Modal Structure and Reliable Inference

Allan Jepson, Whitman Richards, and David Knill

1 Introduction

The world we live in is a very structured place. Matter does not flit about in space and time in a completely unorganized fashion, but rather is organized by the physical forces, biological processes, social interactions, and so on which exist in our world (McMahon, 1975; Thompson, 1952). It is this structure, or regularity, which makes it possible for us to make reliable inferences about our surroundings from the signals taken in from various senses (Marr, 1982; Witkin and Tenenbaum, 1983). In other words, regularities in the world make sense data reliably informative about the world we move around in. But what is the nature of these regularities, and how can they be used for the purposes of perception?

In this chapter,[1] we consider one class of environmental regularities which arise from what we call the modal structure of the world and which has the effect of making sensory information for certain types of perceptual judgements highly reliable (Bobick and Richards, 1986). Our definition of modal regularities is motivated by careful analyses of some simple examples of reliable perceptual inferences. Given the resulting definition, we then briefly discuss some of the implications for the knowledge required of a perceiver in order for it to make reliable inferences in the presence of such modal structure.

2 Modal structure: An example.

2.1 When can we infer that an object is stationary?

A common perceptual inference is that of whether an object is moving or at rest. How can we make this inference given only the two-dimensional projection of a three-dimensional object? When the image of an object is moving the inference is trivial, since a stationary object cannot give rise to image motion (assuming a stationary observer). What about the case in which the image of an object is not moving? Can we then reasonably infer that the object itself is not moving? Intuitively, this is a common inference which is rarely wrong, so it seems that the answer to the question is a strong "yes". We will show that object motions in the world must have a strong type of regularity to support this inference. We will then generalize the discussion to show that the same type of structure in the world must be present to support a wide variety of similarly reliable inferences.

---

[1]This article is to appear in the book "Perception as Bayesian Inference" edited by D. Knill and W. Richards, Cambridge Univ. Press.

One explanation for the reliability of inferences like "stationary in the image implies stationary in the world" is that a stationary image would be an "accidental" view of a moving object (Albert and Hoffman, 1991; Lowe, 1985). That is, one would have to be viewing the object head-on (in the direction of the motion), in order to obtain a stationary image of a moving object. Small perturbations of the viewpoint would destroy the stationarity of the image motion when viewing a moving object. Moreover, only two viewpoints of the infinite number of possible viewpoints would have this special property, thus the probability of obtaining a stationary image from a moving object is zero (or, if we allow for some uncertainty in the measurement of image motion, the probability would be small). On the face of it, this argument is attractive and seems perfectly logical. Unfortunately, the inference described is based on the wrong probability distribution; namely, the probability of obtaining a stationary image, given that an object is moving. Typically, this distribution is called the likelihood function, and we can characterize the inference as being within the class of maximum-likelihood criteria for making statistical decisions. The distribution of interest, however, is the posterior distribution: the probability that an object is moving, given that the image is stationary (Jepson and Richards, 1992; Knill and Kersten, 1991). The difference between the two is critical, as we will show.

We begin by formalizing the general viewpoint argument as a statistical decision based on the ratio of likelihood functions, $p(||\vec{v}_{image}|| = 0 \,|\, ||\vec{v}|| = 0)$ and $p(||\vec{v}_{image}|| = 0 \,|\, ||\vec{v}|| \neq 0)$, where $\vec{v}$ is the 3-dimensional velocity vector of an object, $\vec{v}_{image}$ is the 2-dimensional image velocity vector, and $||\cdot||$ is the usual 2-norm. The general viewpoint argument says that the assumption of a generic view makes the ratio of the two likelihood functions large, so that one can reliably infer that the object is stationary given a stationary image. The likelihood ratio is given by

$$R_{likelihood} = \frac{p(||\vec{v}_{image}|| = 0 \,|\, ||\vec{v}|| = 0)}{p(||\vec{v}_{image}|| = 0 \,|\, ||\vec{v}|| \neq 0)}. \tag{1}$$

The numerator is clearly equal to 1, while the denominator, under the generic viewpoint assumption, is equal to 0, so, the argument goes, one should clearly infer that an object is stationary when one detects no image motion. A seeming advantage of this approach is that it does not appear to depend on any knowledge about the world (with the exception of assuming a generic view). It suggests that one can make a rational decision without knowing about the probabilistic structure of object motions in the world; that is, without assuming any constraints on object motion. We will now proceed to show that this is false, and that, if one assumes a generic, relatively unconstrained model of object motion, the inference of object stationarity from image stationarity will be wrong most of the time (and in the limit, all of the time). In fact, with such a model of object motion, one should infer the opposite, that the object is moving.

Our generic model of object motion is that of idealized gas molecules within a container having constant temperature T. Maxwell's distribution provides the probability density for the 3D velocity of a single molecule as

$$p(\vec{v}) = \frac{1}{(2\pi cT)^{\frac{3}{2}}} exp(-||\vec{v}||^2/(2cT)), \tag{2}$$

where c is a positive constant (Fowler and Guggenheim, 1952). This is just an isotropic

Gaussian distribution in three variables, having mean zero and variance cT. Note that the most probable velocity corresponds to the mean, namely $\vec{v} = 0$. We define the molecule to be at rest whenever the speed is less or equal to some tolerance $\epsilon > 0$. Given this definition, it follows that the prior probability of the particle being "at rest" is nonzero and roughly proportional to ϵ^3 for small values of ϵ (in our notation, $p(\text{at rest}) = p(||\vec{v}|| < \epsilon) = O(\epsilon^3)$). We assume that ϵ is significantly smaller than the standard deviation of the probability distribution, namely $\sqrt{cT}$, since otherwise, being at rest would have no significant meaning.

Suppose our observer has orthographic projection (for simplicity), and can measure the first two components of $\vec{v}$ ($\vec{v}_{image}$) such that the error in the estimation of the image speed, $||\vec{v}_{image}||$, is no larger than $\epsilon > 0$, where we assume some image sensing noise to make the mathematical analysis simpler (we can still examine the limit as $\epsilon \to 0$). Consider the specific case in which the observer measures the image speed to be between 0 and ϵ. Can the observer then infer that the particle is at rest?

The appropriate computation to make is the conditional probability that the particle is at rest, given the data that the image speed has been observed to be less than ϵ. We denote this distribution by $p(||\vec{v}_{world}|| < \epsilon \,|\, ||\vec{v}_{image}|| < \epsilon)$. It is just $1 - p(||\vec{v}_{world}|| > \epsilon \,|\, ||\vec{v}_{image}|| < \epsilon)$, where the latter term is the conditional probability that the particle is not at rest, given the same observation. As with the likelihood ratio analysis, we find it convenient to compute the ratio of these two probabilities. For small values of ϵ we find

$$\frac{p(||\vec{v}_{world}|| < \epsilon \,|\, ||\vec{v}_{image}|| < \epsilon)}{p(||\vec{v}_{world}|| > \epsilon \,|\, ||\vec{v}_{image}|| < \epsilon)} = \frac{p(||\vec{v}_{world}|| < \epsilon \,|\, ||\vec{v}_{image}|| < \epsilon)}{1 - p(||\vec{v}_{world}|| < \epsilon \,|\, ||\vec{v}_{image}|| < \epsilon)} \approx \epsilon / \sqrt{2\pi cT} \qquad (3)$$

to leading order. Since we have assumed that ϵ is significantly smaller than the standard deviation of the distribution, namely $\sqrt{cT}$, it follows that the above ratio of conditional probabilities is significantly smaller than 1. That is, the odds strongly favour the interpretation that the particle is actually moving, even though the image motion is consistent with it being at rest. This is not a problem with the accuracy of the motion measurements. In fact, the same probability ratio is obtained (to leading order) even if the first two components of the velocity are assumed to be measured within a tolerance of some δ taken significantly smaller than ϵ. The difficulty is that the third component of the velocity is not measured at all, and can vary according to the Gaussian distribution with standard deviation $\sqrt{cT}$. For small values of ϵ this third component will rarely fall sufficiently close to zero for the particle to qualify to be at rest. The odds, therefore, consistently favour the inference that the particle is moving no matter how accurately the image velocity is measured. Moreover, having more time frames in which the image motion is measured to be less than δ can actually decrease the odds that the object is at rest.

Our model of the world is clearly missing something which would allow reliable inferences of object stationarity to be made. Since such inferences in our own world our commonplace and are usually correct, the unconstrained model of object motions just presented does not suffice to characterize our world. Similar results may be obtained for a wide range of normally reliable perceptual inferences, including so called "non-accidental properties" such as the colinearity or cotermination of two line segments (see Table 1). What type structure is needed to make such inferences reliable?

2.2 A Bayesian analysis

We wish to explore conditions that must be satisfied in order for an observer to be able to make a reliable inference of, say, a particle being at rest. In fact, it is convenient to approach the problem in a slightly more general fashion. We suppose that, at least in a restricted context C, the occurrence of a world property P can be modeled using the probability $p(P|C)$, and it's absence by $p(notP|C)$. Suppose that some measurements are taken of the objects and events in the world. We refer to a particular collection of such measurements as a feature F. Hence a feature will be identified with the set of all world events having measurements specified by F, and thus probabilities such as $p(F|C)$ are well defined. We wish to study the inference that property P occurs in the world, given both that the world context is C and that the measurements F are satisfied. Note that the probabilities $p(P|C)$ and $p(F|C)$ are considered to be objective facts about the world (or at least an idealization of the world), and are *not* statements about the perceiver's model of the world. Here we keep the issue of whether or not a perceiver needs to use any probabilistic model of the world quite separate from our analysis of a good inference.

