JOURNAL OF MATHEMATICAL PSYCHOLOGY 37, 220–240 (1993)

Modeling Performance in Observer Theory

BRUCE M. BENNETT, DONALD D. HOFFMAN, AND RAMAKRISHNA KAKARALA

University of California, Irvine

We propose a general framework for the study of perceptual capacities, a framework that encompasses theories of perceptual competence and models of perceptual performance. A competence theory of a perceptual capacity (such as stereo vision) describes that capacity in an idealized, information-processing sense, without regard for noise, limited resolution, or limited computational power. A performance model, by contrast, describes the effects of noise, limited resolution, and limited computational power on actual performance as measured in a laboratory. Our framework for studying both competence and performance extends the formal theory of perception known as observer theory. Observer theory provides a canonical form in which to state competence theories of perceptual capacities. This canonical form is called an observer. In this paper we link theories of perceptual competence with models of perceptual performance by constructing an extension to the definition of observer, an extension that we call a "performance extension." To illustrate how the performance extension can aid both the perceptual theorist and the experimental psychologist, we use the performance extension to analyze a psychophysical study of surface interpolation. Finally, we explore a connection between performance extensions and signal detection theory by showing how a signal detection rule can be derived from the performance extension.

1. INTRODUCTION

Researchers into visual perception, and into perception more generally, have found it helpful to distinguish between theories of perceptual competence and models of perceptual performance. A competence theory of a perceptual capacity (such as stereo vision or auditory localization) is a theory that describes the capacity as the solution, given appropriate constraints, to an information processing problem; the competence theory specifies the available sensory data, the desired perceptual solution, and the constraints used to infer the latter from the former, but does so without consideration of noise, resolution, or computational power (Richards, 1988; Ullman, 1986). A performance model, in contrast, describes the effects of noise, limited resolution, and limited computational power, and,

We thank Marc Albert, Myron Braunstein, David Honig, Paresh Murthy, Chetan Prakash, and Scott Richman for many helpful discussions. We also thank two anonymous reviewers for insightful comments which greatly improved the paper. Finally, we thank the editor, Professor Thomas S. Wallsten, for his patience and interest. This work was supported by National Science Foundation Grant IRI-8700924 and by U.S. Office of Naval Research Contract N00014-88-K-0354. Requests for reprints should be sent to Don Hoffman, Department of Cognitive Science, University of California, Irvine, CA 92717.

0022-2496/93 $5.00

ideally, provides a bridge between a competence theory and behavior as measured in the laboratory.

The following example of a pocket calculator, due to Marr and Poggio (1977), illustrates clearly the distinction between competence and performance. A competence theory of a pocket calculator describes the calculator as doing arithmetic on real numbers—adding, multiplying, taking roots, evaluating trigonometric functions, and other similar operations. The theory explains why, for any two entries X and Y, $X+Y$ yields the same answer as $Y+X$. It also dictates that $(\sqrt{2.0})^2 = 2.0$. However, in practice, some calculators yield $(\sqrt{2.0})^2 = 1.999999999$. This departure from the dictates of the competence theory is one of performance: the calculator has finite memory and finite-precision numbers and therefore, in practice, can only approximate the competence theory. We can predict this departure if we know the calculator's precision of decimals and its algorithm for computing roots. This knowledge constitutes part of a model of performance for the calculator.

Competence theories and performance models complement each other. Consider, for example, how one would explain why $\cos(\tan^{-1}(X))$ always yields *approximately* $1/\sqrt{X^2+1}$. The competence theory shows that it is a simple trigonometric identity; the performance model accounts for the error due to truncated Taylor series, memory allocation, and other limiting factors.

For calculators, the differences between performance and the dictates of the competence theory are so slight that we view calculators as actually performing the functions of basic mathematics. For perceptual capacities, however, the differences between performance and the dictates of a competence theory are often much larger. This raises the question: when is the difference between competence theory and actual performance great enough to reject the competence theory? To answer this question, we propose a common framework for analyzing both competence and performance in perception. Our framework uses the formal theory of perception known as *observer theory* (Bennett, Hoffman, and Prakash, 1989). Observer theory proposes a single structure, called an "observer," for modeling all perceptual capacities. An observer model for a perceptual capacity is a competence theory of that capacity stated in a canonical form. After reviewing the observer formalism, we propose the notion of a "performance extension" of an observer; it is intended to model the actual performance—in the presence of noise, limited resolution, or limited computational power—of a perceptual capacity whose competence theory is the given observer. It is thus possible first to describe a perceptual capacity with a competence theory, then to state the competence theory as an observer, and subsequently to investigate performance issues with the performance extension. In this way, the performance extension provides a bridge between theoretical predictions and actual behavior as measured in a laboratory.

