shimshoni and Ponce, Eggert et al. [17] also noted the effect of finite resolution on the complexity of an aspect graph. Using a full perspective model, they introduced the notion of a scale-space aspect graph for 2D polygonal objects.

A number of researchers have studied the closely-related problem of how an object's features vary with view. For example, Burns et al. studied how projections of 3D point sets and lines vary with viewpoint [8]. Ben-Arie introduced

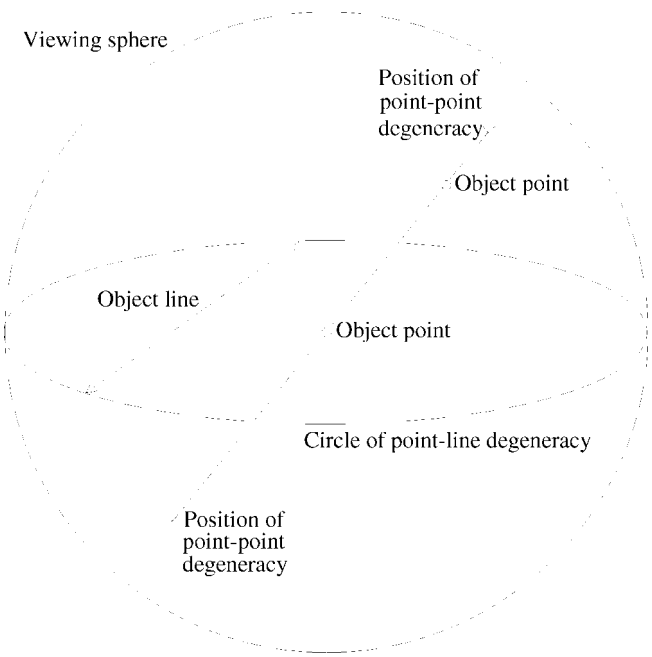


Fig. 6. Loci of degeneracy with infinite resolution. This illustrates degeneracies of both Case 1 type (involving two 0D object features) and Case 2 type (involving a 0D object feature and an infinite line defined by object features).

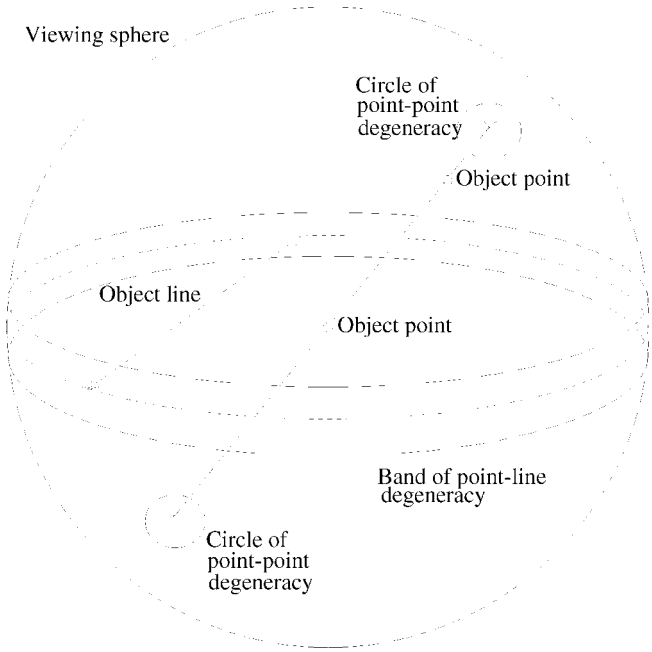


Fig. 7. The loci of degeneracy of the previous figure with finite resolution. The regions of the viewing sphere affected by degeneracy are large. The circle of the previous figure is actually a thick band on the sphere. The points from the Case 1 degeneracy are actually disks.

probabilistic methods for computing the observation probabilities of an object's features and their quantitative attributes [1]. Burns and Kitchen [7] and Swain [35] both proposed view-based, decision-tree-based recognition systems that exploited probability distributions for an object's features, while Weinshall and Werman [38] measured the probability that a certain view is observed (or how stable the view is with changes in viewpoint).

Despite the large body of work in aspect graph construction, e.g., [25], [34], [17], [16], [28], [31], [32], [20], [19], [33], very few aspect graph-based recognition systems have been developed, e.g., [15], [14], [9], [23]. If a degenerate view is ambiguous or excluded from the set of model views in order to reduce complexity, then one must have a mechanism for *detecting* degeneracy when it occurs, and *moving* out of the degenerate view position. This can be accomplished in two ways, each requiring changes in imaging geometry. Wilkes [39] has explored the idea of moving the camera to a viewpoint advantageous to recognition. This approach meets with success, but at the cost of significant viewpoint changes. Dickinson et al. [10], [11] use a probabilistic aspect graph to guide the camera from a position where the view is ambiguous (and often degenerate) to a position where the view is unambiguous.

The second approach to camera motion in order to move out of a degenerate view is to increase the resolution of the camera, i.e., increase the camera's focal length. In 1995, Wilkes et al. [41] applied their computational model of view degeneracy to actively control the focal length of a camera to balance the trade-off between probability of view degeneracy and the field view. A similar trade-off has been noted in the work of Brunnström et al. [6] who use a wide angle of view to

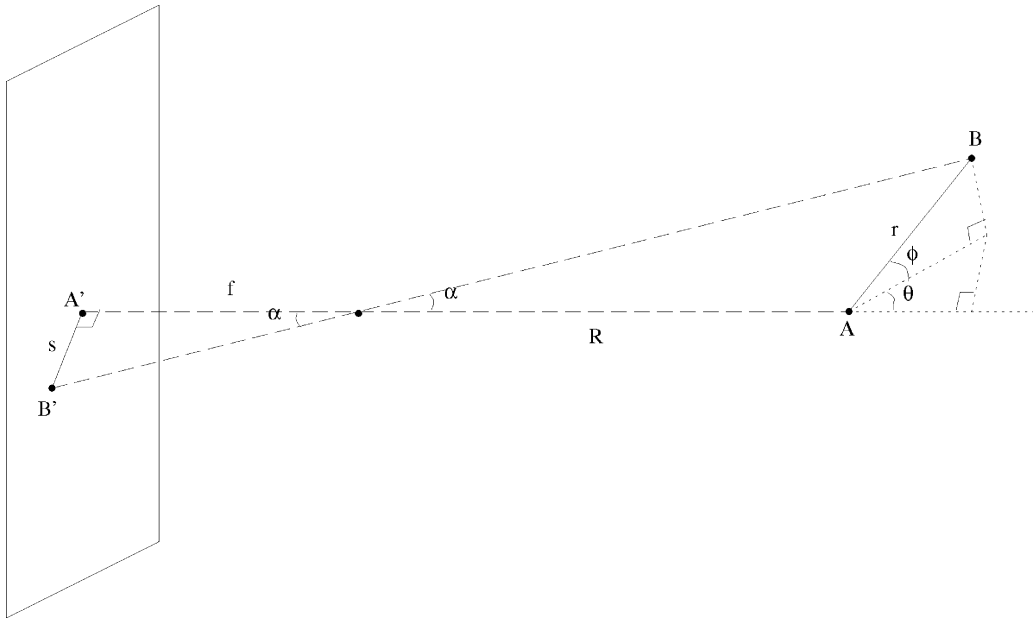


Fig. 8. Perspective projection of a pair of points. We will determine the fraction of all orientations of line AB in relation to the camera for which the image separation s of A and B is small enough to interfere with their resolution into distinct features.

detect object junctions and then zoom in to increase image resolution for junction classification.

1.3 What's Ahead

In the remainder of this paper, we begin by defining view degeneracy precisely (Section 2), and show how our definition covers many interesting classes of degeneracy. Next, in Section 3, we develop a model of view degeneracy under perspective projection. It is parameterized by the distance to the object to be recognized, the separation of features on the object, the minimum separation of the corresponding image features in order for them to be detectable as distinct features, and camera focal length. The model gives the probability of encountering a degenerate view as a function of the model parameters. The evaluation of the model at realistic parameterizations, including a sensitivity analysis, indicates that degenerate views can have very significant probabilities (Section 3.4).

In Section 4, we discuss the implications of our model for object recognition. Choosing a viewer-centered object recognition system, we incorporate our view degeneracy model into a prescription for active focal length control that trades off view degeneracy with field of view. A set of experiments are shown with real images to validate both our view degeneracy model as well as our active focal length control model. Finally, in Section 5, we draw some conclusions about the model.

2 Definitions of Degeneracy

Kender and Freudenstein [24] presented an analysis of degenerate views that pointed out problems with existing definitions of the concept. The definitions that they propose in place of earlier problematic ones incorporate the notion that what constitutes a degenerate view is dependent on what heuristics the system uses to invert the projection from three dimensions to two. One of the definitions incorporating this system-dependence is the negation of the definition of a *general viewpoint*: A general viewpoint is such that there is some positive epsilon for which a camera movement of epsilon in any direction can be taken without effect on the resulting semantic analysis of the image. This definition was noted by Kender and Freudenstein to be incomplete in the sense that there exist general viewpoints on certain objects that are unlikely to allow a system to instantiate a proper model for the object relative to other views of the same object (e.g., the view of the base of a pyramid). They argue for somehow quantifying degeneracy according to the amount of additional work expected to be required to unambiguously instantiate a model corresponding to the object being viewed.