In the probabilistic formalism a measure of the success of inferring property P from F is the a posteriori probability of P given the feature F in the context C. A reliable inference makes this probability, namely $p(P|F\&C)$, nearly one, and the probability of an error, namely $p(notP|F\&C)$, nearly zero. It is convenient to consider the ratio of these two quantities, that is

$$R_{post} = \frac{p(P|F\&C)}{p(notP|F\&C)}. \tag{4}$$

We consider the feature F to provide a reliable inference, in the context C, precisely when this probability ratio R_{post} is much larger than one. Below we consider how such a condition can be ensured.

Bayes' rule can be used to break down the probability ratio R_{post} into two components. The first component, $R_{likelihood}$, is a likelihood ratio and relates to the measurement F of property P. The second component is another probability ratio, R_{prior}, and specifies the relative probabilities of occurrence of P and $notP$ in context C. The decomposition of R_{post} has the simple form:

$$R_{post} = R_{likelihood} \cdot R_{prior} \quad . \tag{5}$$

Here the prior probability ratio R_{prior} is given by (compare equation (4))

$$R_{prior} = \frac{p(P|C)}{p(notP|C)}. \tag{6}$$

and the likelihood ratio $R_{likelihood}$ is defined to be

$$R_{likelihood} = \frac{p(F|P\&C)}{p(F|notP\&C)}. \tag{7}$$

From equation (5) we see that the likelihood ratio $R_{likelihood}$ acts as an amplification factor on the prior probability ratio R_{prior}. In words, we obtain a reliable inference only when the product of the likelihood ratio and the prior probability ratio is significantly larger than one.

To be concrete, consider the gas particle observer discussed in the previous section. In this case the context C denotes the fact that we are using Maxwell's distribution for the particle's velocity. The numerator of the likelihood ratio, $p(F|P\&C)$, is the probability of observing F, namely the image speed to be less than ϵ, given that the object is actually at rest. For the measurement accuracy of ϵ, and a particle moving with 3D speed no larger than ϵ, at least half of the measurements will show an image velocity having a speed of less than ϵ. The actual value doesn't matter for our current argument, so long as it is bounded away from zero. The denominator of the likelihood ratio, $p(F|notP\&C)$, is the probability that the image velocity is small even though the particle is not at rest. This probability is proportional to the square of the tolerance for the image motion, namely ϵ^2. Thus we find that the likelihood ratio is proportional to $1/\epsilon^2$, and is therefore large for sufficiently small values of ϵ. In other words, the image feature F in this case is much more likely to come from a particle at rest than from a moving particle, as expected from our earlier discussion of the likelihood ratio test.

But equation (5) has shown that the likelihood ratio does not provide the whole story, rather it acts as an amplification factor on the prior probability ratio for P versus $notP$ in context C. For our gas particle example, property P denotes that the particle has a velocity of magnitude less than ϵ, and the context specifies that the velocity is distributed according to Maxwell's equation. Therefore R_{prior} is simply

$$\frac{p(P|C)}{p(notP|C)} = \frac{p(||\vec{v}_{world}|| < \epsilon)}{p(||\vec{v}_{world}|| > \epsilon)} = \frac{p(||\vec{v}_{world}|| < \epsilon)}{1 - p(||\vec{v}_{world}|| < \epsilon)} \approx \epsilon^3/(2\pi cT)^{\frac{3}{2}}, \tag{8}$$

to leading order. It should now be clear what the problem is with the inference that the particle is stationary, given that it's image is stationary, in this gas particle context. In particular, note that as $\epsilon \rightarrow 0$ the prior probability ratio R_{prior} decreases to zero like ϵ^3, while we showed in the previous paragraph that the likelihood ratio $R_{likelihood}$ increases like $1/\epsilon^2$. Thus, even though the likelihood ratio becomes large, it is simply not large enough to amplify the prior probability ratio beyond one. Indeed, in agreement with equation (3) above, we find the posterior ratio to be of order ϵ and, for sufficiently small values of ϵ, the odds are therefore strongly against the particle being stationary.

While the above argument puts our conclusions on a solid Bayesian foundation it has not yet answered our basic question of why, when we observe an object to be stationary in the image, can we expect to be able to reliably conclude that the object is actually stationary in the world (still assuming a stationary observer). From equation (5) we see there are only two places to look for this answer, the likelihood ratio and the prior probability ratio.

Considering the likelihood ratio first, we might attempt it raise it by considering a more informative feature. For example, suppose we have accurate *stereo* measurements of the velocity of the particle, so that we obtain accurate constraints on all three components of it's motion. Such a system would increase the likelihood ratio to be of order $1/\epsilon^3$, if the measurement accuracy was order ϵ in each component. However, note that the resulting feature is still not necessarily a reliable indicator that the particle is at rest, since we can only conclude from equation (5) that the posterior probability ratio, R_{post}, would remain bounded away from zero as $\epsilon \rightarrow 0$. In order to get a reliable inference, the measurement accuracy would have to be significantly finer than the tolerance ϵ in our definition of stationarity. This approach of refining the measurements to raise the likelihood ratio does not fit our intuition very well, in which it seems a mere glance at a ball on our desk suffices to assure us that it is stationary. Thus we turn to the second place to look for our answer, namely the prior probability ratio.

2.3 A mode for stationary

Recall that the prior probability ratio represents a fact about the world, such as Maxwell's distribution for the velocities of gas particles, and *not* a perceiver's model of it. Therefore, in asking how the priors may help us in reliably inferring that a ball on our desk is at rest, we are asking about the structure of a suitable prior probability distribution for the velocities of balls. In fact, we will show that only a very simple qualitative property of such prior distributions is needed, not a detailed *quantitative* specification.

One obvious property of a prior distribution for velocities of balls is that, unlike gas particles, friction and gravity together provide strong constraints on the ball's motion. In particular, when in contact with a stationary supporting surface, the ball is often at rest (to within some tolerance ϵ). Thus, the prior distribution is more appropriately modeled using the "mixture" distribution

$$p(\vec{v}|B) = \pi^0 \delta(\vec{v}) + \pi^1 p^1(\vec{v}). \tag{9}$$

Here $\delta(\vec{v})$ is the Dirac delta function which represents a probability distribution concentrated at the point $\vec{v} = \vec{0}$, and $p^1(\vec{v})$ is some smooth probability distribution over the 3-space representing $\vec{v}$. These two 'component distributions' are combined in equation (9) using the 'mixture proportions' π^0 and π^1 to form the overall distribution, $p(\vec{v}|B)$, for such a ball context, B. Of course, to maintain a valid probability distribution we require that π^0 and π^1 are nonnegative and sum to one. Notice that the distribution $p(\vec{v}|B)$ models the property that objects are at rest with probability π^0, which is assumed to be nonzero in context B. This qualitative model for the prior distribution provides an alternative context in which to consider the inference of an object being at rest.

In this new context, the prior probability ratio R_{prior} is easily seen to be

$$\frac{p(P|B)}{p(not P|B)} = \frac{p(||\vec{v}_{world}|| < \epsilon)}{p(||\vec{v}_{world}|| > \epsilon)} = \frac{p(||\vec{v}_{world}|| < \epsilon)}{1 - p(||\vec{v}_{world}|| < \epsilon)} = \pi^0/(1 - \pi^0) + O(\epsilon^3), \tag{10}$$

as $\epsilon \to 0$. Note that, to leading order, all that matters here is the fact that being stationary occurs with the positive probability π^0; none of the details of the smooth distribution for velocities contribute to R_{prior} to first order. Also, in contrast to Maxwell's distribution, we now have a significant prior probability ratio, $\pi^0/(1 - \pi^0)$, bounded away from zero.

Next we need to consider the likelihood ratio. A similar argument to the one given above for the derivation of the likelihood ratio for context C shows that

$$R_{likelihood} = \frac{p(F|P\&B)}{p(F|not P\&B)} = O(1/\epsilon^2). \tag{11}$$

Together these two equations give us a posterior probability ratio, $R_{post} = O(1/\epsilon^2)$, which is much larger than one for ϵ sufficiently small. Therefore, in the context B we can reliably infer that the object is at rest, given that it was observed to be at rest in an image (assuming the measurement accuracy ϵ is sufficiently small); a mere glance should suffice.

The difference between contexts C and B is simply that in the ball context the property of being at rest is what we call a "modal property". That is, the property $\vec{v} = 0$ has nonzero probability in the prior distribution (See Jepson and Richards, 1993, for a formal definition of a "mode".). As we saw above, if a property is modal then the prior probability ratio R_{prior} remains bounded away from zero. Then, given an image feature for which the likelihood ratio $R_{likelihood}$, is known to be large, we might safely conclude that the modal property actually occurs in this instance. It is important to note that this works given fairly weak constraints on the prior distribution; we do not need to have a quantitative model for the prior distribution for the velocity of a ball (nor, in our opinion, can we expect to). Rather we need only assume that:

- The prior distribution is a mixture of two components or 'modes'.
- One component of the mixture, accounting for $1 - \pi^0$ of the probability, is a smooth function of $\vec{v}_{world}$.
- The other component, accounting for probability $\pi^0 > 0$, appears as a delta function at $\vec{v}_{world} = \vec{0}$.

As we saw above, such prior knowledge is needed to license even apparently innocuous inferences such as "objects which are stationary in the image are stationary in the scene".

3 Observability of modes: Key features

The general Bayesian argument presented in the previous section is not limited to inferences about whether or not an object is at rest. A similar argument shows the importance of prior knowledge in other apparently innocuous inferences, involving so called non-accidental properties (Binford, 1981; Lowe, 1985). For example, it can be shown that a nonzero prior for two sticks in the world to form a 'V'-configuration is critical for the reliability of the intuitively plausible inference that a V-configuration observed in an image corresponds to a V-configuration in the world (Jepson & Richards, 1993). Similarly, two colinear line segments in an image can be treated as reliable indicators that the corresponding 3D line segments are colinear in the world *only* when there is a modal prior probability for them to be colinear in the world. The interested reader is referred to Jepson and Richards (1993) for details. Other examples are listed in Table 1, along with the relevant references.

