A Quantitative Analysis of View Degeneracy and its Use for Active Focal Length Control

D. Wilkes
Dept. of Computer Science
University of Toronto
Toronto, CANADA M5S 1A4

S.J. Dickinson
Rutgers University
Center for Cognitive Science
Piscataway, NJ 08855-1179

J.K. Tsotsos
Dept. of Computer Science
University of Toronto
Toronto, CANADA M5S 1A4

Abstract

We quantify the observation by Kender and Freudenstein [6] that degenerate views occupy a significant fraction of the viewing sphere surrounding an object. This demonstrates that systems for recognition must explicitly account for the possibility of view degeneracy. We show that view degeneracy cannot be detected from a single camera viewpoint. As a result, systems designed to recognize objects from a single arbitrary viewpoint must be able to function in spite of possible undetected degeneracies, or else operate with imaging parameters that cause acceptably low probabilities of degeneracy. To address this need, we give a prescription for active control of focal length that allows a principled tradeoff between the camera field of view and probability of view degeneracy.

1 Introduction

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, and also discontinuities described by line segments or curves. Such features may be difficult both to identify and to interpret due to the accidental alignment of spatially distinct scene parts belonging to a single object or multiple objects, as seen from the camera viewpoint. This alignment is often referred to as *view degeneracy*. This superposition of features can lead to erroneous part counts or bad parameterizations of a model. For example, if two object edges are abutted and collinear, one longer edge is seen in their place.

Given arbitrarily high resolution, such alignments would not be a problem, because they would occur only for a vanishingly small fraction of the possible viewpoints [1]. Unfortunately, cameras and feature-extracting operators have finite resolution. We shall see in this paper that this enlarges the set of viewpoints that are problematic.

We begin by defining view degeneracy precisely, and showing how our definition covers many interesting classes of degeneracy. We develop a model of 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 detectible as distinct features, and camera focal length. The model gives the probability of being in a degenerate view as a function of the model parameters. The evaluation of the model at realistic parameterizations indicates that degenerate views can have significant probabilities.

We will show that there is a significant additional problem for recognition systems operating from a single, arbitrary viewpoint. Namely, view degeneracy is not detectible from a single viewpoint. As a result, single-viewpoint systems must be able to perform unambiguous object identification in spite of possible degeneracies. The alternative is to minimize the probability of degeneracy. We will show that the probability of degeneracy is very sensitive to camera focal length. This will lead finally to a prescription for focal length control that allows a tradeoff to be achieved between the width of the field of view and probability of view degeneracy.

1.1 Definitions of degeneracy

Kender and Freudenstein [6] 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. Figure 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:

1. a zero-dimensional (point-like) object feature is collinear with another zero-dimensional object feature and the front nodal point of the lens
2. a zero-dimensional object feature is collinear with some point on either:
    a. an infinite line, or
    b. a finite line segment

   defined by two zero-dimensional object features, and the front nodal point

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:

Vertices imaged in the plane of scene edges This is encompassed by case 2b above, and is depicted in figure 2. The lower right scene edge on the upper object is the line segment of case 2b. The vertex on the lower object is the zero-dimensional object feature.

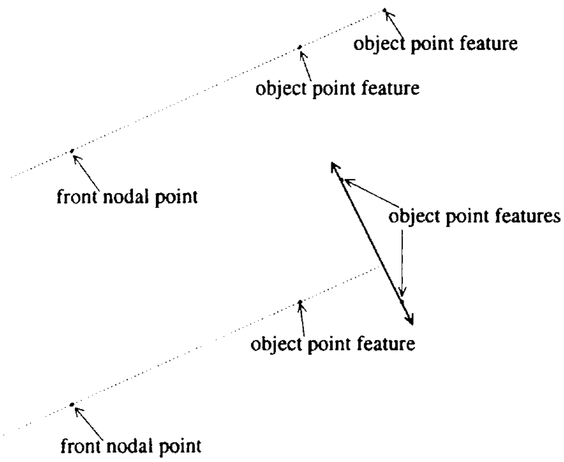

Figure 1: Our definition of view degeneracy. The two cases are the collinearity of the front nodal point of the lens with two object point features (top) or with one object point feature and a point on one line (bottom) or line segment.