We are interested in a definition that is independent of the particular recognition system in use, and are willing to trade away the completeness of the definition to achieve this. Fig. 1 illustrates a definition of degeneracy that covers a large subset of the cases appearing in the literature. The definition has to do with collinearity of the front nodal point of the lens and a pair of points on (or defined by) the object. We shall consider a view to be degenerate if at least one of the following two conditions holds:

> a zero-dimensional (point-like) object feature is collinear with both the front nodal point of the lens and either:
>
> 1. another zero-dimensional object feature, or
> 2. some point on a line (finite or infinite) defined by two zero-dimensional object features.

Each of the specific imaging degeneracies enumerated by Kender and Freudenstein for polyhedral objects belongs to one of these types. Below, we discuss each of their example degeneracies in turn.

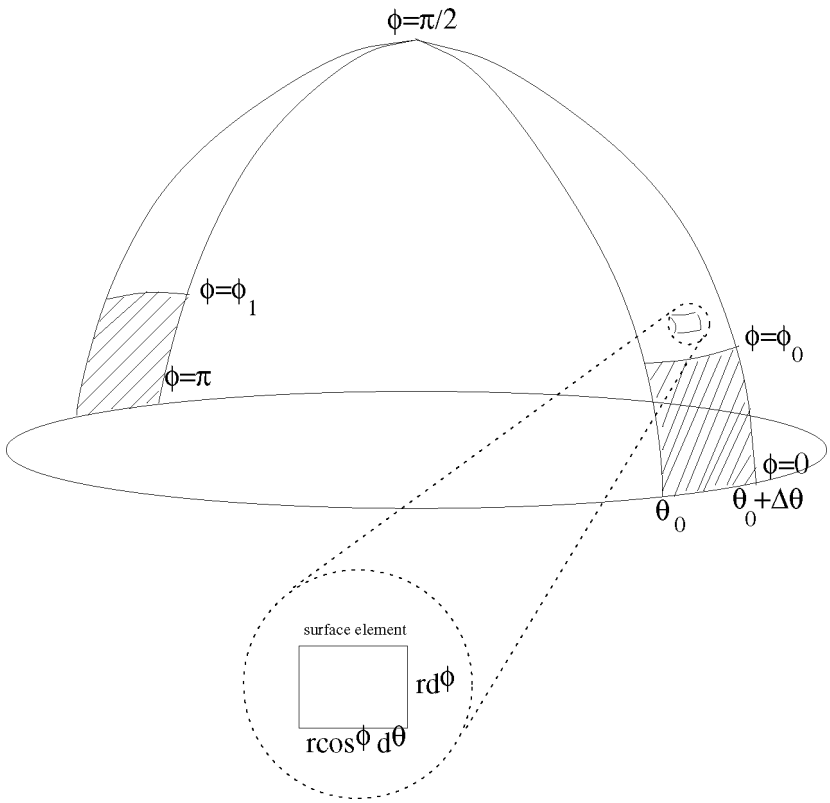


Fig. 9. The regions of the sphere swept by point B for $\theta \in [\theta_0, \theta_0 + \Delta\theta)$ and $\phi \in [0, \pi]$. The shaded parts are the areas for which $s < s_c$.

2.1 Vertices Imaged in the Plane of Scene Edges

This is encompassed by Case 2 (finite line) above and is depicted in Fig. 2. The lower right scene edge on the upper object is the line segment of Case 2 (finite line). The vertex on the lower object is the zero-dimensional object feature.

2.2 Parallel Scene Lines Imaged in Their Own Plane

We depict this situation in Fig. 3. This is Case 2 (infinite line) above. The portion of the viewing sphere in which the degeneracy occurs is that for which the front nodal point of the lens is in the plane defined by an end point of the edge on the front object and the other scene line (extended to have infinite length).

2.3 Coincident Scene Lines Imaged in Their Own Plane

This generalizes to *coplanar lines imaged in their own plane*. We depict this situation in Fig. 4. This is equivalent to Case 2 (infinite line). This reflects the fact that a degeneracy exists (in that noncollinear lines are perceived as collinear) for *any* viewpoint (not just a subset of the viewpoints) in the plane of (but not on) the lines. An endpoint of the edge on the right-hand object may be used as the zero-dimensional feature. The edge on the left-hand object may be extended to define the infinite line.

2.4 Perfect Symmetry

Kender and Freudenstein include perfect symmetry in the enumeration of types of degeneracy because they say that it leads to a tendency to interpret the symmetric structure as flat. An example is shown in Fig. 5. In the case of radial symmetry, this is equivalent to Case 1, with the two zero-dimensional points chosen to be any two distinct points on the axis of symmetry. In the case of symmetry about a plane, this is equivalent to Case 2 (infinite line), with both features chosen arbitrarily to lie in the plane of symmetry. Note that, in the case of symmetry, the features defining the degeneracy are more likely to be abstract points on the object (e.g., the centroid of an object face) rather than points extractable directly by low-level vision. The use of abstract object points rather than obvious feature points has no impact on the estimation of the probability of degeneracy that follows. What matters is that there is some means to enumerate the types of degeneracy that are important to the approach. This enumeration is discussed later in the paper.

2.5 Dependence on Effective Resolution

If a computer vision system had infinite ability to resolve object features that were close to one another, then all of the degeneracies discussed above would have infinitesimal probability of occurrence. Fig. 6 shows typical loci of points on the viewing sphere at which instances of each type of degeneracy occur, assuming infinite resolution. Each Case 1 degeneracy occurs at a pair of points on the sphere surface. Each Case 2 (infinite line) degeneracy occurs at a circle of points on the sphere surface. Each Case 2 (finite line) degeneracy occurs on two sectors of a circle of points on the sphere surface. The sectors are the sets of viewpoints from which the defining point feature appears to lie within the defining line feature. Since these loci of points are of lower dimensionality than the viewing sphere itself, the

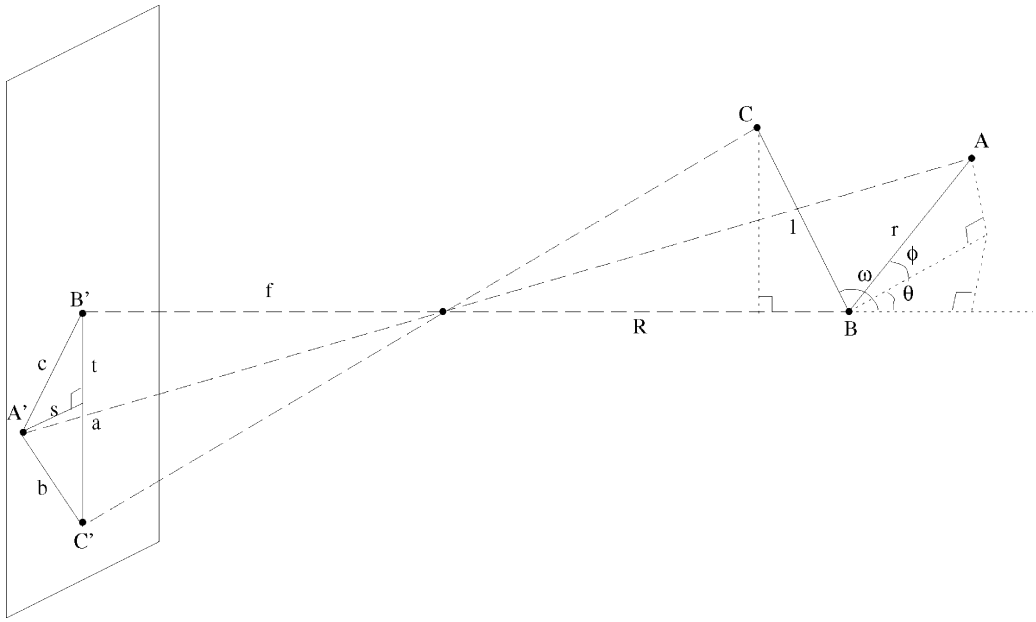


Fig. 10. Perspective projection of a line segment and a point. We will determine the fraction of all possible orientations of A with respect to line BC for which the image separation of A and BC is small enough to interfere with the detection of A as a separate feature.

probability that the viewpoint is on such loci is vanishingly small.

Kender and Freudenstein [24] point out that the finite resolution of real systems has the potential to make the probability of the degenerate views significant or even a certainty. This is due to the thickening of all of the loci of points on the viewing sphere corresponding to degeneracies to include viewpoints from which the key features are "almost collinear" with the front nodal point of the lens. Fig. 7 illustrates this thickening.

We will use the phrase *effective resolution* to capture all system dependencies in our model. As a result, the definition is a general one, to be specialized to each system as needed:

Definition. *The* effective resolution *in the image is the minimum separation of a pair of points, corresponding to distinct object features, needed in order to detect both features.*

Effective resolution is a function of both actual camera resolution and the method used for feature detection. For example, one may estimate the effective resolution to be the width of the convolution kernel used to extract object edges or some function of the kernel width and edge length chosen to model actual system performance.