There are also many examples of modal structures in the world for which a single view may not be enough to obtain a reliable inference. Consider a context consisting of a set of

Property	Psychophysical Evidence	A Posteriori Analysis
Straightness	Watt & Andrews, 1982	1
Cotermination	Julesz, 1971; Nakayama & Shimojo, 1992	3,4,6
Circular	Foster & Wagemans, 1993; Lowe, 1985	1,3
Rectangular	Gregory, 1970	7
Parallel	Rock, 1983; Stevens, 1988	5,7
Rigid Body Motion	Ullman, 1979; Wallach & O'Connell, 1953	1,8
Collinearity	Westheimer & McKee, 1977	2,5
Skew Symmetry	Kanade, 1981; Leyton, 1992	$*$

Table 1: Some familiar "non-accidental" relations often suggested as driving visual perceptual inferences (for acoustic analogs, see Bregman, 1990). The numbers in the third column refer respectively to: (1) Bennett, Hoffman & Prakash, 1989; (2) Albert & Hoffman, 1991; (3) Knill & Kersten, 1991; (4) Jepson & Richards, 1992; (5) Jepson & Richards, 1991; (6) Nakayama & Shimojo, 1992; (7) Richards, Jepson & Feldman, 1994; (8) Reuman & Hoffman, 1986.

elliptical rings, of various eccentricities, scattered on a planar surface. Suppose there is also a mode for perfectly circular rings. Finally, suppose the slant and tilt of the plane with respect to the viewer is randomly chosen using a smooth distribution. Then, given an orthographic image consisting of just one ellipse, can the perceiver reliably determine the mode, namely elliptical or circular, of the ring that generated this image feature? Clearly the answer is no, since the observation of an ellipse in the image is typical for both elliptical and circular rings. Another example of a modal property which cannot be reliably identified is provided in Section 5.5 below, where we exhibit a context in which (amoungst other things) we cannot tell if a ball is sliding along the floor or moving through the air. It is important to emphasize that modal properties may be important to the perceiver, even though some may not be reliably observable from a single image.

Nevertheless, as we see from Table 1, there are a variety of applications for the Bayesian analysis presented in Section 2.2, and therefore it is useful to summarize the essential properties in a definition. In particular, we assume we have a context C which specifies the prior probability distribution for world events in the particular situation being studied. Suppose we are interested in whether or not property P holds in the world, given an image feature F with resolution parameter ϵ. Then we have the following definition of a "key feature":

Key Feature Definition: The feature F, with resolution parameter ϵ, is said to be a key feature for property P in the context C if:

$$\begin{aligned} R_{likelihood} &= \frac{p(F|P\&C)}{p(F|notP\&C)} \text{ is unbounded as } \epsilon \to 0, \\ R_{prior} &= \frac{p(P|C)}{p(notP|C)} \text{ remains bounded away from zero as } \epsilon \to 0. \end{aligned} \tag{12}$$

In this case, the posterior probability ratio R_{post} is unbounded as $\epsilon \to 0$, and thus, for sufficiently small ϵ, the image feature F provides a reliable indicator for world property P in context C.

The term "key feature" was introduced by Jepson and Richards (1993), with the intent that 'key' referred to unlocking reliable inferences about the world. The essential properties of key features, as listed in the above definition, have also been noted or, in fact, anticipated by a number of other authors. For example, Bennett, Hoffman and Prakash (1989) introduced the notion of an ideal observer, which has the same essential ingredients of a high likelihood ratio and a nonvanishing prior probability ratio. Similarly, Knill and Kersten (1991) discuss a different notion of an ideal observer which rests on these same conditions. For example, Knill and Kersten (1991) discuss the inference of the 3D shape of a wire given just one orthographic image. The critical element is the perceiver's prior knowledge of modal structure about the way in which the wires can be bent in the context they treat. Also, in more natural contexts, there is a key feature for the chromaticity of the illuminant (Lee, 1986), and a key feature for rigid 3D motion given perspective projection (Bennett, Hoffman & Prakash, 1989; Jepson & Richards, 1993).

The notion of a key feature is clearly an idealization of both the properties of the world and the sensing process. As we discussed above we need to assume that physical events occur at a variety of scales and, in order to obtain the required sorts of prior distributions involving delta functions, this separation of scales must be assumed to be extreme. This idealization frees us from talking about detection rates, acceptable false target probabilities, and so on. In fact, we view one of the most important contributions of the idea of a key feature to be that, in freeing us from such details, we are left to consider appropriate representations for world structure in a much simpler setting. In the next two sections we illustrate this by describing a suitable representation for a simple domain, along with the use of this representation in determining the reliability of various inferences.

4 Modal analysis

We began this paper by saying that the world we live in is a very structured place, and that it is knowledge of this structure which allows us to make reliable inferences from our various sensory signals (Bobick, 1987; Richards & Bobick, 1988). So far we have provided a glimpse into the details of how this might come about. We saw in Section 2 how the essential ingredients of a key feature, namely a large likelihood ratio and a nonzero prior, were both important for obtaining a reliable inference. Moreover we saw how regularities in our world can give rise to these essential ingredients. For example, we showed that the inference that an object is stationary rests on the fact that our world is structured so that objects often *are* stationary. This regularity, trivial as it sounds, is reflected as a mode in the prior probability distribution for the motion of objects in our world. The existence of this mode was shown to be critical in order to obtain a reliable inference that an object being observed is in fact at rest.

In this section we attempt to broaden the picture we have painted so far. In particular, we wish to show how a property space may consist of an embedded set of modes, and how these modes may interact with those occurring in a different property space. The particular example we consider is an idealization of a ball inside a box. The ball can move around inside the box, bounce off of the walls, roll along the floor or simply rest on the floor. Our goal is to illustrate a modal analysis of this context, and to briefly examine how reliable inferences may arise given perceptual data in this more complex domain. In order to keep the analysis simple we keep a strict idealization of the domain, which allows the essential structure of several natural inferences to be clearly exposed.

4.1 Motion modes: A ball in a box

The position of the ball's center of mass is denoted by $\vec{x}(t)$, which is taken to form a continuous trajectory. The velocity of the ball, $\vec{v}(t)$, is taken to be piecewise continuous in order to model collisions with the box. For the purpose of this section the critical observation is that the velocity of the ball appears naturally at several scales. In particular, we distinguish small velocities of the ball due to vibrations of the box, air currents, and other small perturbations, from the range of velocities achieved by throwing it, dropping it, or hitting it with a squash racket. We idealize the first distribution as the ball being at rest, that is the single point $\vec{v} = \vec{0}$, while the latter corresponds to a range of velocities which occur with the ball in the air. In addition, there is the set of velocities which occur while the ball is rolling on the floor. Here again there is a separation of scales, with the component of the velocity in the direction perpendicular to the floor taken to be negligible.

This notion of a separation of scales between the velocities which occur due to different physical processes does not mean an exclusive separation. In particular, we allow that the ball can momentarily be moving slowly while it is in the air even though this speed is more typical of the ball being on the floor. For example, if the ball bounces nearly vertically, then it will be nearly stationary at the top of it's trajectory. Our point about the separation of scales is that the distribution of velocities of the ball while it is in the air has a broad range and it is the *exceptional situation* which can produce a small velocity. Our idealization is then to take this separation of scales to be extreme. In particular, we take the small motions due to vibrations, etc. to be negligible, and treat the "at rest" state as the single point $\vec{v} = \vec{0}$.

The modes in the prior probability distribution on a configuration space correspond to the effects of different physical processes, operating at different scales and existing primarily on different sets within this configuration space. For example, consider the configuration space consisting only of the set of velocities, $\vec{v}(t)$, of the ball. Then, as depicted in Figure 1a we have a mode for being at rest (i.e. at the point $\vec{v} = \vec{0}$, another for rolling or sliding on the floor (i.e. in the plane $v_3 = 0$), and a third mode for velocities which occur during free fall. This third mode is taken to be a smooth distribution over the three dimensional configuration space $\vec{v}$. The critical point is that the processes operating at different scales have resulted in modes that exist on sets which have different scales in particular directions. We have idealized this separation of scales to the extreme of changing the *dimension* of the various subsets corresponding to modes in the configuration space. For example, smooth free fall generates velocities smoothly distributed over the 3D configuration space, rolling motion

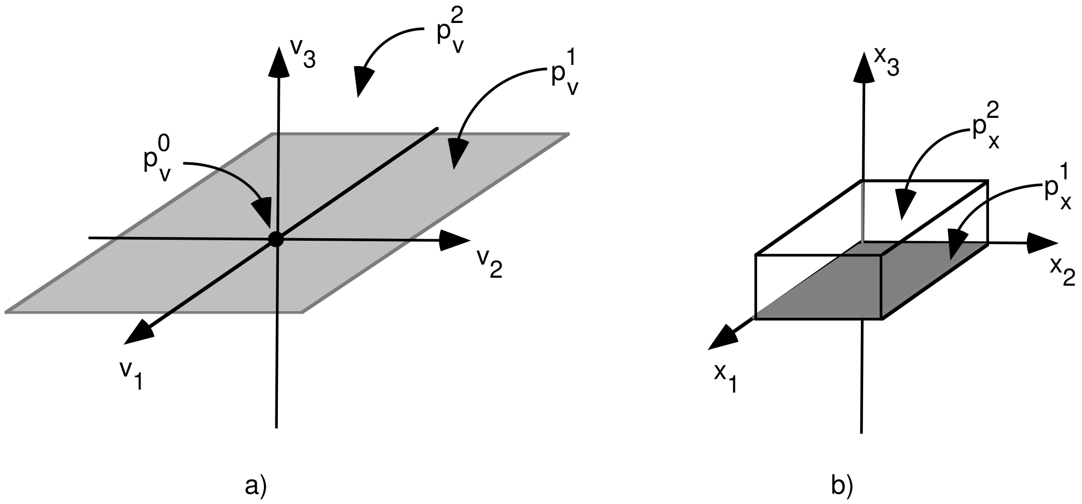

Figure 1: Modes for the velocity (a) and the position (b) of the ball in the box.

only exists in the 2D subset consisting of $v_3 = 0$, while being at rest occurs only at a single point, namely the origin. Our prior model then is to put smooth probability distributions on each of these sets.