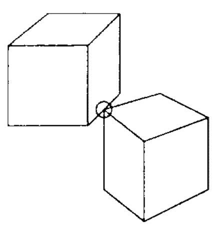

Figure 2: An example of view degeneracy in which a vertex is imaged in the plane of a scene edge

Parallel scene lines imaged in their own plane We depict this situation in figure 3. This is case 2a, with an infinite line. 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).

Coincident scene lines imaged in their own plane This generalizes to *coplanar lines imaged in their own plane.* We depict this situation in figure 4. This is equivalent to case 2b, using an infinite line. 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.

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 figure 5. In the case of a 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 2a, 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 extractible directly by low-level vision. The use of abstract object points rather than obvious feature points has no impact on the estimation of probabilities 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.

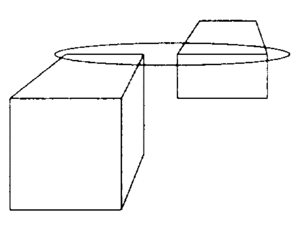

Figure 3: An example of view degeneracy in which parallel scene lines are imaged in their own plane

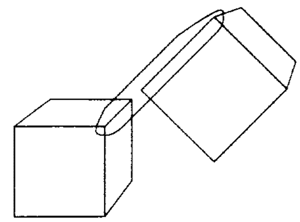

Figure 4: An example of view degeneracy in which coplanar scene lines are imaged in their own plane

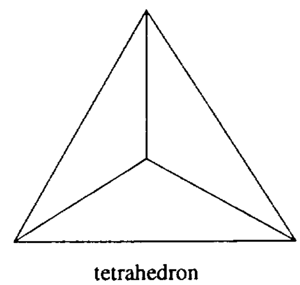

Figure 5: An example of view degeneracy in which perfect symmetry gives preference to a flat, 2D interpretation of the lines.

1.2 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. Figure 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 2a degeneracy occurs at a circle of points on the sphere surface. Each case 2b degeneracy occurs on a sector of a circle of points on the sphere surface. Since these loci of points are of lower dimensionality than the viewing sphere itself, the probability that the viewpoint is on such loci is vanishingly small.

Infinitesimal probabilities of view degeneracy would potentially allow strong inferences to be made about scene structure [1]. Consider a certain coincidence, C, of simple object features A and B. If C occurs in the image much more frequently than expected due to view degeneracy, then many of the occurrences must be due to a regularity in the domain of interest (i.e. the frequent physical coincidence of A and B). This regularity is useful for recognition if C reduces the uncertainty in object identity. The utility of C for recognition may be expressed more formally as the average number of bits of information gained by observing C. An expression for this is:

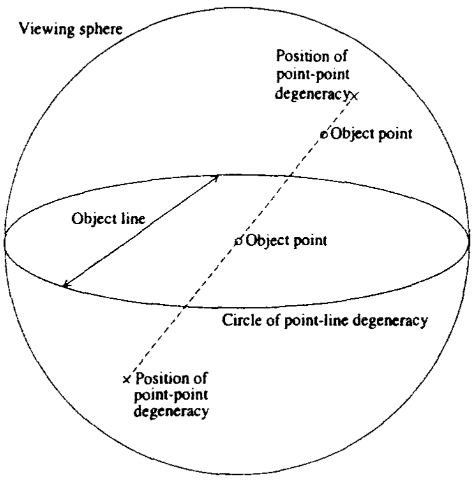

Figure 6: Loci of degeneracy with infinite resolution. This illustrates degeneracies of both case 1 type (involving two 0D object features) and case 2 type (involving a 0D object feature and an infinite line defined by object features).

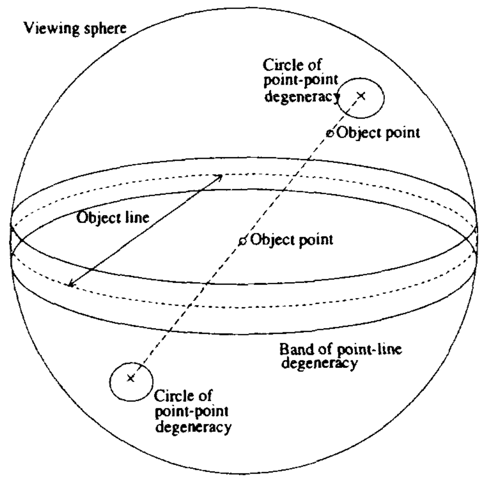

Figure 7: The loci of degeneracy of the previous figure, with finite resolution. The regions of the viewing sphere affected by degeneracy are large. The circle of the previous figure is actually a thick band on the sphere. The points from the case 1 degeneracy are actually disks.