3 A Computational Model of View Degeneracy

3.1 Degeneracy Based on a Pair of Points:

Fig. 8 shows the perspective projection of a pair of feature points, A and B. For simplicity, our analysis assumes that A is centered in the image.[1] We are interested in the separation s of the images of the two points as a function of the other parameters shown in the figure. The parameters are:

- f distance from image plane to rear nodal point
- R distance of A from the lens front nodal point
- r separation of the two object points
- (θ, ϕ) orientation of B with respect to the optical axis
- α apparent angular separation of the points

We have that

$$\tan\alpha = \frac{s}{f} \tag{1}$$

and that

$$\tan\alpha = \frac{\sqrt{r^2\sin^2\phi + r^2\cos^2\phi\sin^2\theta}}{R + r\cos\phi\cos\theta} \tag{2}$$

Let s_c be the effective resolution, as defined earlier. Equating the two right-hand sides, we will solve for ϕ given $s = s_c$ and θ. The choice to solve for ϕ is arbitrary. We obtain two solutions, ϕ_0 and ϕ_1, at which $s = s_c$:

$$\phi_0 = \cos^{-1}\left(\frac{-Rs_c^2 + f\sqrt{r^2f^2 - s_c^2R^2 + s_c^2r^2}}{r(f^2 + s_c^2)\cos\theta}\right) \tag{3}$$

$$or \quad \phi_1 = \pi - \cos^{-1}\left(\frac{Rs_c^2 + f\sqrt{r^2f^2 - s_c^2R^2 + s_c^2r^2}}{r(f^2 + s_c^2)\cos\theta}\right), \tag{4}$$

ϕ_0 and ϕ_1 exist provided that the arguments for their respective $\cos^{-1}$ functions are in the interval $[-1, 1]$. Where they exist, there are intervals of ϕ for which $s < s_c$. The intervals are $[0, \phi_0)$ and $(\phi_1, \pi]$. Where they do not exist, it is because the entire range of ϕ values from 0 to π gives $s \geq s_c$ at the given θ value. Fig. 9 shows the sections of the sphere swept out by point B in Fig. 8 as θ ranges over a small

1. This simplification is based on our assumption that the off-axis effects are small, assuming that the camera is foveating on the object.

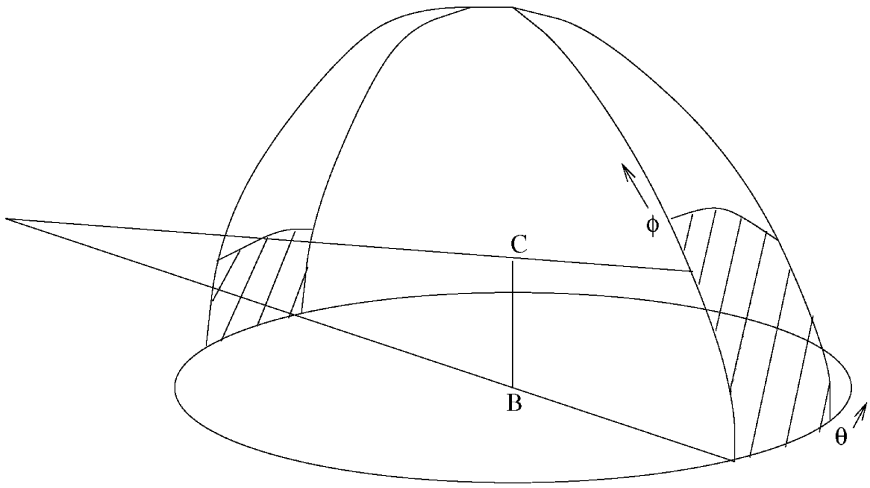


Fig. 11. In the point-line case, the point A sweeps a sphere as θ and ϕ are varied for each value of ω. Here, $\omega = \frac{\pi}{2}$. The shaded regions are the areas for which $s < s_c$. For ease of presentation, the lower half of the sphere is not shown and only half of each region in the upper half of the sphere is shown. The lines passing through B and C intersect at the front nodal point of the lens.

interval $[\theta_0, \theta_0 + \Delta\theta)$ and ϕ ranges over the interval $[0, \pi]$. For each small interval in θ, the ratio of the area for which $s < s_c$ to the total area of the sphere swept by point B is $g(\theta)$, where

$$g(\theta) = \frac{1}{total\ area} \times (shaded\ area) \tag{5}$$

$$= \frac{1}{2r^2\Delta\theta} \times \left(\int_0^{\phi_0} \int_{\theta_0}^{\theta_0+\Delta\theta} r^2 \cos\phi d\theta d\phi + \int_{\phi_1}^{\pi} \int_{\theta_0}^{\theta_0+\Delta\theta} r^2 \cos\Big(\pi - \phi\Big) d\theta d\phi \right) \tag{6}$$

$$= \frac{1}{2r^2\Delta\theta} \times r^2 (\int_0^{\phi_0} \Delta\theta \cos\phi d\phi + \int_{\phi_1}^{\pi} \Delta\theta \cos\Big(\pi - \phi\Big) d\phi \Big) \tag{7}$$

$$= \frac{1}{2\Delta\theta} \times \Delta\theta(\sin\phi_0 + \sin(\pi - \phi_1)) \tag{8}$$

$$= \frac{\sin\phi_0 + \sin\phi_1}{2}. \tag{9}$$

Thus, we may estimate the probability that s is smaller than some critical value s_c by doing a numerical integration of $g(\theta)$ over all θ values to determine the fraction of the sphere swept by θ and ϕ that gives $s < s_c$.

3.2 Degeneracy Based on a Point and a Line

Fig. 10 shows the perspective projection of a line segment of length l, with endpoints B and C and a point A. We simplify our calculations by centering B in the image. Without loss of generality, we choose an image coordinate system that causes the image of the line to lie along one of the coordinate axes. Let A', B', and C' be the images of A, B, and C, respectively. In addition to the parameters used in the case of two points, we have the additional parameters:

- w angle between the line segment and the optical axis
- a length of the line segment $B'C'$ (possibly infinite)
- b separation of A' and C'
- c separation of A' and B'
- t distance between B' and the nearest point to A', on the infinite line through B' and C'

The parameter t is used to determine whether a perpendicular dropped to the line through B' and C' falls on the segment between B' and C'.

We have

$$s = r\cos\phi\sin\theta\frac{f}{R + r\cos\phi\cos\theta} \tag{10}$$

$$t = r\sin\phi\frac{f}{R + r\cos\phi\cos\theta} \tag{11}$$

$$a = l\sin\omega\frac{f}{R + l\cos\omega} \tag{12}$$

$$b = \sqrt{s^2 + (t-a)^2} \tag{13}$$

$$c = \sqrt{s^2 + t^2}, \tag{14}$$

We are interested in the distance of the image of the point from the line segment. This distance d is

$$d = \begin{cases} s \text{ if } t \geq 0 \text{ and } t \leq a \\ b \text{ if } t > a \\ c \text{ if } t < 0. \end{cases} \tag{15}$$

We considered two possibilities for the integration for the probability that d is less than some critical value d_c. The

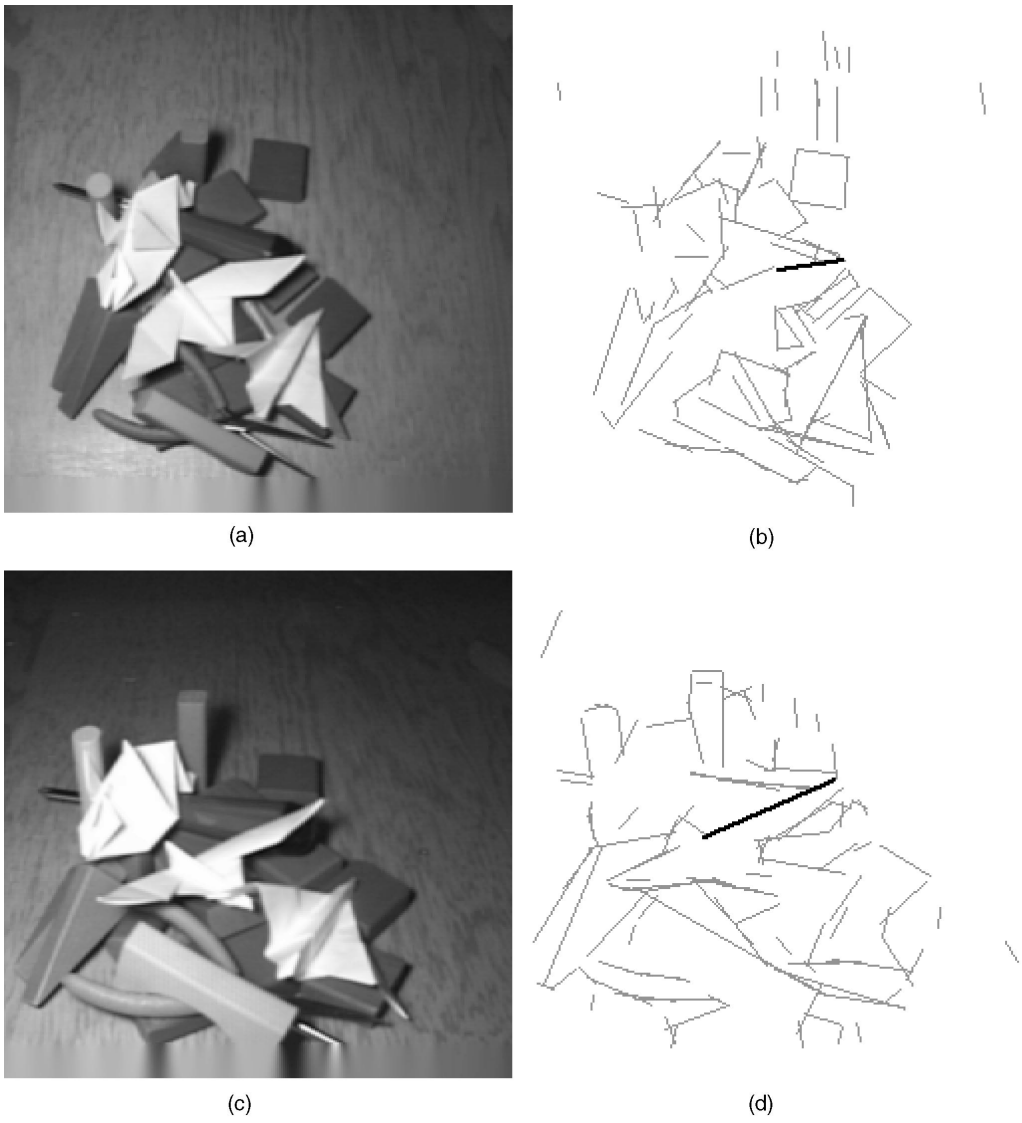


Fig. 12. Segmented Images of Origami Object. (a) Nondegenerate gray-level image of origami figures. (b) Edge image with the edge of interest highlighted. (c) Gray-level image of same origami figures from a slightly different viewpoint. (d) Edge image of (c) showing the absorption of a second edge due to the apparent collinearity (point-line degeneracy) in the new view.