To be concrete, we take the prior distribution for velocities $\vec{v}(t)$ to have the form of the mixture model

$$p_v(\vec{v}) = \pi_v^0 p_v^0(\vec{v}) + \pi_v^1 p_v^1(\vec{v}) + \pi_v^2 p_v^2(\vec{v}). \tag{13}$$

Here p_v^n denotes the *component distribution* associated with the n^{th} mode or process. Each of these component distributions is combined with it's *mixing proportion* π_v^n, and the sum provides the desired prior distribution. The sum of the π_v^n over the n processes should be 1. For processes which exist on sets having a dimension smaller than the full configuration space, such as rolling, the prior distribution is taken to be a smoothly modulated delta function along this set. In particular, for our current example, we have component distributions of the following form

$$\begin{array}{l} p_v^0(\vec{v}) = \delta(\vec{v}), \\ p_v^1(\vec{v}) = \delta(v_3) q_v^1(v_1, v_2), \\ p_v^2(\vec{v}) \text{ is a bounded function of } \vec{v}. \end{array} \tag{14}$$

Note that p_v^0 represents the "at rest" mode, p_v^1 represents rolling or sliding on the floor with $q_v^1(v_1, v_2)$ some bounded distribution over the plane of horizontal velocities $(v_1, v_2, 0)$. The final mode, for particles on a ballistic trajectory for example, is accounted for by the last component, $p_v^2(\vec{v})$. We assume that both q_v^1 and p_v^2 are bounded functions in order to rule out the possibility of additional modal structure not already represented by the delta functions in (14).

Given our analysis in the preceding sections, one might expect that these modes play a critical role in making inferences such as whether the ball is at rest or moving at time t. Before considering a Bayesian analysis of such an inference it is useful to first examine a separate configuration space for the ball-in-a-box example which represents spatial properties of the system.

4.2 Spatial modes: A ball in a box

The position of the ball's center of mass, $\vec{x}(t)$, provides another example of a configuration space for which the prior distribution has a modal structure. Here, we have one mode consisting of a smooth distribution over the inside of the box which corresponds to positions of the ball during free fall. Also, there is a mode on the floor, say $x_3 = 0$, which corresponds to situations in which the ball is rolling or simply at rest on the floor (see Figure 1b). Finally, we assume that the collisions are effectively instantaneous[2] so there are no additional modes arranged around the walls and ceiling of the box.

Following our previous example, such a prior distribution can be conveniently written as the mixture model

$$p_x(\vec{x}) = \pi_x^1 p_x^1(\vec{x}) + \pi_x^2 p_x^2(\vec{x}). \tag{15}$$

Again the quantities π_x^n provide the mixing proportions for the two modes. The component distributions p_x^n are given by

$$\begin{array}{l} p_x^1(\vec{x}) = \delta(x_3) q_x^1(x_1, x_2). \\ p_x^2(\vec{x}) \;\text{ is a bounded function of } \vec{x}. \end{array} \tag{16}$$

Note that p_x^1 represents the both the "at rest" mode and the "rolling" mode, with $q_x^1(x_1, x_2)$ representing some bounded distribution over the plane of horizontal positions $(x_1, x_2, 0)$. Since there is no special position on the floor for the ball to be at rest we do not have a pure delta function in the spatial domain. Also, no point on the floor should be forbidden, so we take $q_x^1(x_1, x_2)$ to be a smooth nonzero distribution over the floor. The second mode, p_x^2, accounts for the positions achieved during it's various possible trajectories, and we take it to be a smooth (nonzero) distribution over the interior of the box.

4.3 Mode coupling

Given the spatial and motion modes described in the previous sections we need to consider how they can be combined to derive the prior probability for the ball to be at position $\vec{x}(t)$ with velocity $\vec{v}(t)$ at time t. Our basic point is a simple one, namely that the appropriate joint distribution is *not* just what one obtains by treating the priors for position and velocity independently. Rather, the mixture distributions for position and velocity must be coupled in a nontrivial way (see also Pearl, 1988; Yuille et al., 1994).

To see this, consider the distribution $p_{indep}(\vec{x}, \vec{v})$ obtained by treating the position $\vec{x}$ and

[2] This is yet another form of the argument that processes occur at different scales. Here we are taking the time duration of a collision as negligible compared to the temporal resolution dt of the system.

velocity $\vec{v}$ as independent, that is

$$p_{indep}(\vec{x}, \vec{v}) = p_x(\vec{x}) p_v(\vec{v}). \tag{17}$$

Our claim is that this distribution does not provide the correct modal properties. In particular, consider the conditional prior probability density that the ball is at rest, given that it is at position $\vec{x}$, for some $\vec{x}$ in the interior of the box. Since there is no physical process generating a mode for the ball to be at rest, we should expect the prior to similar to the one obtained in Section 2 for the gas particle (see equation (8)). That is, we expect the prior to be roughly

$$p(||\vec{v}|| < \epsilon \mid x_3 > 0) = p_v^2(\vec{0})\epsilon^3 + O(\epsilon^4),$$

with $p_v^2(\vec{v})$ the prior distribution for velocities while the ball is in the air. The actual coefficient $p_v^2(\vec{0})$ of the leading order term does not matter here; what is significant is that this probability is of order ϵ^3. In contrast, it follows from equation (17) that

$$p_{indep}(||\vec{v}|| < \epsilon \mid x_3 > 0) = \pi_v^0 + O(\epsilon^2),$$

which is just the mixing proportion, π_v^0, of the rest state to leading order.

The reason for this difference is that the independent combination of the spatial and motion distributions, namely p_{indep}, does not take into consideration the appropriate coupling between the modes in the different configuration spaces. In particular, notice that when the product in (17) is written out we obtain six different modes in the resulting mixture model. Some of these modes do not make sense physically. For example, when the ball is above the floor there should be no mode for being at rest (i.e. the mixture proportion π_v^0 should be 0), and no mode for horizontal motion (i.e. $\pi_v^1 = 0$). Also, when the ball is on the floor, there should not be a mode for a general 3D velocity (i.e. $\pi_v^2 = 0$). Together these give the mode coupling conditions

$$\begin{array}{l} \pi_v^0 = \pi_v^1 = 0 \text{ whenever } x_3 > 0, \\ \pi_v^2 = 0 \text{ whenever } x_3 = 0. \end{array} \tag{18}$$

These coupling conditions can also be incorporated into a mixture model for the joint distribution of $\vec{x}$ and $\vec{v}$, in which only the physically plausible modes are kept. In particular, an appropriate model is

$$p_{x,v}(\vec{x}, \vec{v}) = \pi_{x,v}^0 p_{x,v}^0(\vec{x}, \vec{v}) + \pi_{x,v}^1 p_{x,v}^1(\vec{x}, \vec{v}) + \pi_{x,v}^2 p_{x,v}^2(\vec{x}, \vec{v}). \tag{19}$$

Here the three modes, or component processes, are

$$\begin{array}{l} p_{x,v}^0(\vec{x}, \vec{v}) = \delta(\vec{v})\delta(x_3) q_x^1(x_1, x_2), \\ p_{x,v}^1(\vec{x}, \vec{v}) = \delta(v_3)\delta(x_3) q_{x,v}^1(x_1, x_2, v_1, v_2), \\ p_{x,v}^2(\vec{x}, \vec{v}) \text{ is a bounded distribution.} \end{array} \tag{20}$$

Here q_x^1 is the same as introduced in equation (16). Also, we constrain the new distributions $q_{x,v}^1$ and $p_{x,v}^2$ to be bounded functions, which eliminates the possibility of further modal structure within this model. Note that the three mixture components in equation (19) correspond to the processes at rest on the floor, rolling or sliding on the floor, and moving through the air.

4.4 Reliable inferences

Suppose the image feature F_{floor} includes the fact that the image velocity of the ball satisfies $||\vec{v}_{image}(t)|| < \epsilon$, along with some positional information about the ball and the box. In particular we assume that the image feature F_{floor} also shows that the position of the ball in the image is consistent with the ball being on the floor. Should we infer from the observation that the ball is stationary? Let property P denote that the ball is stationary in the world (i.e. $||\vec{v}|| < \epsilon$), given our "ball-in-a-box" context, B. First consider the likelihood ratio $R_{likelihood} = p(F|P\&B)/p(F|notP\&B)$. Then $R_{likelihood}$ is of order $1/\epsilon^2$, as follows: The numerator of the likelihood ratio is nonzero, since image motion is typically small when the ball is at rest (property P). On the other hand, in the absence of property P, the image motion will be small only if both coordinates of the image velocity also happen to be small. However, because the prior distribution for $\vec{v}$ requires that these components be smoothly distributed, $||\vec{v}_{image}||$ will be less than ϵ with a probability proportional to ϵ^2. The likelihood ratio is therefore of order $1/\epsilon^2$ as $\epsilon \rightarrow 0$, and we see that the first condition of the key feature definition is satisfied, given our observation of no image motion.

Next consider the prior probability ratio $R_{prior} = p(P|B)/p(notP|B)$. The mode coupling condition (18) allows a mode in which the ball is stationary so long as it is on the floor. Moreover, there is also a spatial mode for the ball to be on the floor. By equation (19), the prior $p(P|B)$ is just $\pi^0_{x,v}$ to leading order in ϵ, and therefore the ratio R_{prior} is given by $\pi^0_{x,v}/(1-\pi^0_{x,v})$ to leading order. Therefore the second condition of our key feature definition is satisfied, and we can reliably conclude that the ball is at rest. Notice that we can also conclude that the ball must be *on the floor*. In fact, it can be shown that this state in which the ball is on the floor and at rest has the a posteriori probability of $1 - O(\epsilon^2)$, and is therefore a reliable inference for sufficiently small ϵ.