$$-\sum_i Pr(object\ i)\log\left(\frac{Pr(object\ i|C)}{Pr(object\ i)}\right) \qquad (1)$$

or

$$\sum_i Pr(object\ i)\log\left(\frac{Pr(object\ i)}{Pr(object\ i|C)}\right) \qquad (2)$$

This is also called the asymmetric divergence, or mean information gain per observation of C [7].

Kender and Freudenstein [6] 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. Figure 7 illustrates this thickening. This would have a negative impact on systems that assume no degeneracy. If the coincidence C discussed above occurs more frequently due to degeneracy than expected, then expression 2 above will be smaller than expected, since the apparent coincidence of component features A and B will now occur more often for objects in which A and B are physically separated than with infinite resolution.

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.

1.3 Overview

The reason for using the above definitions of degeneracy and effective resolution is that they have allowed us to construct a system-independent, quantitative model of view degeneracy. We can predict the probability of being in a degenerate view as a function of the geometric parameters of the key object features and the effective resolution. This is useful, because it allows one to decide the importance of degenerate views in various imaging situations. In particular, we note that degenerate views have significant probability in many realistic situations.

In the next section, we present our model and its analysis. Section 3 presents an example parameterization of the model corresponding to a real application. The result shows a significant probability of view degeneracy. The implications of this for recognition systems is discussed in Section 4, along with various means of coping with degeneracy. The sensitivity of the probability of view degeneracy to focal length leads to a prescription for focal length control, discussed in Section 5. Finally, we conclude with a recapitulation of the results.

2 Model

Degeneracy based on a pair of points Figure 8 shows the perspective projection of a pair of feature points, A and B. For simplicity, our analysis assumes that A is centred in the image. 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

$$\tan\ \alpha = \frac{s}{f} \qquad (3)$$

and that

$$\tan\ \alpha = \frac{\sqrt{r^2\sin^2\phi + r^2\cos^2\phi\sin^2\theta}}{R + r\cos\phi\cos\theta} \qquad (4)$$

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.

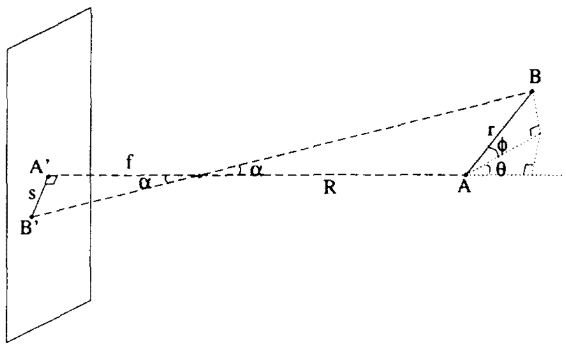

Figure 8: Perspective projection of a pair of points. We will determine the fraction of all orientations of line AB in relation to the camera for which the image separation s of A and B is small enough to interfere with their resolution into distinct features.

The choice to solve for ϕ is arbitrary. We obtain two solutions, ϕ_0 and ϕ_1, at which $s = s_c$:

$$\phi_0 = \cos^{-1}\left(\frac{-Rs_c^2 + f\sqrt{r^2f^2 - s_c^2R^2 + s_c^2r^2}}{r(f^2+s_c^2)\cos\theta}\right) \tag{5}$$

$$\text{or } \phi_1 = \pi - \cos^{-1}\left(\frac{Rs_c^2 + f\sqrt{r^2f^2 - s_c^2R^2 + s_c^2r^2}}{r(f^2+s_c^2)\cos\theta}\right) \tag{6}$$