approach that is analogous to the approach used in the first case leads to numerical evaluation of a three-dimensional integral. Fig. 11 shows the sections of the sphere swept out by the point A, assuming a vertical line BC (with B at center of sphere). Note that, for clarity, the portion swept out on the lower half of the sphere is not shown and only half of each section (on either side) on the top half is shown. We chose an easier method, consisting of performing a Monte Carlo integration by randomly choosing triples (ω, θ, ϕ) uniformly on the surface of the viewing sphere and tabulating the fraction of the trials for which $d < d_c$.

3.3 Combining Probabilities of Individual Degeneracies

Finally, if we have estimates of the probabilities of degenerate views based on individual pairs of points and (point, line segment) pairs, we need to combine the probabilities of individual degeneracies to get the probability that at least one degeneracy exists. This final step is somewhat application-dependent because different classes of objects have different degrees of interaction among their features. If we were to assume independence among all of the points and line segments in the object set and if the objects were wire-frame objects so that all features were visible all of the time, then the task would be easy: All features would be able to interact with each other to cause degeneracies. This model is appropriate with sensors or objects for which transparency or translucency is common (e.g., X-ray images).

For computer vision, there typically exist pairs of features that do not interact to form a degeneracy simply because they are at opposite positions on an opaque object. We will examine the case of a polyhedral world in order to provide an example of the analysis necessary to determine overall probabilities of occurrence of degeneracy. In a polyhedral world, features are clustered to be in coplanar groups (faces), reducing the probability of degeneracy from

Perturbation in Functional Form	
f	dp_d/df
0.004000	-194.166926
0.005000	-115.379056
0.006000	-70.300510
0.007000	-47.414434
0.008000	-35.003845
0.009000	-24.745112
0.010000	-18.982178
0.011000	-15.836535
0.012000	-12.349427
0.013000	-10.273039
0.014000	-7.810763
0.015000	-7.117619
0.016000	-6.041048
0.017000	-5.176331
0.018000	-3.847985

Fig. 13. Functional perturbation analysis: p_d/df as a function of f.

the value that would be estimated assuming independence of the features.

In the case of single isolated convex polyhedral objects, we can take a face-by-face approach to estimating the probability of degeneracy. Specifically, we will only examine interactions among points and lines belonging to the same face since features not sharing a face are relatively less likely to interact. An edge-on view of a face is equivalent to our Case 2 (infinite line) degeneracy where the point used is an arbitrary vertex and the line used is an edge on the opposite side of the face extended to have infinite length.

To arrive at a global estimate of degeneracy, we could simply sum the probabilities of degeneracy for each face. However, since parallel faces cause strongly overlapping regions of the viewing sphere to be degenerate, faces should be divided into equivalence classes, with one class per face orientation. For example, a cube consists of three sets of faces, with three distinct orientations. We may estimate the probability of degeneracy for the cube by summing the probabilities of three faces, one representing each equivalence class. Finally, we can take into account the degeneracy due to radial symmetry by summing the point-pair degeneracy estimates for each axis of radial symmetry in the object.

If we define terms as follows:

- p_{00} the probability of an individual point-pair degeneracy
- p_{01} the probability of an individual line-point degeneracy
- n_r the number of axes of radial symmetry
- n_f the number of face equivalence classes

where n_r does not include axes of symmetry lying in planes of symmetry, then the overall probability of a degenerate view, p_d, is estimated by

Basic parameterization	
parameter	value
f	0.007 meters
R	0.50 meters
r	0.04 meters
l	0.04 meters
s_c	0.0001548 meters (3 pixels)
n_f	3
n_r	0

Results at basic parameterization	
probability	value
p_{00}	0.039181
p_{01}	0.073 ± 0.003
p_d	0.20

Perturbations of basic parameterization			
perturbation	p_{00}	p_{01}	p_d
$f = 0.0035$ m	0.167181	0.217	0.52
$f = 0.014$ m	0.009657	0.028	0.08
$R = 0.25$ m	0.009835	0.028	0.08
$R = 1.00$ m	0.166838	0.220	0.52
$r = 0.02$ m	0.166838	0.198	0.48
$r = 0.08$ m	0.009835	0.015	0.04
$l = 0.00$ m	0.039181	0.039	0.11
$l = 0.02$ m	0.039181	0.045	0.13
$l = 0.08$ m	0.039181	0.075	0.21
$l = 10^5$ m	0.039181	0.078	0.22
$s_c = 1$ pixel	0.004281	0.016	0.05
$s_c = 5$ pixels	0.112892	0.161	0.41
$n_f = 1$	0.039181	0.074	0.07
$n_f = 5$	0.039181	0.074	0.32
$n_r = 3$	0.039181	0.073	0.29
human fovea	7×10^{-7}	2×10^{-4}	5×10^{-4}

Fig. 14. Results for a tabletop vision system. The basic parameterization corresponds to our experimental setup. One fifth of the views on a single object are degerate. The probability of degeneracy varies significantly as model parameters are perturbed.

$$p_d = 1.0 - \underbrace{(1.0 - p_{00})^{n_r}}_{no\ symmetry\ degeneracies} \times \overbrace{(1.0 - p_{01})^{n_f}}^{no\ face\ degeneracies} . \qquad (16)$$

The above expression corresponds to an assumption of independence of the axes of radial symmetry and face groups. We conjecture that this is more realistic than assuming no overlap between the regions of degeneracy (as one would assume by simply summing the individual probabilities of degeneracy). In a world with multiple objects in each image, there is a greater degree of independence among the features since the positions of the objects with respect to one another are less likely to be as ordered as the positions of individual features within an object. The above enumeration would provide an optimistic lower bound on the probability of degeneracy. Summation over all feature pairs (as in the wireframe case) would

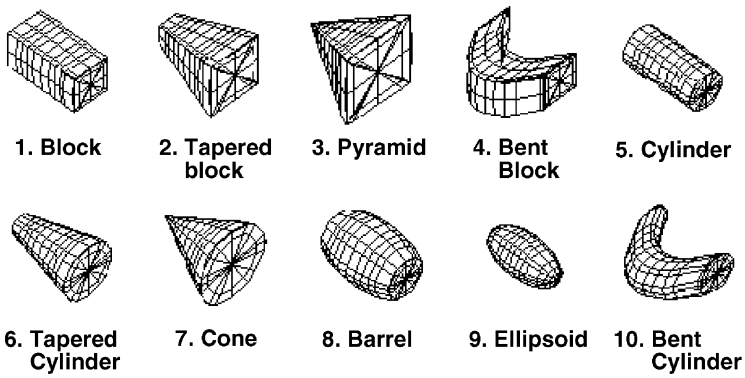


Fig. 15. The 10 modeling primitives.

provide a pessimistic upper bound on the probability of degeneracy.

Our analysis has ignored the fact that objects in the world have preferred orientations, sit on tables, have tops and bottoms, and so forth. In fact, there may be surfaces on an object which are rarely visible. In a more thorough analysis, the probability of view degeneracy contributed by a face should be weighted by the probability of the visibility of the face. Finally, it should be noted that in the case of nonconvex objects, (16) provides an even more optimistic lower bound on the probability of degeneracy. For convex polyhedra, the point-line degeneracies involve points and lines on adjacent faces. However, in the case of nonconvex polyhedra, nonadjacent faces (in addition to adjacent faces) may contribute to point-line degeneracies involving a given face.

3.4 Sensitivity Analysis of the Model

We have parameterized the model to match the situation in some of our own work on active object recognition. The experimental apparatus consists of a gray-level CCD camera mounted on a light robot arm which is, in turn, mounted on a mobile robot. Recognition of origami figures lying on a tabletop is accomplished by using low-level image data to drive the camera to a special viewpoint on the unknown object. The recognition work is described in detail in [39] and presented in overview in [42]. We have used the actual camera pixel size to compute the critical feature separation s_c, while the object-related parameters are those of the origami object set. Example segmented images with and without a view degeneracy are shown in Fig. 12.