After giving the general definition of a performance extension, we use the performance extension to analyze a specific psychophysical task, viz., a task that explores the visual interpolation of subjective surfaces in displays of structure from motion (as reported by Saidpour, Braunstein, and Hoffman, 1991). In this instance, the analysis shows that the data of Saidpour *et al.* disconfirm an existing competence theory of surface interpolation. We then use the performance analysis to suggest an alternative theory, one which is consistent with the results of Saidpour *et al.*

We assume in this paper that the reader is familiar with abstract probability theory, preferably at the level presented in Billingsley (1986) or Bauer (1981). In Section 2 we review the concept of an observer; since observer theory is relatively new, this might serve as an introduction for most readers. In Section 3 we provide a motivated definition of the performance extension. In that section, we discuss at length the principles and modeling strategies that lead to our definition, and provide some predictions for experimental verification. We use the performance extension in Section 4 to analyze the surface interpolation task mentioned above. Finally, in Section 5 we explore some connections between this work and statistical decision theory, showing that a signal detection rule is canonically derivable from each performance extension.

2. Definition of Observer

To make this paper self-contained we now review briefly the definition of observer. More detailed discussions of the definition can be found in Bennett, Hoffman, and Prakash (1989, 1991).

An observer is a competence theory of a perceptual capacity stated in a canonical form. Recall that a competence theory, as defined in the Introduction, describes a perceptual capacity as the solution—under appropriate constraints—to an information processing problem: the competence theory specifies the available sensory data, the desired perceptual solution, and the constraints used to infer the solution from the sensory data. Any competence theory can in turn be described by an observer, $\mathcal{O}$, where $\mathcal{O}$ is a six-tuple $\mathcal{O} = (X, Y, E, S, \pi, \eta)$. In this notation, Y is the set of all possible sensory data, X is the set of all possible perceptual solutions, π is a map from X to Y, and E and S are sets that represent the constraint respectively in the solution space and in the data space. The last component, η, is a probabilistic representation of the perceptual inference from S to E, i.e., from the data that satisfy the constraint to the solutions that satisfy the constraint.

We introduce the concepts underlying an observer by using Ullman's (1979) theory of structure from motion. Ullman analyzes a remarkable ability of the human visual system: we can perceive the relative three-dimensional (3D) positions of moving points when viewing a sequence of two-dimensional images of their motion. Ullman shows how the relative 3D positions of points can be determined from only three distinct views of the points. In particular, Ullman and Fremlin prove the following theorem (see Ullman, 1979, p. 146):

> Given three distinct orthographic views of four (or more) noncoplanar points in a rigid configuration, the structure and motion compatible with the three views are uniquely determined.

As Ullman points out elsewhere, there are actually two solutions, which are mirror images of each other.

The observer representation of this structure-from-motion theory is as follows.

Ullman and Fremlin's theorem states that structure and motion can be recovered from three views of four points—if the points are noncoplanar, the points move rigidly, and the views are obtained by orthographic projection. Let us examine the space of possible sensory data for this task. A single datum contains exactly four orthographic views of three points. Let us assume, with Ullman, that one of the four points is "foveated," i.e., always fixed at the origin. Then, if we assign two-dimensional (2D) coordinates to the image plane projections of the points, we find that each image datum has 18 degrees of freedom—one degree for each coordinate not at the origin. Let Y be the set of all such data; then $Y = R^{18}$. Now let us examine the solution space. For each sensory datum, the corresponding perceptual solution specifies the missing depth coordinates that were lost under orthographic projection. Therefore each solution contains exactly twelve points—three views of four points—where each point lies in R^3. Thus each solution is a set of 27 real numbers (one for each coordinate not fixed at the origin). Consequently the space of all possible solutions, denoted X, is R^{27}. The relationship between the two spaces X and Y is given by the orthographic projection $(x, y, z) \mapsto (x, y)$; the induced map π from X to Y strips off the z coordinates from each element of X.