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. Figure 9 shows the sections of the sphere swept out by point B in Figure 8 as θ ranges over a small 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

$$g(\theta) = \frac{1}{total\ area} \times (shaded\ area) \tag{7}$$

$$= \frac{1}{2r^2\Delta\theta} \times \left(\int_0^{\phi_0}\int_{\theta_0}^{\theta_0+\Delta\theta} r^2\cos\phi\, d\theta d\phi + \int_{\phi_1}^{\pi}\int_{\theta_0}^{\theta_0+\Delta\theta} r^2\cos(\pi-\phi)\, d\theta d\phi\right) \tag{8}$$

$$= \frac{1}{2r^2\Delta\theta} \times r^2\left(\int_0^{\phi_0}\Delta\theta\cos\phi\, d\phi + \int_{\phi_1}^{\pi}\Delta\theta\cos(\pi-\phi)\, d\phi\right) \tag{9}$$

$$= \frac{1}{2\Delta\theta} \times \Delta\theta(\sin\phi_0 + \sin(\pi-\phi_1)) \tag{10}$$

$$= \frac{\sin\phi_0 + \sin\phi_1}{2} \tag{11}$$

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 $s < s_c$.

Degeneracy based on a point and a line Figure 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,

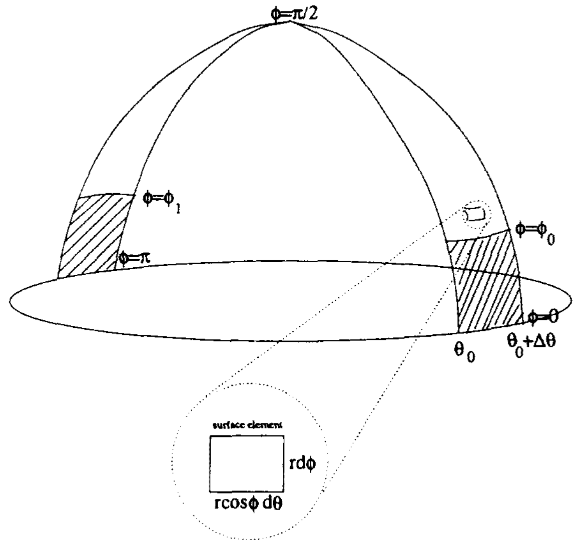

Figure 9: The regions of the sphere swept by point B for $\theta \in [\theta_0, \theta_0+\Delta\theta)$ and $\phi \in [0,\pi]$. The shaded parts are the areas for which $s < s_c$.

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

$$s = r\cos\phi\sin\theta\frac{f}{R + r\cos\phi\cos\theta} \tag{12}$$

$$t = r\sin\phi\frac{f}{R + r\cos\phi\cos\theta} \tag{13}$$

$$a = l\sin\omega\frac{f}{R + l\cos\omega} \tag{14}$$

$$b = \sqrt{s^2 + (t-a)^2} \tag{15}$$

$$c = \sqrt{s^2 + t^2} \tag{16}$$

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

$$d = \begin{cases} s & \text{if } t \geq 0 \text{ and } t \leq a \\ b & \text{if } t > a \\ c & \text{if } t < 0 \end{cases} \tag{17}$$

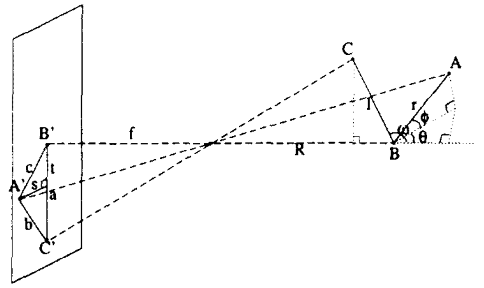

Figure 10: Perspective projection of a line segment and a point. We will determine the fraction of all possible orientations of A with respect to line BC for which the image separation of A and BC is small enough to interfere with the detection of A as a separate feature.

We considered two possibilities for the integration for the probability that d is less than some critical value d_c. The approach that is analogous to the approach used in the first case leads to numerical evaluation of a three-dimensional integral. An easier method would be to perform 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 $d < d_c$.

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 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 2a 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

$$p_d = 1.0 - \underbrace{(1.0 - p_{00})^{n_r}}_{\textit{no symmetry degeneracies}} \times \overbrace{(1.0 - p_{01})^{n_f}}^{\textit{no face degeneracies}} \quad (18)$$

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 provide a pessimistic upper bound on the probability of degeneracy.

3 Parameterization of the model

We have parameterized the model to match the situation in our own experimental setup, as described in detail in [8], and presented in overview in [9]. We have used the actual camera pixel size to compute the critical feature separation s_c. The object-related parameters are those of an object set, consisting of origami figures, that is used in our experiments. Figure 11 tabulates the results with the basic parameterization, and examines the sensitivity of the probability of a degeneracy to each of the model parameters by perturbing each parameter in isolation.