Fig. 14 tabulates the results with the basic parameterization, and examines the sensitivity of the probability of encountering a degeneracy to each of the model parameters by perturbing each parameter in isolation. The parameters were derived as follows:

- f was the nominal focal length of the camera lens.
- l was the average (to one significant digit) of the edge lengths over all discriminable edges in the eight origami figures used in the experiment.
- r was set equal to l on the grounds that the average interfeature distance would be of the same order of magnitude as the edge length in typical scenes.
- s_c was three times the actual CCD pixel size.
- n_f was the nearest integer to the average value over the eight origami figures.
- R was the approximate operating distance (to one significant digit) of the camera from the objects, given that the example recognition system actively controlled the camera to give an edge of interest in the scene a fixed initial length in the image (so average camera distance is dependent on the length of scene edges).

The above perturbations were given as sample data points, due to the discrete nature of the parameters. And, although the models in Fig. 8 and Fig. 10 can be scaled up or down at will (so all the parameters could be divided by either R or f), it was a matter of convenience to parameterize the model directly from measured quantities. Nevertheless, in Fig. 13, we show one of the perturbations in functional form.

We see that, for example, for our model system with a parameterization to recognize small tabletop objects from a range of half a meter, there are one in five odds that the view encountered will be degenerate. These high odds were reflected empirically by the actual recognition system [39]. The sensitivity analysis demonstrates that there are realistic parameterizations of the model for which probabilities of degeneracy are very significant. Most of the variations of p_d with the individual parameters are quite intuitive. For example, as distance to the object increases or focal length decreases, the probability of degeneracy increases, exceeding 50 percent in some cases. Notice that the probability of degeneracy also increases somewhat as object line length l increases. This is due to the increased probability of interference between the point and line features, as described in Case 2 of our definition. In the other direction, the probability of a point-line degeneracy becomes equal to the probability of a point-point degeneracy as line length l becomes zero. We also tried a parameterization with f and s_c set to match human foveal acuity of 20 seconds of arc. The probabilities of degeneracy were negligibly small. This may explain why the importance of degenerate views to computer vision has traditionally been underestimated.

Although our model has been defined for polyhedral objects, empirical studies using curved objects have come to the same conclusion. In Dickinson et al. [15], we examined the aspect graphs of 10 classes of volumetric parts, including a number of curved objects (cylinder, bent cylinder, cone,

truncated cone, ellipsoid, truncated ellipsoid, and bent block). Empirically estimating the probabilities of the views of the various parts over a tessellated viewing sphere, we found that the degenerate views of curved objects also had nontrivial probabilities. In a recognition system which used the probabilities to guide an aspect recovery process, we often encountered such views.

4 Implications for Recognition

Our model and realistic parameterizations of it provide quantitative arguments that so-called degenerate views should be taken into account by designers of view-based object recognition algorithms. This was the approach taken by the work of Dickinson et al. [13], [15], [14] in which an object's volumetric parts are recovered from a single 2D image using part-based aspect matching. Integral to this approach is the estimation of the probabilities of different aspects, including the degenerate ones. The probabilities were estimated by determining what set of regions was visible from each point on a discretized viewing sphere. The degenerate views (or aspects) were found to have nontrivial probabilities or nontrivial regions on the viewing sphere. As mentioned in Section 1.2, Shimshoni and Ponce [32] have accounted for these enlarged regions in their finite resolution aspect graph, although using only an orthographic projection model. Eggert et al. [17] have dealt with the issue of finite resolution in their construction of a scale-space aspect graph. The inclusion of degenerate views, when numerous, can lead to increased search complexity, although, as shown in [13], [15], [14], [11], canonical views can be hypothesized before degenarate views.

Degenerate views in 3D from 2D recognition systems can also lead to increased search complexity. Recall from Section 1 that any correspondence between an image feature group (containing a degeneracy) and a model feature group will be incorrect; solving for model pose or verifying the correspondence under these conditions will be fruitless. However, 3D object recognition systems based on simple indexing features such as points or lines can still function with various degeneracies if they can recover a sufficient number of nondegenerate features from the image (e.g., [27], [22]). With a few features, accurate pose estimation is possible and the projected aligned model can be used to verify other model features in the image. Voting techniques can also accumulate evidence for a degenerate view or model orientation with a small number of nondegenerate features [36], [26].

If our goal is to avoid the degenerate views in a viewer-centered object representation or avoid making inferences from accidental viewpoints in a 3D from 2D system, then one must have a mechanism for *detecting* degeneracy when it occurs and *moving* out of the degenerate view position. This can be accomplished in two ways, each requiring changes in the camera's extrinsic or intrinsic parameters.[2] As mentioned in Section 1.2, some approaches attempt to move the camera to an advantageous viewpoint, e.g., [39], [10], [11]. A second, and less costly, alternative is to reduce the probability of view degeneracy. We see from the sensitivity analysis in Section 3.4 that the probability of degeneracy p_d is sensitive to both focal length and effective resolution. Thus, we might consider controlling the fixation of a variable-resolution (foveated) sensor, using an attentional mechanism of some sort, e.g., [37]. Alternatively, we could control the focal length of the camera.

Basic parameterization	
parameter	value
f	0.02 meters
R	1.55 meters
r	0.04 meters
l	0.04 meters
s_c	0.0001548 meters (3 pixels)
n_f	3
n_r	0

Fig. 16. Parameters used in model validation.

There is a clear trade-off between wide angle of view and probability of view degeneracy. In the presence of uncertainty in camera or object position, a wide viewing angle is desirable for initial object detection since the object of interest is more likely to fall in the field of view if focal length is small. On the other hand, small focal length increases the probability of view degeneracy. As noted in Section 1.2, a similar trade-off has been noted in the work of Brunnström [6], who uses a wide angle of view to detect object junctions and then zooms in to increase image resolution for junction classification. The existence of such a trade-off suggests that there can be a principled choice of focal length in many circumstances. In the following subsections, we illustrate the use of our view degeneracy model in selecting the optimal focal length of a camera to maximize object recognition performance for an example object recognition system. Note that, although the example system is a probabilistic view-based recognition system, the reader should not conclude that our analysis of view degeneracy favors one recognition approach over another.

4.1 An Example Object Recognition Problem

Our demonstration of the above trade-off lies in the domain of 3D object recognition from a single 2D image. One of the authors has developed an approach to object recognition based on a hybrid object representation combining object-centered and viewer-centered models [12], [13], [15], [14]. 3D objects are composed of object-centered volumetric parts drawn from some arbitrary finite vocabulary of part classes. The volumetric part classes, in turn, are mapped to a finite set of views, or aspects, that capture the appearance of each part class from all possible viewpoints. To capture the likelihood of a particular part aspect as well as the ambiguity of inferring a part from a given aspect, a set of upward and downward conditional probabilities relates the parts and the aspects.

2. In fact, as proven in [41], detection of view degeneracy cannot be accomplished from a single view.

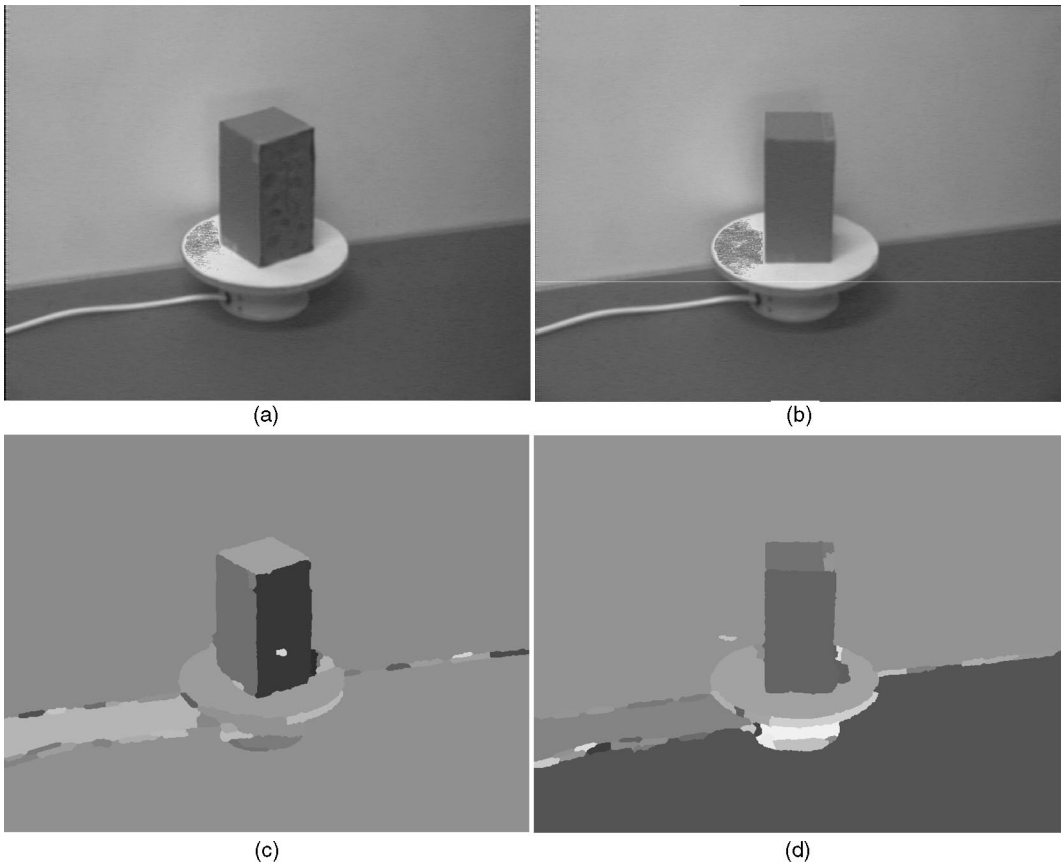


Fig. 17. Example images (original and processed) in the first sequence. (a) Nondegenerate view. (b) Degenerate view. (c) Region segmentation for image in (a). region rsegmentation for image in (b).