Now consider another image feature F_{air} for the ball-in-a-box where again $||\vec{v}_{image}(t)|| < \epsilon$, but $x_3 > 0$, and hence the ball cannot be on the floor. Then, from the mode coupling condition (18), it follows that there is no mode for the ball to be at rest. Indeed, the prior probability ratio R_{prior} for it to be at rest is only of order ϵ^3. Hence the feature F_{air} does not satisfy the second key feature condition, namely that R_{prior} remain bounded away from zero, as $\epsilon \rightarrow 0$. Given the likelihood ratio $R_{likelihood}$ is of order $1/\epsilon^2$, the posteriori ratio R_{post} for property P being correct is only $0(\epsilon)$. Hence the odds strongly favor the conclusion that the ball must be in motion, even though that motion must be special in that it is nearly directed along the line of sight (see Figure 6.4 in Jepson and Richards, 1992, for a different perceptual example of the possible use of such mode coupling conditions).

5 Qualitative probabilities

An important property of the modal analysis presented in the previous section is that the conclusions, namely the reliability of the inference that the ball is at rest, do not depend on quantitative details of the various smooth component distributions in the prior probability

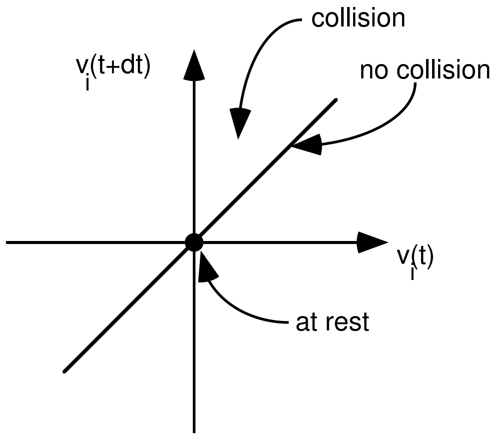

Figure 2: A 2D slice of velocity configuration space with modes for at rest, smooth motion, and discontinuous motion.

model (19). In fact, all that matters in the limit as $\epsilon \to 0$ is that the mixture proportion for the mode "at rest on the floor" is nonzero (i.e. $\pi^0_{x,v} > 0$), along with various nondegeneracy assumptions which ensure the density functions q^1_v, p^2_v, etc. are smooth and bounded. A perceiver can therefore gain some freedom from knowing precise quantitative details about the appropriate prior distributions, yet still expect to make reliable inferences, at least for contexts that support key features (see also Doyle & Sacks, 1989). Here we illustrate this point by considering inferences about the occurrence of a collision, again using the "ball-in-a-box" context.

5.1 Velocity discontinuities

Recall that the position of the ball's center of mass, $\vec{x}(t)$, is taken to form a continuous trajectory but that the velocity of the ball, $\vec{v}(t)$, is taken to be piecewise continuous. Behind this assumption there is again a separation of scales of various physical processes, just as we had for the distinction of moving versus at rest in the previous section. Here we note that forces on the ball arise naturally at several scales. At one scale we have gravity, aerodynamic effects, and so on. It is assumed that the forces due to effects at this scale are significantly smaller than the typical forces due to collisions. We emphasize again that this separation is not absolute since collisions can, of course, produce small forces, but it is an exceptional collision which produces a force as small as gravity.

The analysis is simplest when we take this separation of scales to be extreme. In particular, over a time interval of duration dt, which is just resolvable by the system, we take the change of velocity due to the gravity and aerodynamic effects to be negligible while collisions can produce significant changes in the velocity. At the resolution determined by dt then, the velocity of the ball may be discontinuous due to collisions with the walls. Otherwise the motion of the ball is well approximated by a constant velocity over each interval $(t, t + dt)$.

Consider then the configuration space made up of the 6D-points, $(\vec{v}(t), \vec{v}(t+dt))$. As discussed above, the modes in the prior distribution on this space are meant to correspond to the effects of different physical processes, operating at different scales and existing primarily on different sets within this space. For example, smooth trajectories appear on the 3D surface $\vec{v}(t) = \vec{v}(t+dt)$, where we have used the idealization that the smaller forces produce negligible velocity changes over time intervals of length dt. In Figure 2 we depict a 2D slice of the velocity configuration space, and this smooth motion mode is represented simply by the line $v_1(t) = v_1(t+dt)$. The modes we studied in the previous section, namely for the ball to be at rest and for the ball to be rolling or sliding on the floor will also appear in this 6D configuration space.

The large forces due to collisions can generate large changes in the velocity during a time step of dt. We take these collision events as producing a smooth distribution in the configuration space $(\vec{v}(t), \vec{v}(t+dt))$. In fact, there is additional structure within the set of velocities consistent with a collision occurring between time t and $t+dt$. For example, if the ball can only collide with the stationary box then conservation of energy (neglecting spin) requires that the speed $||\vec{v}(t+dt)||$ can be no larger than $||\vec{v}(t)||$. In addition, there are further restrictions on the direction of the outgoing trajectory due to the surface normal and spin on the ball. In order to keep our analysis simple we ignore these further regularities and just require that the component of the prior distribution due to collisions is a bounded function of $(\vec{v}(t), \vec{v}(t+dt))$.

5.2 Collisions

In order to make inferences about collisions we need to consider the prior distribution for the ball to be at position $\vec{x}(t+dt)$ with velocity $\vec{v}(t+dt)$ at time $t+dt$, given that it was at $\vec{x}(t)$ with velocity $\vec{v}(t)$ at time t. Again it should be clear that the desired distribution cannot be represented with an independent combination of a distribution for the position of the ball with a second distribution for the velocity of the ball at $t+dt$. Instead, we need to take into account further structure about the context, such as the fact that velocity discontinuities only occur when the ball collides with the box.

The desired prior can be broken up into two cases, depending on whether or not a collision occurs during the time interval $(t, t+dt)$. Since we are taking the time duration dt to be short enough such that the velocity of the ball is essentially constant, the ball's trajectory can be initially predicted by the linear model

$$\vec{x}_p(\tau) = \vec{x}(t) + (\tau - t)\vec{v}(t), \tag{21}$$

for $\tau \in (t, t+dt)$. This prediction is taken to be accurate so long as this trajectory does not penetrate the walls of the box. Here t, $\vec{x}(t)$ and $\vec{v}(t)$ are treated as constants, and they together completely specify the trajectory $\vec{x}_p(\tau)$. We say that the trajectory $\vec{x}_p(\tau)$ *transversally* intersects the box at a point $\vec{x}_c$ if the path intersects the box at that point and is not tangent to the box. Thus rolling or sliding motions are non-transversal, while a typical collision is a tranversal intersection. In the cases where no collision occurs during $[t, t+dt)$

we take the prior model to be

$$p(\vec{x}(t+dt), \vec{v}(t+dt) \,|\, \vec{x}(t), \vec{v}(t), noCollision) = \delta(\vec{x}(t+dt) - \vec{x}_p(t+dt))\delta(\vec{v}(t+dt) - \vec{v}(t)). \quad (22)$$

Here the conditional term $noCollision$ refers to the fact that the predicted trajectory $\vec{x}_p(\tau)$ does not transversally intersect the box for $\tau \in [t, t+dt)$. The two delta functions in the above equation impose the constraint that the predicted trajectory $\vec{x}_p(t+dt)$ is accurate in the absence of collisions. In practice one would need to replace these delta functions with distributions which have nonzero variances. But for simplicity here we are assuming the scale of the errors in the prediction is negligible over a time step of length dt.

In the second case, for which a collision occurs, the predicted trajectory $\vec{x}_p(\tau)$ is piecewise linear. The discontinuity occurs at the point of transversal intersection, namely $(\vec{x}_c, t_c)$, between the line described in (21) and the box. For simplicity we neglect the special cases in which several discontinuities may appear in the time interval $(t, t+dt)$. The predicted trajectory is thus

$$\vec{x}_p(\tau) = \begin{cases} \vec{x}(t) + (\tau - t)\vec{v}(t), & \text{for } \tau \in [t, t_c]; \\ \vec{x}_c + (\tau - t_c)\vec{v}(t+dt), & \text{for } \tau \in [t_c, t+dt). \end{cases} \quad (23)$$

Here we have approximated the impact to be instantaneous, and taken the reflected velocity, $\vec{v}(t+dt)$ to be constant after the impact. Clearly, the assumption here is that the reflected velocity points back into the interior of the box or, at least, along the wall from the contact position $\vec{x}_c$. Given this predicted trajectory $\vec{x}_p(\tau)$, the prior distribution given a collision is taken to be

$$p(\vec{x}(t+dt), \vec{v}(t+dt) \,|\, \vec{x}(t), \vec{v}(t), \; Collision \;) = \delta(\vec{x}(t+dt) - \vec{x}_p(t+dt)) p_{refl}(\vec{v}(t+dt) | \vec{v}(t), \vec{x}_c). \quad (24)$$

Here p_{refl} provides the distribution for the reflected velocity, $\vec{v}(t+dt)$, given the incoming velocity $\vec{v}(t)$ and the point of collision $\vec{x}_c$. In order to model relatively unstructured scattering of the ball off of the wall, due to spins or imperfections in the wall, we take p_{refl} to be a smooth bounded function of $\vec{v}(t+dt)$, subject to the constraint that $\vec{v}(t+dt)$ cannot point out of the box from $\vec{x}_c$.