We see that (for example) for a model system with a parameterization to recognize small tabletop objects from a range of half a metre, there are one in five odds that the view encountered will be degenerate. The sensitivity analysis demonstrates that there are realistic parameterizations of the model for which probabilities of degeneracy are very significant. We also tried a parameterization with f and s_c set to match human foveal acuity of twenty seconds of arc. The probabilities of degeneracy were negligably small. This may explain why the importance of degenerate views to computer vision has traditionally been underestimated.

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 recognition algorithms. If one excludes degenerate views from the object representations or viewpoints used by a system, then one must have a mechanism for *detecting* degeneracy when it occurs, and *moving* out of the degenerate view position. It is important to note that *even detection of degeneracy requires camera motion*: Using our definition of view degeneracy, we have

Theorem: It is impossible to detect view degeneracy from a single camera viewpoint.

Basic parameterization	
parameter	value
f	0.007 metres
R	0.50 metres
r	0.04 metres
l	0.04 metres
s_c	0.0001548 metres (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}

Figure 11: Results for a tabletop vision system. The basic parameterization corresponds to our experimental setup. One fifth of the views on a single object are degenerate. The probability of degeneracy varies significantly as model parameters are perturbed.

Proof: Assume to the contrary, that we have a "degeneracy detector" D, that takes a single image as input, and determines whether it contains instances of view degeneracy. We may cause D to fail as follows. We construct a pair of scenes such that each scene results in the same camera image, but one is degenerate and the other is not. Then, whether D reports that the image is degenerate or not, D will be in error on one of the two scenes. Figure 12 diagrams such a pair of scenes. The idea is to acquire a camera image at a position of degeneracy, then print the image acquired. If one then acquires an image of the print from an appropriate face-on viewpoint, the result can be made indistinguishable from the image of the original scene. The first image, a face-on view of a cube, is degenerate: a small motion of the camera will change the number of faces of the cube that are visible. This second image, however, is from a non-degenerate viewpoint: small motions of the camera result only in minor distortion of the camera image. In fact, changes in viewpoint with respect to the print of the scene of up to 90 degrees result in smooth changes in the image features. □

For recognition to be successful from a single viewpoint, therefore, degenerate views must be usable even if the degeneracy is not explicitly detectible.

Single view approaches to recognition can be successful provided that the camera view, even if degenerate, is not ambigu-

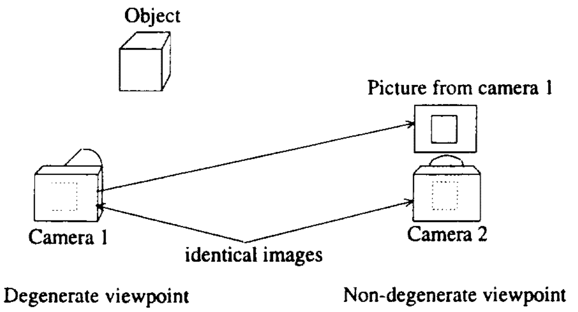

Figure 12: The camera in the left scene is at a position of degeneracy. The camera in the right scene is not, (since no alignments are present) yet the two camera images are identical. This demonstrates that degeneracy is not detectable from a single viewpoint.

ous in the sense of belonging to more than one stored object model. If, the mapping from a recovered view to an object model, or part, is ambiguous, then the various interpretations should be rank-ordered to provide a more efficient search [3, 4]. A more effective solution to dealing with ambiguity is to take an active approach. Wilkes [8] has explored the idea of moving the camera to a viewpoint advantageous to recognition. This approach also meets with success, but at the cost of significant viewpoint changes. Dickinson, Christensen, Tsotsos, and Olofsson [5] use a probabilistic aspect graph to guide the camera from a position where the view is ambiguous to a position where the view is unambiguous.

A final alternative is somehow to reduce the probability of view degeneracy. We see from the sensitivity analysis in Section 3 that the probability of degeneracy p_d is sensitive to both focal length and effective resolution. Thus, we may consider controlling the aiming of a variable-resolution (foveated) sensor, using an attentional mechanism of some sort, or we may control focal length. This latter approach is the subject of the next section.