Part of the motivation for the above representation is to support view-based recognition while avoiding its intractable complexity; recall that, for a general polyhedra with n faces, the complexity of its aspect graph is $O(n^9)$ [28]. Applying the "recognition by parts" paradigm to an aspect graph of a complete object, we can break down the aspect graph into a set of manageable parts to support local feature indexing, effectively reducing the prohibitive complexity of the overall aspect graph.

Our approach is based on two assumptions. First, we assume that the aspect graph's "parts" correspond to the 3D object-centered volumetric parts that make up the object. And second, we assume the views of an object's parts to be independent. By effectively discarding the covisibility constraints between an object's parts, we dramatically reduce the complexity of the representation. For example, the 10 volumetric part classes shown in Fig. 15 are entirely described by a total of 40 views.[3] Unlike "whole object" aspect graphs, whose size (in number of aspects) is proportional to the size of the object database, the set of 40 views is independent of the number of objects in the database.

3. The 40 aspects were defined by taking the union of all aspects for each volumetric part class and removing redundant views (symmetries). For example, the rectangular block maps to three aspects: three faces, two faces, and one face.

Given the above framework, part recovery can be formulated as local part-based aspect matching. Once a part aspect is recovered, its 3D volumetric class can be inferred. A set of adjacent part-based aspects therefore yields an object-centered indexing structure (consisting of connected 3D volumetric parts) that is rich enough to yield a small number of object candidates for verification [15], [14]. Note that, by assuming independence of the parts' views, we effectively weaken our 3D model by encoding only part connectivity; relative part orientation is therefore lost.

4.2 The Effect of View Degeneracy

Given a recovered part aspect, we employ the aspect hierarchy to infer the part's identity. Using the conditional probability tables mapping aspects to volumes, we generate a part hypothesis for each nonzero mapping. For example, if a block is viewed degenerately as a single rectangular face, this aspect could, in fact, be used to infer six of the 10 volumetric shapes shown in Fig. 15, albeit with different probabilities. Nevertheless, when such a view is encountered, there is no way to disambiguate the hypothesis without moving the camera [11].

A degenerate view is not a problem for a view-based recognition system if its appearance is unique. However, if a degenerate view is ambiguous, then a change in viewpoint is necessary. Consider our set of 10 volumetric part classes and their 40 aspects. From our conditional

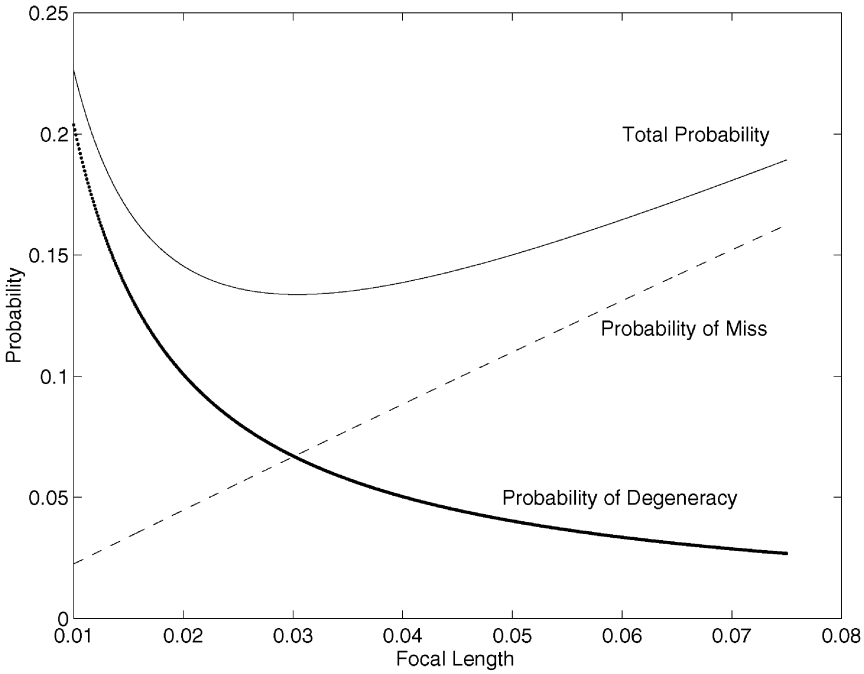


Fig. 18. Predicted p_d, p_{miss} and p_{total} for object recognition experiment as a function of focal length.

probability table mapping volumes to aspects, we know $P(aspect_i|volume_j)$. Therefore, if we assume that all volumetric parts are equally likely in the image, we can compute the probability of occurrence of any of the 40 aspects as follows:

$$P(aspect_i) = \sum_{j=1}^{10} P(volume_j)P(aspect_i|volume_j), \qquad (17)$$

where $P(volume_j) = 1/10$.

If we examine our set of 40 aspects, exactly 13 are degenerate views of at least one volumetric part class in that they contain either a point-point or point-line degeneracy. Furthermore, exactly eight of the 40 views are ambiguous, requiring camera motion to disambiguate them. However, in calculating the probability that a degenerate view is ambiguous, i.e.,

$$P(AmbiguousView|DegenerateView) = \frac{\sum_{view_i \text{ is both degenerate and ambiguous}} P(view_i)}{\sum_{view_j \text{ is degenerate}} P(view_j)} = 0.900, \quad (18)$$

we find that there is a strong correlation between degenerate views and ambiguous views; in fact, 90 percent of the degenerate views are ambiguous! Thus, if we can reduce the probability of a degenerate view, we can, in turn, reduce the probability that a costly camera movement needs to be made to disambiguate the object (volumetric part).

4.3 Reducing the Probability of View Degeneracy

One solution to the problem of reducing the probability of view degeneracy (or probability of having to move the camera to disambiguate the object) is to increase the focal length of the camera, as reflected in our model. However, there is an important trade-off that must be noted here. Although the probability of view degeneracy is *decreased* as the camera is zoomed, the probability that the object is no longer contained in the field of view is *increased*. If the object is only partially visible or leaves the field of view entirely, then a costly camera movement must be made to bring it back into the field of view. The important question now emerges: What is the optimal setting of the camera's focal length that will minimize the probability that the camera must be moved in order to recognize the object (volumetric part)?

We will begin by assuming that, in order to recognize the object within the field of view, the entire object must be visible. Furthermore, we will assume that the location of the projected center of the object in the image can lie anywhere in the image with equal probability. Under this assumption, we can compute the probability that some portion of the object will fall outside the field of view as follows:

$$P_{miss} = 1.0 - \frac{(XSize - 2 * ImgRadius_{avg}) * (YSize - 2 * ImgRadius_{avg})}{XSize * YSize} \qquad (19)$$

where $XSize$ and $YSize$ are the X and Y dimensions of the image, respectively, and $ImgRadius_{avg}$ is the projected length of the object's average radius, $Radius_{avg}$, as defined below:

$$ImgRadius_{avg} = \frac{FocalLength * Radius_{avg}}{DistanceToCamera}. \qquad (20)$$

To minimize the probability (cost) of camera movement, we must therefore minimize the total probability of encountering a degenerate view and losing part of the object from the field of view:

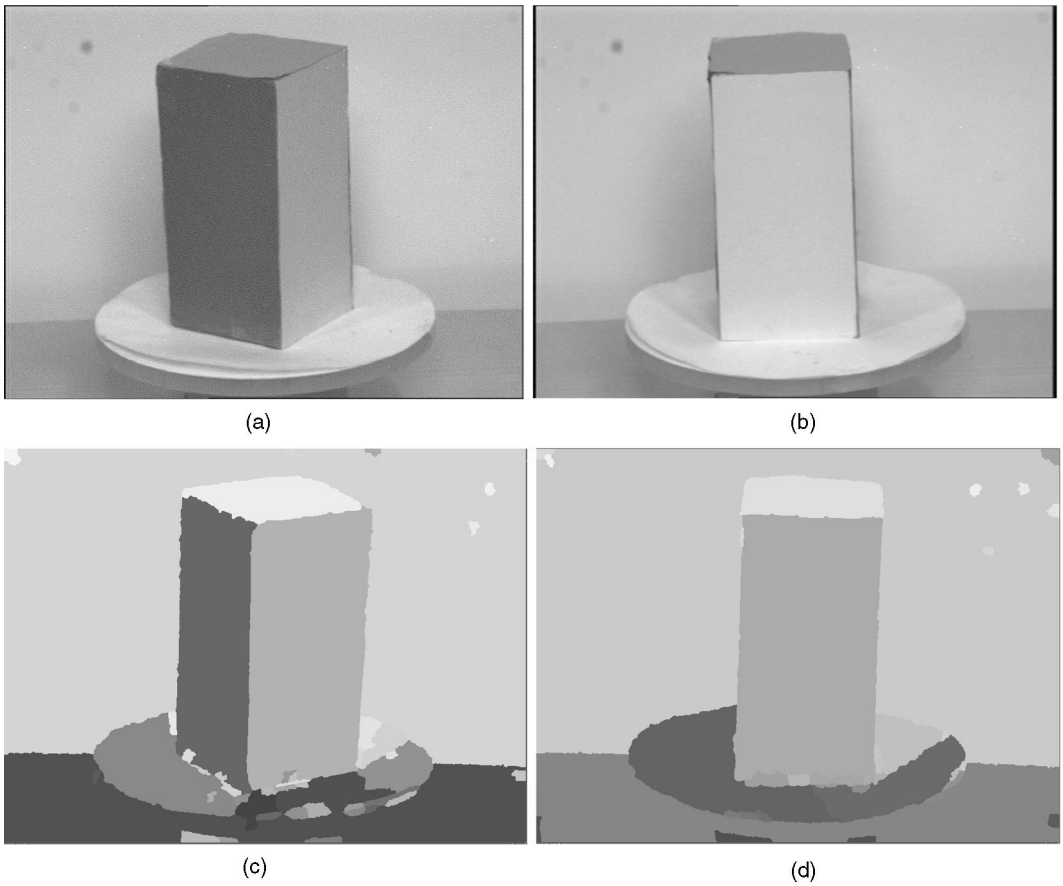


Fig. 19. Example images (original and processed) in the second sequence with the increased focal length. (a), (b) contain a nondegenerate image and a degenerate image from the sequence, while (c), (d) contain their corresponding region segmented images.