It is also possible to model more specialized structure in the process of the ball bouncing off of the walls. For example, the distribution p_{refl} above could be augmented with a mixture of different modes of reflection. In addition, to the general scattering used in (24), we could include a second mode in which the reflected velocity is in the direction of the perfect reflection. A third mode could be added in which the motion immediately after the collision is directed along the planar surface at $\vec{x}_c$, as in a rolling or sliding motion. Thus our basic framework in terms of mixture models is sufficiently expressive to capture a wide range of phenomena. However, for our purposes here, it is convenient to keep the model simple and ignore these more detailed structures.

5.3 Non-degeneracy conditions

In order to complete the specification of our ball-in-a-box context B, we need to impose some non-degeneracy conditions on the various prior components which we have introduced so far. For example, consider the component distribution $p^2_{x,v}(\vec{x},\vec{v})$, introduced in (19), which is just the prior probability distribution for the ball to be at location $\vec{x}(t)$ with velocity $\vec{v}(t)$ during it's free motion through the air. Up to this point we have only required that $p^2_{x,v}$ is a bounded function of $\vec{x}$ and $\vec{v}$. But clearly this is too broad a class of distributions, since entirely unintended behaviours can arise by choosing distributions with particular structures. For example, $p^2_{x,v}$ might be zero within a wide layer all around the inside of the box, in which case the prior probability of a collision would also vanish. To avoid such bogus properties we need to bound the various component distributions from below. Just such a bound was used in Section 4.4, where we discussed the inference that the ball was resting on the floor of the box. There we needed to assume that the density q^1_x was bounded away from zero over the floor of the box. The main point of this section is that such a bound can be conveniently expressed in terms of a single *canonical* distribution.

First consider the component distribution $p^2_{x,v}(\vec{x},\vec{v})$. We take the canonical model, $m^2_{x,v}(\vec{x},\vec{v})$ to be a uniform distribution over the inside of the box for all velocities having speed less than some constant v_{max}. That is,

$$m^2_{x,v}(\vec{x},\vec{v}) = \begin{cases} K & \text{for } \vec{x} \text{ in the box and } ||\vec{v}|| < v_{max}; \\ 0 & \text{otherwise.} \end{cases} \tag{25}$$

Here K is a positive normalization constant, which depends on the volume of the box and v_{max}, such that the model distribution $m^2_{x,v}$ integrates to one. Our nondegeneracy condition on $p^2_{x,v}(\vec{x},\vec{v})$ is simply that there exists a positive constant c such that

$$p^2_{x,v}(\vec{x},\vec{v}) \geq c m^2_{x,v}(\vec{x},\vec{v}) \tag{26}$$

for all values of $\vec{x}$ and $\vec{v}$. In words, our component density $p^2_{x,v}(\vec{x},\vec{v})$ must be bounded below by some constant multiple of our canonical model $m^2_{x,v}(\vec{x},\vec{v})$. Such a condition, ensures the prior density $p^2_{x,v}$ is nonzero at any point $(\vec{x},\vec{v})$, with $\vec{x}$ in the box and $\vec{v}$ having a speed of at most v_{max}. As we show in the next subsection this condition is sufficient to ensure that collisions between the ball and the box do occur with a positive prior probability. Notice that equation (26) is a rather loose constraint on the component $p^2_{x,v}(\vec{x},\vec{v})$ in that many different distributions satisfy this constraint.

Similarly, to specify context B, we also need to impose a non-degeneracy condition on the conditional prior $p_{refl}(\vec{u}|\vec{v},\vec{x}_c)$ used in equation (24). In this case we take our canonical model m_{refl} to be

$$m_{refl}(\vec{u}|\vec{v},\vec{x}_c) = \begin{cases} K' & \text{for } \vec{n}(\vec{x}_c)\cdot\vec{u} \leq 0 \text{ and } ||\vec{u}|| \leq ||\vec{v}||; \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

Again K' is a positive normalization factor. Here $\vec{n}(\vec{x}_c)$ is an outward pointing normal vector for the side of the box which contains the point of collision $\vec{x}_c$. The above model states that the reflected velocity $\vec{u}$ is uniformly distributed in any direction, other than those that penetrate the box, with the speed limited to be no larger than the speed $||\vec{v}(t)||$ along the incoming trajectory. Our nondegeneracy condition for the distribution p_{refl} is then simply that there exists a positive constant c such that

$$p_{refl}(\vec{u}|\vec{v}, \vec{x}_c) \geq c m_{refl}(\vec{u}|\vec{v}, \vec{x}_c) \tag{28}$$

for all $\vec{u}$, $\vec{v}$, and $\vec{x}_c$.

Finally, non-degeneracy conditions are required for the remaining distributions, q^1_x and $q^1_{x,v}$ in equation (20). For $q^1_x(x_1, x_2)$ we take the canonical distribution, say m^1_x, to be a uniform distribution over the floor of the box. Recall that the distribution q^1_{xv} models the position and velocity of the ball while it is rolling on the floor. A canonical model distribution, say $m^1_{x,v}$ similar to the one used in equation (25) can be formulated, but with x_3 and v_3 constrained to be zero.

This completes the specification of the context B. In particular the prior distribution in this context B is given by equations (19), (22), and (24). Each of the mixture proportions $\pi^n_{x,v}$ in equation (19) are taken to be positive. Moreover, the various component distributions q^1_x, $q^1_{x,v}$, $p^2_{x,v}$, and p_{refl} must each satisfy a non-degeneracy condition of the form (26) and (28) using the corresponding canonical distribution m^1_x, etc. A large class of particular quantitative distributions satisfy these conditions, and it is in this sense that our probabilistic model is qualitative. The perceiver need not precisely specify the prior distributions, but rather only needs to describe it's modal structure. Our last task is to consider what sort of reliable inferences can be made given only that the prior distribution is in this class of distributions which are consistent with the canonical model, that is, within the context B.

5.4 Key feature for a collision

As Rubin has pointed out (1986), the observation of a velocity discontinuity in an image can be a reliable indicator of a velocity discontinuity in 3D. For our context B, a velocity discontinuity occurs only if the ball collides transversally with the box. In fact, as we show in this section, the observation of a discontinuity in the image velocity provides a key feature for such a collision.

It is useful to walk through the argument in terms of the Bayesian analysis we presented in Section 2.2 and the key feature definition provided in Section 3. Let the feature F denote the observation of image velocities $\vec{v}_{image}(t)$ and $\vec{v}_{image}(t + dt)$ such that the norm of the difference satisfies $||\vec{v}_{image}(t) - \vec{v}_{image}(t + dt)|| > dv$. Here the threshold $dv > 0$ should be taken smaller than the rough scale of velocities which occur while the ball is in the air. That is, given our canonical model, for which speeds at least up to the value v_{max} are known to occur, we should take $dv << v_{max}$. Next, let property P denote the occurrence of a collision of the ball with the box at some time in the interval $(t, t + dt)$. Similarly, let $notP$ denote the property that no collision occurs during the same time interval. Finally, suppose the (absolute) error in the measured image velocities is no larger than ϵ.

Consider the likelihood ratio, $p(F|P\&B)/p(F|notP\&B)$, as the measurement resolution ϵ goes to zero. The denominator is just the likelihood of observing the image feature F given that no collision (i.e. $notP$) occurs. In our context B, the change in the 3D velocity over a time interval of duration dt is negligible unless a collision occurs. Given orthographic projection, the same holds true for the exact image velocity. Thus, given $notP$, the observation of a significant change in the observed image velocity (i.e. the observation of F) can only be due to measurement error. Therefore, as the velocity measurement resolution, ϵ, goes to zero it must be the case that $p(F|notP\&B)$ also goes to zero. That is, a smooth 3D trajectory is increasingly unlikely to produce a sudden large jump in the image motion as the measurement resolution becomes finer.

Next consider the numerator in the likelihood ratio, $p(F|P\&B)$, which is just the likelihood of observing the image feature F given a collision occurs. In general this likelihood depends on the nature of the collisions. For example, if all the collisions occurred with the ball moving extremely slowly, then it is possible that the likelihood $p(F|P\&B)$ might vanish since no velocity step will be larger than dv, as is required to trigger the feature F. Therefore, to ensure a positive likelihood we must use the nondegeneracy conditions discussed in the previous section.

In order to check that $p(F|P\&B)$ is bounded away from zero, it is convenient to consider the collisions which *fail* to generate a suitable image feature F. For each choice of incoming velocity, $\vec{v}(t)$, the detection of a velocity discontinuity will fail precisely when $||\vec{v}_{image}(t) - \vec{v}_{image}(t+dt)|| \leq dv$. The set of all reflected velocities $\vec{v}(t+dt)$ which satisfy this inequality is a tube through the 3D velocity space, having radius dv, which is aligned with the viewing direction. Since dv is chosen to be much smaller that the maximum speed v_{max} in the canonical model, then for all sufficiently large incoming velocities $\vec{v}(t)$, there must be a large region outside the tube in which the canonical density $m^2_{refl}(\vec{v}(t+dt)|\vec{v}(t), \vec{x}_c)$ is strictly positive. The result then is that, given the canonical model, $p(F|P\&B)$ is strictly positive. And, by our nondegeneracy conditions, the same must hold for any suitable prior. As a result, we have shown that the likelihood ratio is unbounded as ϵ goes to zero.