We now consider the constraints and the inference in Ullman and Fremlin's theorem. The constraints are noncoplanarity and rigidity of motion of the points in space. These constraints define a subset E of the solution space X. Intuitively, each point of E corresponds to a 3D visual interpretation in which four points move rigidly together over three instants of time. Now consider the subset S of the sensory data space Y given by $S = \pi(E)$. Intuitively, each point of S corresponds to a visual display (three 2D views of four points) which can be given two rigid interpretations. To each point of S, Ullman and Fremlin's theorem assigns exactly two solutions in E, where the only difference between the solutions is that one is a mirror reflection of the other. Both solutions are equally valid given the data; the fact that there are two solutions suggests that displays of structure from motion can lead to multistable percepts: a viewer may "flip" back and forth between the two solutions in the same way that a viewer often flips between the different 3D interpretations of a Necker cube. To model multistable percepts, observer theory uses a markovian kernel denoted η. This kernel can be thought of simply as a collection of probability measures on the rigid interpretations E. For each element s of S, $\eta(s, \cdot)$ is a probability measure on the two points in E that are solutions according to Ullman and Fremlin's theorem; these points lie in the set $\pi^{-1}(s) \cap E$. Intuitively, the probability assigned to a point of E by the measure $\eta(s, \cdot)$ indicates the relative likelihood, given the visual display s, of perceiving the 3D interpretation represented by that point.

This structure-from-motion example illustrates the ideas behind an observer; the precise mathematical definition of an observer is as follows.

An *observer* $\mathcal{O}$ is a collection (X, Y, E, S, π, η) with the following properties (see Fig. 1):

(a) X and Y have measurable structures $(X, \mathscr{X})$ and $(Y, \mathscr{Y})$, respectively. The points of X and Y are measurable.

(b) $E \in \mathscr{X}$ and $S \in \mathscr{Y}$.

(c) $\pi: X \to Y$ is a measurable surjective function, such that $\pi(E) = S$.

(d) Let $\mathscr{E}$ denote the σ-algebra on E induced from that of X. η is a markovian kernel on $S \times \mathscr{E}$ such that for each s, $\eta(s, \cdot)$ is a probability measure supported in $\pi^{-1}(s) \cap E$.

A *preobserver* is a collection (X, Y, E, S, π) satisfying conditions (a)–(c). X is the *configuration space* and Y is the *premise space* of $\mathcal{O}$. E is the *distinguished configurations* and S is the *distinguished premises* of $\mathcal{O}$. Finally, π is the *perspective* and η is the *interpretation kernel* of $\mathcal{O}$.

An observer $\mathcal{O}$ "works" as follows. As a result of an interaction with an object of perception, $\mathcal{O}$ receives a premise y in Y. If y is in S, then $\mathcal{O}$ gives interpretations in $\pi^{-1}(y) \cap E$ with distribution $\eta(y, \cdot)$. If y is not in S, then $\mathcal{O}$ gives no interpretations; i.e., it remains inert. In other words, $\mathcal{O}$ gives interpretations only for the distinguished premises (points of S), and the interpretations are given only among the distinguished configurations (points of E). It is in this sense that E and S represent the bias of a perceptual capacity.

In addition to the structure-from-motion example presented above, observers have been formulated for competence theories of stereoscopic vision, light source detection, and other perceptual capacities (Bennett *et al.*, 1989; pp. 30–41). In fact, the *observer thesis* holds that the competence theory of any perceptual capacity is representable by an observer; this thesis, like the Turing hypothesis in computer science, is not provable but is subject to disconfirmation by counterexample (loc. cit.).

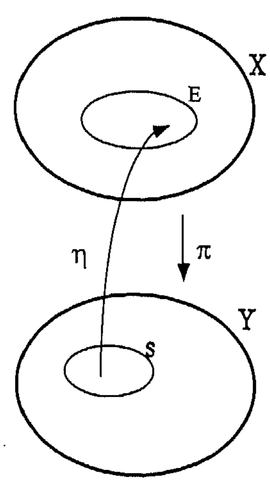

FIG. 1. An observer. Y is the premise space, X is the configuration space, and S and E are respectively the distinguished premises and configurations. π is the perspective map. η is a markovian kernel on $S \times \mathscr{E}$; for each point s of S, $\eta(s, \cdot)$ is supported in $\pi^{-1}(s) \cap E$.