$$P(degeneracy \cup miss) = P(degeneracy) + P(miss) - P(degeneracy \cap miss), \quad (21)$$

4.4 Limitations of the Model

It is important to stress that the strong correlation between degenerate views and ambiguous views is a function of the particular parts chosen to make up the part vocabulary. For our database, the correlation is high while, for other databases, the correlation might be very low. For example, if the part vocabulary contains a single object, then there is no ambiguity despite the fact that the probability of encountering a degenerate view may be high. Or, conversely, the part database may contain, for example, a set of different-sized spheres, none of which has an associated degenerate view and all of which are ambiguous from any viewpoint. It is only with our particular system, with its set of shapes that share high-entropy views, that we can effectively apply this technique. If an ambiguous view is not degenerate, then changing the focal length will not disambiguate the object (part), ultimately requiring a change in viewpoint.

It is also important to note that our formulation of the cost of the camera movement is overly simple, and is presented solely to illustrate the trade-off between smaller field of view and higher probability of degeneracy. A more complete formulation would factor in the cost (in time) of changing the focal length and changing the pan/tilt of the camera. Furthermore, the equal probability assumption on the object's position in the image is also overly simple. In this case, a more effective model would assume, for example, that the object sits on a table and that its possible positions are equiprobable. In this case, the projected positions in the image would yield a more complex P_{miss}.[4]

4.5 Experimental Results

We validate our model in the context of the OPTICA system [15], [14], [11]. We begin by placing an isolated volumetric part (volume 1 in Fig. 15) on a turntable in front of a camera. We assume that we know the average inter-feature distance of the edges and vertices of the planar faces comprising the volume, the camera focal length, and the approximate distance to the object. Furthermore, we know the size of the feature extraction operator used by OPTICA to extract regions from the image. The values of these parameters are presented in Fig. 16. The turntable is then rotated 360 degrees and a dense sequence of images is acquired (one image approximately every one degree). Fig. 17 shows a few frames from the image sequence containing both nondegenerate and degenerate views. Since the elevation of the camera is fixed, the top face of the block volume is always visible and, thus, cannot be viewed degenerately. Fig. 17 also shows the results

4. The authors would like to thank one of the anonymous reviewers for suggesting these enhancements to the model.

of OPTICA's region segmentation operator on the two images.

When a view degeneracy occurs, it will appear in the image as two visible faces on the volume (including the top face). This view is, in fact, an ambiguous view, given our part vocabulary, and would require camera movement in order to disambiguate (recognize) the object. Conversely, when no view degeneracy occurs, the three faces of the volume (including the top face) will be visible. In this case, the view is unambiguous and OPTICA would be able to make a unique inference from the aspect without additional camera motion.

Given the two possible views that we will encounter, we need only compute P_{01}, the probability of a point-line degeneracy that corresponds to the degenerate view of the block volume. From the formulation of point-line degeneracy given in Section 3.2, we can derive p_{01} by noting that the angle ϕ is fixed due to the fixed elevation of our camera. Making the appropriate modifications to the model in Section 3.2 and using the parameters in Fig. 16, the resulting plot showing the probability of view degeneracy as a function of focal length is shown in Fig. 18. In addition to showing $P_{degeneracy}$, Fig. 18 also shows P_{miss} and $P_{total} = P_{degeneracy} + P_{miss}$, under the simplifying assumption that $P(degeneracy \cap miss)$ is negligible.[5]

To validate our model, we begin by applying the OPTICA object recognition system to every image in the dense test sequence and noting the number of views which were identified as ambiguous, i.e., images for which OPTICA was able to recover only two distinct faces. Given this empirically-derived value of $P_{degeneracy}$, we can compare it to the predicted value of $P_{degeneracy}$ at $f = 0.02$ (the focal length used in acquiring the images). The predicted value from Fig. 18 (0.1008) is, in fact, quite close to the empirically-derived degeneracy of 0.1111.

Given the above parameterization, we can now compute the optimal focal length that will minimize the probability of having to move the camera in order to disambiguate the volume (recognize the object). We should choose that value of f in Fig. 18 that minimizes the function p_{total}. Choosing the value $f = 0.03$, p_d is calculated to be 0.0670. Now, returning to our experiment, we can zoom the camera to $f = 0.03$, acquire another dense sequence of images, and reapply our recognition algorithm. Fig. 19 shows two example images from this second sequence. Again, our predicted probability (0.0670) compares favorably to our empirically observed probability (0.0666). As expected, due to the increased focal length (resolution), the probability of view degeneracy has gone down, while the probability of losing the object in the field of view has risen.

5 Conclusions

We have introduced the first computational model of view degeneracy for the class of polyhedral objects under perspective projection. Our experimental results indicate that degenerate views are neither accidental nor unlikely. View degeneracy is a surprisingly frequent occurrence that warrants consideration in the design of object recognition systems. In viewer-centered recognition systems, degenerate views must be encoded, while, in object-centered systems, view degeneracy can increase the number of pose hypotheses that need to be verified. In both these paradigms, recognition performance can be increased by reducing the probability of view degeneracy. We have shown how the trade-off between the probability of view degeneracy and the field of view can lead to a focal length prescription that can be used to reduce the probability of having to perform a costly camera viewpoint change in order to recognize an object.

5. In the feasible region of our optimization problem, both events have low probability.

Acknowledgments

The authors would like to first thank Jorgen Bjornstrup and Henrik Christensen at Aalborg University, in Denmark, and Ali Shokoufandeh at Rutgers University for assisting in the image acquisition and analysis for the active focal length control experiment. The authors wish to thank the Natural Sciences and Engineering Research Council of Canada and the Information Technology Research Center for financial support. The authors are members of the Institute for Robotics and Intelligent Systems (IRIS) and wish to acknowledge the support of the Networks of Centers of Excellence Program of the Government of Canada and the participation of PRECARN Associates Inc. John K. Tsotsos is the CP (Unitel) Fellow of the Canadian Institute of Advanced Research. Sven J. Dickinson would like to gratefully acknowledge the support of the National Science Foundation (NSF IRI 96-23913).

References

[1] J. Ben-Arie, "The Probabilistic Peaking Effect of Viewed Angles and Distances with Application to 3D Object Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12 no. 8, pp. 760–774, Aug. 1990.

[2] R. Bergevin and M.D. Levine, "Part Decomposition of Objects from Single View line Drawings" *CVGIP: Image Understanding,* vol. 55, no. 1, pp. 73–83, Jan. 1992.

[3] R. Bergevin and M.D. Levine, "Generic Oobject Recognition: Building and Matching Coarse 3D Descriptions from Line Drawings," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, no. 1, pp.19–36, Jan. 1993.

[4] I. Biederman, "Human Image Understanding: Recent Research and a Theory," *Computer Vision, Graphics, and Image Processing,* vol. 32, pp. 29–73, 1985.

[5] T. Binford, "Inferring Surfaces from Images," *Artificial Intelligence,* vol. 17, pp. 205–244, 1981.

[6] K. Brunnström, T. Lindeberg, and J. Eklundh, "Active Detection and Classification of Junctions by Foveation with a Hand-Eye System Guided by the Scale-Space Primal Sketch," Technical Report ISRN KTH/NA/P-91/31-SE, Computer Vision and Active Perception Laboratory (CVAP), Royal Inst. of Technology, Stockholm, 1991.

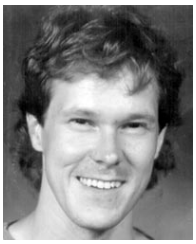
[7] J. Burns and L. Kitchen, "Recognition in 2D Objects from Large Model Bases Using Prediction Hierarchies," *Proc. Int'l Joint Conf. Artificial Intelligence,* pp. 763–766, Milan, Italy, 1987.

[8] J. Burns, R. Weiss, and E. Riseman, "View Variation of Point-Set and Line-Segment Features," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, no. 1, pp. 51–68, Jan. 1993.

[9] I. Chakravarty and H. Freeman, "Characteristic Views as a Basis for Three-Dimensional Object Recognition," *Proc. SPIE Conf. Robot Vision,* pp. 37–45, Arlington, Va., 1982.