In order for discontinuities in the image velocity to be a key feature for collisions we need show that the prior probability ratio satisfies the second condition in the key feature definition. That is, the prior probability ratio, $p(P|B)/p(notP|B)$, must be positive (or, more precisely, bounded away from zero as ϵ goes to zero). Here again we need to use the nondegeneracy conditions, this time to argue that a collision during a time interval of duration dt has a positive prior probability. To do this, consider the region close to one wall of the box, say within a distance dx of the wall. Then the canonical model $m^2_{x,v}(\vec{x}(t), \vec{v}(t))$ ensures a positive probability for the ball to be in such a region at time t. In order for a collision to occur before time $t+dt$ we also require that the velocity, $\vec{v}(t)$, satisfies $\vec{n}_{wall} \cdot \vec{v}(t) > dx/dt$, where $\vec{n}_{wall}$ is the outward pointing normal. But with dx chosen sufficiently small (i.e. $dx < v_{max}dt$), there is a positive probability that the ball will also have a velocity such that a transversal impact will occur before time $t+dt$. Thus collisions occur with a strictly positive probability under our canonical model, and therefore the same must be true for any prior consistent with this model. Thus we find $p(P|B)$ is bounded away from zero, which is sufficient to ensure the second requirement in the key feature definition. As a result, a

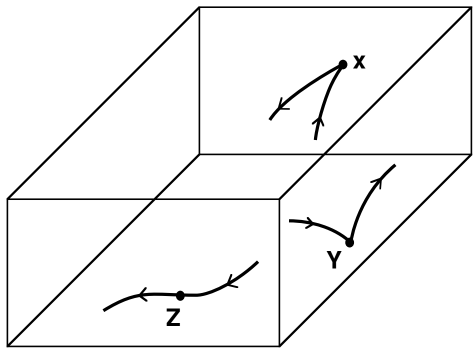

Figure 3: Collisions can be reliably inferred from the trajectories through X and Y, while at point Z the contact of the ball with the ground cannot be reliably inferred given the current prior model.

discontinuity in the image velocity is a key feature for a collision in 3D in our ball-in-a-box context B.

5.5 Reliable inferences

Given that we have a key feature, suppose you observe a velocity discontinuity in the trajectory near the point 'X' in Figure 3. Then the calculations in the previous subsection show that the posterior probability ratio, $p(P|F\&B)/p(notP|F\&B)$, is large and the odds overwhelmingly favour the conclusion that the ball hit the back wall (or a transparent front one) somewhere near the visual ray through the point X.

In fact, the priors introduced above also sanction the conclusion that, after the collision, the ball bounced away from the wall rather than sliding along it. To show this, consider the set of all collisions which are consistent with the images of the three points $\vec{x}(t)$, $\vec{x}_c$, and $\vec{x}(t+dt)$. The situation in which the ball slides along the wall is uniquely specified by these observations, and a single value of the reflected velocity $\vec{v}(t+dt)$ is determined. Alternatively, for bounces in which the ball comes away from the wall, the set of reflected velocities consistent with the observations form a line directed along the visual ray. Including some image noise changes these structures to a small neighbourhood and a tube, respectively. The basic structure is similar to our "at rest"/"moving along the line of sight" distinction we had before. In fact, since our canonical distribution for 3D reflected velocities (namely $m_{refl}(\vec{u}|\vec{v},\vec{x}_c)$) does not have a mode directed along the wall, it is much more probable that the observed reflected velocity is away from the wall (i.e. $\vec{v}(t+dt)$ lies somewhere in the tube, not just near one end).

As an alternative example, suppose you observe an abrupt stop at X. That is, there is a discontinuity in the image velocity at time t, and that image of the ball remains at rest immediately after this event. The best bet is then that the ball hit the wall and is now coming towards/going away from you. The analysis again provides a tube of possible motions consistent with the ball moving towards or away from the viewer (to within the resolution ϵ). The prior $m_{refl}(\vec{u}|\vec{v},\vec{x}_c)$ has no mode at the rest state, and as a result the posterior probability is strongly in favour of motion (nearly) along the line of sight.

Next, suppose you observe a velocity discontinuity in the trajectory near the point Y in Figure 3, so the image observations are consistent with the ball bouncing off of the floor. Actually, the same observations are consistent with the ball colliding with the floor and then beginning to roll, or rolling and then suddenly leaving the floor, or, finally, executing a velocity discontinuity as it rolls on the floor. Does our current qualitative probabilistic model have anything to say about the choice between these various possibilities? The latter two interpretations are eliminated by our model since velocity discontinuities do not occur unless the trajectory of the motion is transverse to a surface (i.e. the floor is smooth). We are therefore left with two interpretations, one in which the ball bounces off of the floor, and another in which the ball impacts the floor but then immediately begins rolling. This case is different than the one discussed previously for the collision occurring at the point X, since now there is a mode for rolling or sliding along the floor. Which interpretation, if either, is more probable?

The issue can be resolved using the model for the reflected velocity, $m_{refl}(\vec{u}|\vec{v},\vec{x}_c)$, just as for the collision at X treated above. In particular, the majority of collisions (i.e. $1-O(\epsilon)$ versus $O(\epsilon)$) predicted by this model will be away from the floor. Thus we can reliably infer that the ball is in the air shortly after the collision with the floor. (Note that we cannot apply the prior distribution $p_{x,v}$ to the analysis of the state $(\vec{x}(t+dt),\vec{v}(t+dt))$ at time $t+dt$, since the conditional distribution p_{refl} shows that this state is not independent of the initial state $(\vec{x}(t),\vec{v}(t))$).

Finally, suppose you observe the smooth trajectory through the image point Z at time t, and that $||\vec{v}_{image}(t+dt)-\vec{v}_{image}(t)||<\epsilon$ (call this F_{smooth}). The reader is left to check that the two most plausible interpretations are: 1) the ball is moving smoothly (no collision) through the air during $(t,t+dt)$; 2) the ball is rolling/sliding smoothly along the floor during $(t,t+dt)$. In particular, the probability that a collision will give rise to this image observation is $O(\epsilon^2)$. Therefore it is improbable that a collision has occurred (in fact F_{smooth} is a key feature for smooth motion during $(t,t+dt)$). We are therefore left with the above two possibilities. How should we decide between the two?

One way to decide is to seek the most probable category, that is, either the trajectory is in the air or it is on the ground. Our canonical distribution cannot uniquely designate one of these two categories as being necessarily more probable. The result depends on the mixture proportions for the prior probabilities of motion through the air and motion on the ground ($\pi^2_{x,v}$ and $\pi^1_{x,v}$, respectively), which have not been specified. The appropriate conclusion then is that without more information, we cannot decide on the most probable category.

A second way to decide is to seek the most probable *initial state* $(\vec{x}(t),\vec{v}(t))$, say up to our ubiquitous resolution parameter ϵ. Here, because of the mode for the ball to be rolling on the floor, there is a clear cut winner. The a posteriori probability for the (unique) rolling state is some constant, bounded away from zero. Each of the other states, accounting for

3D velocities in a neighbourhood of radius ϵ, have posterior probabilities of only $O(\epsilon)$. Thus rolling on the floor is the most probable state, even though it is possible that, as a category it is less probable than motion through the air.

This is not a paradox so much as a warning. The reason the category for motion through the air can be more probable is simply that it is made up of lots (i.e. $O(1/\epsilon)$) of states each of which have probability $O(\epsilon)$. While no single state may be very probable, the total probability over the category can be significant. The choice of which method to use to come up with a preferred interpretation depends on the task. For example, we might consider using the maximally probable state if we wish to intercept the ball. While, for purposes of predicting future events such as possible collisions with the floor, we may wish to consider the most probable category instead.

Our point with these examples is not to generate a detailed model of a ball in a box. Rather, we hope it has demonstrated that by using a qualitative model of the prior probability distribution we have been able to make simple back-of-the-envelope calculations to arrive at plausible inferences about events within an interesting natural context.

6 Summary

We have argued that for a perceiver to make reliable inferences about it's world, the perceiver should make use of world regularities. The class of regularities stressed are those that give rise to modal structures in the distribution of events. The essential ingredient of modal structure is simply that different physical processes generate effects distributed over significantly different scales. For example, the range of velocities of an object while it is in the air is significantly larger than the range due to small vibrations while the object is at rest on the ground. Similarly, differences of scale occur in the positions objects achieve during free fall versus where they come to rest, and also in the accelerations objects undergo due to gravitational or aerodynamic factors versus those during collisions. Once this basic notion that physical processes generate effects over significantly different scales is understood, then other examples that support reliable inferences become apparent, such as rigid 3D motion, articulated motion, skew or reflectional symmetry, or various patterns and groupings of objects.

We have taken such scale separations to be extreme, by assuming that the variation generated by a particular process is negligible, at least in particular directions. This assumption frees us from considering many details, such as precisely how concentrated a particular process needs to be in order for a particular sensor measurement to indicate a world event with 95% confidence. Such details can be added when there is sufficient knowledge about the prior distributions and errors in the sensing process. Their inclusion here, however, would simply mask the important role of modal structure in making certain perceptual inferences reliable.

One application of modal structure is to the notion of a key feature, which is a specialization of the so called non-accidental properties. Roughly speaking, the basic criteria for a particular sensor measurement to be a key-feature for some world property are: i) the feature is highly unlikely to occur in the absence of the property, but often occurs in it's presence; and ii) the prior probability of the property occuring is not negligible. A non-accidental property need only satisfy the first condition here, which ensures the likelihood ratio is large. As we have discussed at some length, this condition alone is not sufficient to lead to a reliable inference. Rather, the reliability of the inference that the particular world property occurs depends critically on the second condition, namely that there must be a significant prior probability.

The presence of this condition on the prior probability has serious implications for perceptual systems. For example, consider a frog that reflexively responds to a dark blob that moves faster than a certain speed, but not at all when the blob moves slower than this speed. Clearly this beast is not computing a posteriori odds in the manner suggested by equation (5). But equally clearly, in more general contexts, the frog certainly would be better off if it could identify situations in which there was no known process which could cause a dark blob, observed to be at rest in the image, to actually be at rest in the world. This is analogous to our ball observed to be at rest in the image and projected against the back wall of the box. The appropriate inference in this case is that the blob is moving along the line of sight and, hence, such a blob would deserve further scrutiny. The more advanced perceptual system, then, should have the ability to represent the critical information needed to make the appropriate inference. Our notion of a context as a class of prior probability distributions over a configuration space, in which various modes are assumed to have nonzero mixture proportions, gives the advanced perceiver this capability. In particular, the mode coupling conditions were shown to be essential in order for a perceiver to arrive at the most probable conclusion, both when the ball is seen against the floor, and when it is seen against the back wall.