3. Performance Observers and Performance Extensions

As we state in the Introduction, our goal is provide a common framework for analyzing theories of perceptual competence and models of perceptual performance. The competence part of our framework is the observer defined in Section 2; the performance part is formulated below as the "performance observer" and the "performance extension." For an observer $O = (X, Y, E, S, \pi, \eta)$, a performance extension is a pair $(\mathcal{O}', R)$ consisting of a performance observer $\mathcal{O}' = (X, Y, \pi, \eta')$ and a "retraction kernel" R (see Fig. 2); the pair satisfies certain compatibility conditions, which we now discuss.

Let $O = (X, Y, E, S, \pi, \eta)$ represent an observer, and let $\mathcal{O}'$ denote a single instantiation of O, e.g., the structure-from-motion capacity of a particular human subject. Observe that the bias of $\mathcal{O}'$ need not match that of O, as represented by the sets E and S. In Ullman's observer, for example, there are many elements of S which represent displays that are too small or too large for humans to see. Furthermore, elements of $Y - S$ that are "very near" S may be indistinguishable from points of S due to limitations in resolution. Moreover, points of $Y - S$ still further from S may be given interpretations as "nearly rigid" configurations—configurations that are rigid except for a noticeable flaw, e.g., jitter in one of the points. These nearly rigid configurations necessarily lie in $X - E$.

To model performance data, we define for $\mathcal{O}'$ the sets S' and E' as respectively the *maximal* set of premises given interpretations, and the *maximal* set of configurations used as interpretations. We make four comments to clarify this definition. First, the meaning of "maximal" becomes precise when the interpretation kernel η' and the retraction kernel R are defined below. Second, when we say a configuration

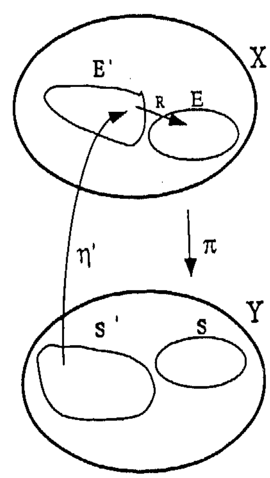

Fig. 2. A performance extension. X, Y, E, S, π are the same as in the observer. η' is the interpretation kernel and R is the retraction kernel.

x is "used as an interpretation," we mean that x represents mathematically a perceptual conclusion given to a premise y. Third, S' and E' are specific to a single instantiation. By experimental evaluation, it may be possible to determine "average" sets S' and E' (i.e., average over different subjects), but that is not required here. Fourth and finally, E' need not contain E as a subset. For example, E in the structure-from-motion observer contains configurations consisting of four noncoplanar points that rotate by an angle too small to detect. These configurations would never be used as interpretations by $\mathcal{O}'$, and hence would not be in E'.

Let us now consider the inference process for $\mathcal{O}'$. For the competence-observer O, the inference process is described by a markovian kernel η, which is defined on $S \times \mathscr{E}$. A performance counterpart to η, which we denote η', should have the following characteristics. First, the "maximality" criterion in the definitions of S' and E' given above requires that η' give interpretations for *all* points of S', and also that η' give these interpretations *only* in E'. Second, η' should be sufficiently general to capture such phenomena as multistable interpretations (such as reversals of the Necker cube) and perceptual learning. A markovian kernel does this effectively: it is a well-defined entity that for a single premise allows more than one interpretation, as required for multistability, and it can allow the relative weights given to different interpretations to change over time, as required for learning.

Should we require the kernel η' to be markovian? Were we to do this, it would follow that for each sensory premise $s' \in S'$, the measure $\eta'(s', \cdot)$ would be a probability measure on E'. This condition is too restrictive for η' for the following reason. Psychophysical studies of perception near absolute threshold indicate that for premises in S', it must be possible to give interpretations in E' only *some* of the time. As the sensory stimulus varies in intensity from below threshold to above threshold a subject's perceptual response is characterized by a smooth transition from never giving interpretations to always giving interpretations. For example, in auditory signal detection tasks, the probability that a subject will identify a stimulus as "signal" gradually increases from zero to one as the signal's intensity is raised from barely audible to perfectly detectable (Green and Swets, 1974). We can account for this by defining η' to be a *submarkovian* kernel on $Y \times \mathscr{X}$; for each $y \in Y$, $\eta'(y, \cdot)$ is a subprobability measure on $(X, \mathscr{X})$. S' is then precisely the measurable subset of Y where η' gives interpretations with nonzero probability:

$$S' = \{y : \eta'(y, X) > 0\}.$$

The number $1 - \eta'(y, X)$ gives, for each y, the probability of no response (or a negative response).