[10] S. Dickinson, H. Christensen, J. Tsotsos, and G. Olofsson, "Active Object Recognition Integrating Attention and Viewpoint Control," *Proc. European Conference Computer Vision, ECCV '94,* Stockholm, May 1994.
[11] S. Dickinson, H. Christensen, J. Tsotsos, and G. Olofsson, "Active Object Recognition Integrating Attention and Viewpoint Control," *Computer Vision and Image Understanding,* vol. 67, no. 3, pp. 239–260, Sept. 1997.
[12] S. Dickinson, A. Pentland, and A. Rosenfeld, "A Representation for Qualitative 3D Object Recognition Integrating Object-Centered and Viewer-Centered Models," Technical Report CAR-TR-453, Center for Automation Research, Univ. of Maryland, 1989.
[13] S. Dickinson, A. Pentland, and A. Rosenfeld, "A Representation for Qualitative 3D Object Recognition Integrating Object-Centered and Viewer-Centered Models," *Vision: A Convergence of Disciplines,* K. Leibovic, ed. New York: Springer Verlag, 1990.
[14] S. Dickinson, A. Pentland, and A. Rosenfeld, "From Volumes to Views: An approach to 3D Object Recognition," *CVGIP: Image Understanding,* vol. 55, no. 2, pp. 130–154, 1992.
[15] S. Dickinson, A. Pentland, and A. Rosenfeld, "3D Shape Recovery Using Distributed Aspect Matching," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 14, no. 2, pp. 174–198, Feb. 1992.
[16] D. Eggert and K. Bowyer, "Computing the Orthographic Projection Aspect Graph of Solids of Revolution," *Pattern Recognition Letters,* vol. 11, pp. 751–763, 1990.
[17] D. Eggert, K. Bowyer, C. Dyer, H. Christensen, and D. Goldgof, "The Scale Space Aspect Graph," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, no. 11, pp. 1,114–1,130. Nov. 1993.
[18] R. Fairwood, "Recognition of Generic Components Using Logic-Program Relations of Image Contours," *Image and Vision Computing,* vol. 9, no. 2, pp. 113–122, 1991.
[19] Z. Gigus, J. Canny, and R. Seidel, "Efficiently Computing and Representing Aspect Graphs of Polyhedral Objects," *Proc. Second Int'l Conf. Computer Vision,* pp. 30–39, 1988.
[20] Z. Gigus and J. Malik, "Computing the Aspect Graph for Line Drawings of Polyhedral Objects," *Proc. Computer Vision and Pattern Recognition,* pp. 654–661, Ann Arbor, Mich., 1988.
[21] J. Hummel and I. Biederman, "Dynamic Binding in a Neural Net Model for Shape Recognition," *Psychological Review,* vol. 99, pp. 480–517, 1992.
[22] D. Huttenlocher and S. Ullman, "Recognizing Solid Objects by Alignment with an Image," *Int'l J. Computer Vision,* vol. 5, no. 2, pp. 195–212, 1990.
[23] K. Ikeuchi and T. Kanade, "Automatic Generation of Object Recognition Programs," *Proc. IEEE,* vol. 76, pp. 1,016–1,035, 1988.
[24] J. Kender and D. Freudenstein, "What is a Degenerate View?" *Proc. 10th Int'l Joint Conf. Artificial Intelligence,* pp. 801–804, Milan, Italy, Aug. 1987.
[25] D. Kriegman and J. Ponce, "Computing Exact Aspect Graphs of Curved Objects: Solids of Revolution," *Int'l J. Computer Vision,* vol. 5, no. 2, pp. 119–135, 1990.
[26] Y. Lamdan, J. Schwartz, and H. Wolfson, "Affine Invariant Model-Based Object Recognition," *IEEE Trans. Robotics and Automation,* vol. 6, no. 5, pp. 578–589, Oct. 1990.
[27] D. Lowe, *Perceptual Organization and Visual Recognition.* Norwell, Mass.: Kluwer Academic, 1985.
[28] H. Plantinga and C. Dyer, "Visibility, Occlusion, and the Aspect Graph," *Int'l J. Computer Vision,* vol. 5, no. 2, pp. 137–160, 1990.
[29] N. Raja and A. Jain, "Recognizing Geons from Superquadrics Fitted to Range Data," *Image and Vision Computing,* vol. 10, no. 3, pp. 179–190, Apr. 1992.
[30] N. Raja and A. Jain, "Obtaining Generic Parts from Range Images Using a Multiview Representation," *CVGIP: Image Understanding,* vol. 60, no. 1, pp. 44–64, July 1994.
[31] M. Sallam and K. Bowyer, "Generalizing the Aspect Graph Concept to Include Articulated Assemblies," *Pattern Recognition Letters,* vol. 12, pp. 171–176, 1991.
[32] I. Shimshoni and J. Ponce, "Finite Resolution Aspect Graphs of Polyhedral Objects," *Proc. IEEE Workshop on Qualitative Vision,* pp. 140–150, New York, June 1993.
[33] T. Sripradisvarakul and R. Jain, "Generating Aspect Graphs for Curved Objects," *Proc. IEEE Workshop Interpretation of 3D Scenes,* pp. 109–115, Austin, Tex., 1989.
[34] J. Stewman and K. Bowyer, "Direct Construction of the Perspective Projection Aspect Graph of Convex Polyhedra," *Computer Vision, Graphics, and Image Processing,* vol. 51, pp. 20–37, 1990.
[35] M. Swain, "Object Recognition from a Large Database using a Decision Tree," *Proc. DARPA Image Understanding Workshop,* pp. 690–696, Cambridge, Mass., 1988.
[36] D. Thompson and J. Mundy, "Model-Directed Object Recognition on the Connection Machine," *Proc. DARPA Image Understanding Workshop,* pp. 93–106, Los Angeles, 1987.
[37] J. Tsotsos, "An Inhibitory Beam for Attentional Selection," *Spatial Vision in Humans and Robots,* L. Harris and M. Jenkin, eds. Cambridge: Univ. Press, 1993.
[38] D. Weinshall and M. Werman, "On View Likelihood and Stability," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 2, pp. 97–108, Feb. 1997.
[39] D. Wilkes, "Active Object Recognition," PhD dissertation, 1994.
[40] D. Wilkes, S. Dickinson, and J. Tsotsos, "Quantitative Modeling of View Degeneracy," *Proc. Eighth Scandinavian Conf. Image Analysis,* Univ. of Tromsö, Norway May 1993.
[41] D. Wilkes, S. Dickinson, and J. Tsotsos, "A Quantitative Analysis of View Degeneracy and Its Application to Active Focal Length Control," *Proc. Int'l Conf. Computer Vision,* Cambridge, Mass., June 1995.
[42] D. Wilkes and J. Tsotsos, "Active Object Recognition," *Proc. Computer Vision and Pattern Recognition '92,* Urbana, Ill., June 1992.
[43] A. Witkin and J. Tenenbaum, "On the Role of Structure in Vision," *Human and Machine Vision,* J. Beck, B. Hope, and A. Rosenfeld, eds. New York: Academic Press, 1983.

Sven Dickinson received the BSc. degree in systems design engineering from the University of Waterloo, Canada, in 1983, and the MS and PhD degrees in computer science from the University of Maryland in 1988 and 1991, respectively. He is currently an assistant professor of computer science at Rutgers University, where he also holds a joint appointment in the Rutgers Center for Cognitive Science (RuCCS) and a membership in the Center for Discrete Mathematics and Theoretical Computer Science (DIMACS). He also currently holds or has held affiliations with the MIT Media Laboratory (visiting scientist, 1992–1994), the University of Toronto (visiting assistant professor, 1994–present), and the Center for Automation Research at the University of Maryland (assistant research scientist, 1993–1994, visiting assistant professor, 1994–1997). From 1991 to 1994, he was a research associate at the Artificial Intelligence Laboratory, University of Toronto. From 1985 to 1991, he was a graduate research assistant in the Computer Vision Laboratory, Center for Automation Research, University of Maryland. Prior to this, he worked in the computer vision industry, designing image processing systems for Grinnell Systems Inc., San Jose, California, 1983–1984, and optical character recognition systems for DEST, Inc., Milpitas, California, 1984–1985. His major field of interest is computer vision, with an emphasis on shape representation, object recognition, and mobile robot navigation. Dr. Dickinson was co-chair of both the 1997 and 1999 IEEE Workshops on Generic Object Recognition held in San Juan, Puerto Rico, and Corfu, Greece, respectively, while, in 1999, he co-chaired the DIMACS Workshop on Graph Theoretic Methods in Computer Vision. In 1996, he received the NSF Career award for his work in generic object recognition and currently serves as an associate editor of the *IEEE Transactions on Pattern Analysis and Machine Intelligence*.



David Wilkes received his PhD in computer science from the University of Toronto, where he did work in computer vision and robotics. He is currently president of Wilkes Associates Inc., a software and firmware development company specializing in end-to-end product development and project rescue.

John K. Tsotsos received his PhD in 1980 from the University of Toronto in computer science. He currently is a professor in that department and maintains a status professorship in the university's Department of Medicine. He has served on numerous conference committees and the editorial boards of *Computer Vision and Image Understanding*, *Image and Vision Computing Journal*, *Computational Intelligence*, and *AI & Medicine*. He is the general chair of the Seventh IEEE International Conference on Computer Vision, Corfu, Greece, 1999. His research focuses on biologically plausible models of visual attention, the development of a visually guided robot to assist physically disabled children, and perceptually guided robot control mechanisms.