5 Prescription for focal length control

There is a clear tradeoff between wide angle of view and probability of view degeneracy. Wide angle is desirable for initial object detection, in the presence of uncertainty in camera or object position, because 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. A similar tradeoff has been noted in the work of Brunnström [2] who uses a wide angle of view to detect object junctions and then zooms in to increase image resolution for junction clasification. The existence of such a tradeoff suggests that there can be a principled choice of focal length in many circumstances.

Suppose that we have known constants p_0, the maximum acceptable probability of view degeneracy, and s_c, the critical separation of image features. Then the focal length should be set to the smallest value that gives a probability of degeneracy p_d no larger than p_0. This may be done automatically by a vision system, provided that the necessary parameters of our model can be determined.

We may rewrite all of the expressions involving parameters l, R and r (equations 5, 6, 12, 13, 14) in terms of the ratios $\frac{r}{R}$ and $\frac{l}{R}$:

$$\phi_0 = \cos^{-1}\left(\frac{-\left(\frac{r}{R}\right)^{-1} s_c^2 + f\sqrt{f^2 - s_c^2\left(\frac{r}{R}\right)^{-2} + s_c^2}}{(f^2 + s_c^2)\cos\theta}\right) \quad (19)$$

$$or\ \phi_1 = \pi - \cos^{-1}\left(\frac{\left(\frac{r}{R}\right)^{-1} s_c^2 + f\sqrt{f^2 - s_c^2\left(\frac{r}{R}\right)^{-2} + s_c^2}}{(f^2+s_c^2)\cos\theta}\right) \tag{20}$$

$$s = \frac{r}{R}\cos\phi\sin\theta\frac{f}{1+\frac{r}{R}\cos\phi\cos\theta} \tag{21}$$

$$t = \frac{r}{R}\sin\phi\frac{f}{1+\frac{r}{R}\cos\phi\cos\theta} \tag{22}$$

$$a = \frac{l}{R}\sin\omega\frac{f}{1+\frac{l}{R}\cos\omega} \tag{23}$$

Thus, if we can estimate $\frac{r}{R}$ and $\frac{l}{R}$, we can completely parameterize the model, computing p_d via the above equations plus equations 11, 15, 16, 17, and 18.

If the current focal length f is known, then we may make a rough estimate of $\frac{l}{R}$ and $\frac{r}{R}$ as follows. We may sample the distribution for the apparent separation s of neighbouring features in the image, getting an average apparent separation $\bar{s}$; to minimize the effects of view degeneracy in this sampling, we should use a large focal length. If all pairs of adjacent features had exactly the separation r in 3-D space, and all relative orientations of such feature pairs in 3-D space were equally likely, then the expected value for the image separation s of feature pairs would be given by

$$E[s] = \int_0^{\frac{\pi}{2}} \underbrace{\frac{fr}{R}\cos\phi}_{\text{image length if pair tilted by }\phi} \times \overbrace{\cos\phi}^{\text{probability density for tilt }\phi}\, d\phi \tag{24}$$

$$= \frac{fr}{R}\times\frac{\pi}{4} \tag{25}$$

Thus $\bar{s} \approx E[s]$ gives us an estimate of $\frac{r}{R}$:

$$\frac{r}{R} = \frac{4\bar{s}}{\pi f} \tag{26}$$

We may set $\frac{l}{R} = \frac{r}{R}$ if the predominant degeneracies in the domain of interest are of the type of case 2b, or set $\frac{l}{R} = \infty$ if the predominant degeneracies are of the type of case 2a. Which parameterization is chosen typically does not affect the order of magnitude of the resulting p_d values. Figure 13 shows how p_d varies with focal length for our basic parameterization of Section 3. In general, the focal length f giving $p_d = p_0$ is found by applying a zero-finding method to the function $g(f) = p_d - p_0$, where p_d is found by evaluating our model as described above. Since the function is smooth and has only one zero, we recommend a Newton iteration $f_{n+1} = f_n - \frac{g(f_n)}{g'(f_n)}$, where f_0 is any possible focal length value for the system, and $g'(f_n)$ is the derivative of g with respect to f, approximated by $\frac{g(f+\epsilon)-g(f-\epsilon)}{2\epsilon}$ and the iteration is terminated when successive iterations change f by less than 2ϵ. To summarize, given a sampling of image feature separation at a known large focal length, we can solve for an updated focal length that gives a probability of view degeneracy not to exceed a given target value.