More generally, the context sensitivity of modal properties raises the possibility that a perceptual system may use an explicit representation of modal structure. In particular, instead of making an implicit use of modal structure, such as in the speed threshold of our hypothetical frog discussed above, a perceptual system may attempt to maintain an explicit representation of where and when different physical processes, responsible for different modal structures, are active. This is, after all, the critical ingredient in terms of the modal prior probability distributions discussed in this paper. What would one expect to be able to observe about such a system? Perhaps the primary property is that such a system should be able to rapidly learn about a particular class of novel environments, but should be much slower to learn about others. The environments which are rapidly learnable are ones in which the perceiver already knows the various modes, that is, their locations in configuration space, but does not know which modes have nonzero mixing proportions.

A preliminary psychophysical experiment down these lines has already been done (Feldman, 1992). The results provide some support for the hypothesis that an explicit representation of modal structure is involved in human perception, in that modal properties in a novel domain were shown to be learnable from a single example. Further investigation of how we learn to perceive novel environments promises to shed considerable light on how we represent structure in our world.

Acknowledgements

The first author wishes to thank Richard Mann for many fruitful discussions. This work was supported by CIAR, NSERC Canada, IRIS (Canada).

7 References

Albert, M.K. & Hoffman, D.D. (1991) Generic visions. Univ. Calif., Irvine, Math. Behav. Sciences Technical Report MBS-91-23. (Sci. Ameri.?)

Ames, A. (1951) Visual perception and the rotating trapezoidal window. *Psych. Monographs*, 65: No. 329 (entire issue).

Attneave, F. (1982) Pragnanz and soap bubble systems: a theoretical explanation. In J. Beck (Ed.) *Organization and Representation in Perception*, Hillsdale, NJ: Lawrence Erlbaum Assn.

Attneave, F. (1982) The determination of perceived tridimensional orientation by minimum criteria. *Percept. Psychophys.*, 6:391-396.

Binford, T.O. (1981) Inferring surfaces from images. *Artificial Intelligence*, 17: 205-244.

Bennett, B.M., Hoffman, D.D. & Prakash, C. (1989) *Observer Mechanics: A Formal Theory of Perception.* NY: Academic Press.

Biederman, I. (1987) Recognition-by-components: a theory of human image understanding. *Psychol. Rev.*, 94(2): 115-147.

Binford, T.O. (1981) Inferring surfaces from images. *Art. Intell.*, **17**: 205-244.

Bobick, A. (1987) Natural categorization. MIT Art. Intell. Lab. Tech. Report 1001.

Bobick, A. & Richards, W. (1986) Classifying objects from visual information. MIT Art. Intell. Lab Memo 879.

Bregman, A.S. (1990) *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press.

Clark, J.J. & Yuille, A.L. (1990) *Data fusion for sensory information processing systems.* Boston: Kluwer.

Doyle, J. & Sacks, E.P. (1989) Stochastic analysis of qualitative dynamics. MIT Lab for Comp. Sci. Report TM-418.

D'Zmura, M. & Lennie, P. (1986) Mechanisms of color constancy. *Jrl. Opt. Soc. Am. A*, **3**: 1662-1672.

Fahle, M. (1991) Parallel perception of vernier offsets, curvature, and chevrons in humans. *Vis. Res.*, 31: 2149-2184.

Feldman J. (1992) Perceptual categories and world regularities. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA USA

Feldman, J. (1991) Perceptual simplicity and modes of structural generation. *Proc. 13th Ann. Conf. Cog. Sci.*, August.

Foster, D.H. & Wagemans, J. (1993) Viewpoint in variant Weber fractions and standard contour-curvature discrimination. *Biol. Cyber.*, 70: 29-36.

Fowler, R. & Guggenheim, E.A. (1952) *Statistical Thermodynamics.* The University Press, Cambridge.

Geisler, W. (1989) Sequential ideal-observer analysis of visual discriminations. *Psychol. Rev.*, 96: 26-314.

Gershon, R., Jepson, A.D. & Tsotsos, J.K. (1987) Highlight identification using chromatic information. *Proc. IEEE 1st International Conference on Computer Vision, London, England*, pp. 161-170.

Gregory, R.L. (1970) *The Intelligent Eye.* New York: McGraw Hill.

Helmholtz, H. (1963) *Handbook of Physiological Optics.* (Dover reprint of 1925 edition, Ed JPC Southall, 3 volumes).

Helmholtz, H. (1925) *Physiological Optics, Vol. III: The Perceptions of Vision.* (J.P. Southall, trans.) Rochester, NY: Optical Society of America (original publication in 1910.)

Hochberg, J. (1987) Machines should not see as people do, but must know how people see. *Compt. Vision Graphics & Image Proc.*, 37:221-237.

Jepson, A. & Richards, W. (1993) *What is a Percept?* University of Toronto, Dept. of Computer Science Tech Report RBCV-TR-93-43. (Also MIT Cognitive Science Memo 43, 1991.)

Jepson A, Richards W, 1993 "What makes a good feature?" To appear in *Spatial Vision in Humans and Robots* Eds L Harris, M Jenkin (Cambridge University Press) See also Massachusetts Institute of Technology Artificial Intelligence Lab Memo 1356 (1992)

Jepson, A. & Richards, W. (1992) A lattice framework for integrating vision modules, *IEEE Transactions on Systems, Man and Cybernetics*, 22: 1087-1096.

Jepson, A. & Richards, W. (1991) *What's a Percept?* (MIT Cognitive Science Memo 43).

Julesz, B. (1971) *Foundations of Cyclopean Perception.* Chicago: Univ. Chicago Press.

Kanade, T. (1981) Recovery of the three-dimensional shape of an object from a single view. *Art. Intell.*, 17: 409-460.

Kendall, D.G. & Kendall, W.S. (1980) Alignments in two-dimensional random sets of points. *Adv. Appl. Prob.*, **12**: 380-424.

Knill, D.C. & Kersten, D.K. (1991) Ideal perceptual observers for computation, psychophysics and neural networks. In R.J. Watt (Ed.), *Pattern Recognition by Man and Machine*, London: MacMillan.

Koenderink, J.J. (1990) *Solid Shape.* Cambridge, MA: MIT Press.

Koffka, K. (1935) *Principles of Gestalt Psychology.* NY: Harcourt Brace.

Lee, H.-C. (1986) Method for computing the scene illuminant chromaticity from specular highlights. *Jrl. Opt. Soc. Am. A*, **3**: 1694-1699.

Leyton, M. (1992) *Symmetry, Causality, Mind.* Cambridge, MA: MIT Press.

Lowe, D. (1985) *Perceptual Organization and Visual Recognition.* Boston: Klewer.

McMahon, T.A. (1975) Using body size to understand the structural design of animals. *J. Applied Physiology*, 39: 619-627.

Marr, D. (1982) *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* San Francisco: Freeman.

MacKay, D.M. (1985) The significance of 'feature sensitivity'. Chapt. 5 in D. Rose & V.G. Dobson (Eds.), *Models of the Visual Cortex*, New York: Wiley.

Nakayama, K. & Shimojo, S. (1992) Experiencing and perceiving visual surfaces. *Science*, **257**: 1357-1363.

Pearl, J. (1988) *Probabilistic Reasoning in Intelligent Systems.* San Mateo, CA: Morgan Kauffman.

Reuman, S. & Hoffman, D. (1986) Regularities of nature: the interpretation of visual motion. In A. Pentland (Ed.), *From Pixels to Predicates*, Norwood, NJ: Ablex.

Richards, W. & Bobick, A. (1988) Playing twenty questions with nature. In *Computational Processes in Human Vision: An Interdisciplinary Perspective*, Z. Pylyshyn (ed.), Ablex Publishing, Norwood.

Rock, I. (1983) *The Logic of Perception.* Cambridge, MA: MIT Press.

Rubin, J. (1986) Categories of visual motion. Ph.D. Thesis, Dept. of Brain & Cognitive Sciences, Mass. Inst. of Tech., Cambridge, MA.

Shafer, S.A. (1984) Using color to separate reflection components. *Color Res. & Appl.*, **10**: 210-218.

Stevens, K.A. (1988) Line of curvature constraint and the interpretation of 3D shape from parallel surface contours. In W. Richards (Ed.) *Natural Computation*, Cambridge, MA: MIT Press.

Thompson, D'Arcy (1952) *On Growth and Form.* Cambridge: The University Press.

Ullman, S. (1979) *The interpretation of visual motion.* Cambridge, MA: MIT Press.

Wallach, H. & O'Connell, D.N. (1953) The kinetic depth effect. *Jrl. Expt. Psych.*, 45: 205-217.

Waltz, D. (1975) Understanding line drawings of scenes with shadows. In P.H. Winston (Ed.) *The Psychology of Computer Vision*, NY: McGraw-Hill, pp. 19-91.

Watt, R.J. & Andrews, D.P. (1982) Contour curvature analysis: Hyperacuities in the discrimination of shape. *Vis. Res.*, 22: 449-460.

Westheimer, G. & McKee, S. (1977) Spatial configurations for visual hyperacuity. *Vis. Res.*, 17: 941-947.

Witkin, A. & Tenenbaum, J.M. (1983) On the role of structure in vision. In J. Beck, B. Hope & A. Rosenfeld (Eds.), *Human and Machine Vision*, New York: Academic.

Yuille, A.L. & Bülthoff, H.H. (1994) A Bayesian approach to vision: sensor fusion and competitive priors. In W. Richards & D. Knill (Eds.) *Perception as Bayesian Inference*, Cambridge University Press, 1994.