6 Conclusions

We have demonstrated that view degeneracy is a frequent enough occurrence to warrant consideration in the design of computer vision systems. The fact that view degeneracy is not detectible from a single viewpoint means that systems that attempt recognition from a single, arbitrary viewpoint must be able to do so from degenerate as well as non-degenerate views.

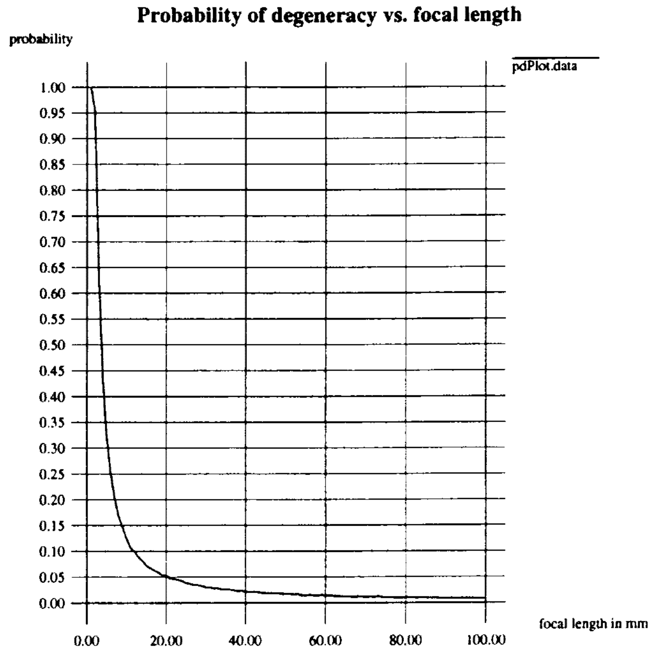

Figure 13: Probability p_d as a function of focal length f, at the basic parameterization.

The alternative is to take steps to minimize the probability of degeneracy.

The probability of view degeneracy may be reduced by several means. One is to move the camera to a viewpoint advantageous for recognition [9, 5]. Another is to use an attentional mechanism to aim a high-resolution fovea-like sensor for feature detection. A final, economical possibility is the control of focal length discussed above.

Acknowledgements The authors wish to thank the Natural Sciences and Engineering Research Council of Canada, the Information Technology Research Centre, and the Canadian Institute for Advanced Research 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 Centres of Excellence Program of the Government of Canada and the participation of PRECARN Associates Inc. The second author is the CP (Unitel) Fellow of the Canadian Institute of Advanced Research.

References

[1] Binford, T.O., "Inferring Surfaces from Images," *Artificial Intelligence*, Vol. 17, pp. 205-244. 1981.

[2] Brunnström, K., "Active Exploration of Static Scenes", Ph.D. Dissertation, Computational Vision and Active Perception Laboratory, Department of Numerical Analysis and Computing Science, Royal Institute of Technology, Sweden, October, 1993.

[3] Dickinson, S.J., Pentland, A.P., and Rosenfeld, A., "3-D Shape Recovery Using Distributed Aspect Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, February, 1992.

[4] Dickinson, S.J., Pentland, A.P., and Rosenfeld, A., "From volumes to views: An approach to 3-D object recognition." *CVGIP: Image Understanding*, Vol. 55, No. 2, 1992.

[5] Dickinson, S., Christensen, H., Tsotsos, J., and Olofsson, G., "Active Object Recognition Integrating Attention and View-point Control", Proceedings, ECCV '94, Stockholm, May, 1994.

[6] Kender, J., and Freudenstein, D., "What is a 'degenerate' view?" *Proc. DARPA Image Understanding Workshop*, pp. 589-598, Los Angeles, CA, February 1987.

[7] Kullback, S., *Information Theory and Statistics*, Dover Publications, New York, 1968.

[8] Wilkes, D., *Active Object Recognition*, Ph.D. Dissertation, Department of Computer Science, University of Toronto, 1994.

[9] Wilkes, D., and Tsotsos, J.K., "Active Object Recognition," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition '92*, IEEE Computer Society Press, Los Vasqueros, CA, 1992.