Let us now examine the difference between E' and E. Suppose you observe a motion display in which the points appear to move through space almost, but not quite, rigidly. Perhaps the points are moving rigidly, but with a little vibration that makes their motion not strictly rigid. Often, in this case, you have a good idea of what the rigid motion would be were there no vibration, and you also have a good idea of what the vibration is that is perturbing the points from this rigid motion.

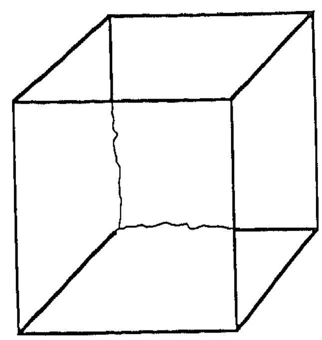

FIG. 3. A Necker cube with warped edges.

In the language of observers, your perceptual interpretation corresponds to a point x that is in $X-E$; i.e., that is not distinguished. However, you have a good idea which point e in E is "closest" to x, i.e., which rigid motion it is that is being perturbed by the vibration. As another example, one can easily identify the object in Fig. 3 as a Necker cube with warped edges; the distinguished part of this percept can be identified as a perfect cube and the flaws as the warpings of its edges.

We describe this effect by a "retraction" R of E' to E (see Fig. 2): The idea of the retraction is to associate to each nondistinguished configuration x_0 in E' a probability measure on E which, intuitively, describes the probability that a given distinguished configuration in E is viewed as the "distinguished part" of x_0. In order to specify R mathematically, we adduce the following considerations. First, the meaning of E' is that every configuration in it should have *some* identifiable distinguished part, although it may be impossible to identify that part with a *specific* element of E. Second, R should be sufficiently general to allow for multistability and learning. Thus we are led to define R to be a submarkovian kernel on $X \times \mathscr{E}$, where $\mathscr{E}$ is the σ-algebra on E induced from that of X. An additional benefit of defining R in this way is that now E' can be defined rigorously:

$$E' = \{x \colon R(x, E) > 0\}.$$

Finally we note that on the set $E \cap E'$, there is no feed for a retraction. Here all the configurations are distinguished, and therefore the identification of the distinguished part is trivial; i.e., it is the configuration itself. Consequently, for every $x_0 \in E \cap E'$, we require $R(x_0, \cdot)$ to be the Dirac measure $\varepsilon_{x_0}(\cdot)$ supported on x_0.

The kernel product $\eta' R$ is a submarkovian kernel on $Y \times \mathscr{E}$. For each $y \in S'$, $\eta' R(y, \cdot)$ represents the distribution of identifications of distinguished parts in E.

Let us now analyze the structure of interpretations given by $\mathcal{O}'$. Here the "logic" of the competence observer O proves useful (for more details, see Bennett, Hoffman, and Murthy, 1993). Recall that for O, the sets E and S enjoy two compatibility relations: first, $\pi(E) = S$, and second, for every $s \in S$, $\eta(s, \cdot)$ is supported in the set $\pi^{-1}(s) \cap E$.

We expect that $\pi(E')=S'$ for the instantiation $\mathcal{O}'$, i.e., that the first compatibility condition is satisfied. This relation states simply that for each atomic interpretation $e \in E'$, the corresponding atomic premise $\pi(e)$ gives rise to an interpretation, and thus $\pi(E') \subset S'$; conversely, the only atomic premises $s \in S'$ that are given interpretations are those compatible with an atomic interpretation $e \in E'$, and thus $\pi(E') \supset S'$.

The second compatibility relation gives η a fibre structure. (A *fibre* of a measurable map $\pi: X \to Y$ is a set $\pi^{-1}(y)$ for $y \in Y$.) A consequence of the fibre structure is that for distinct premises $s_0, s_1 \in S$, the corresponding conclusions $\eta(s_0, \cdot)$ and $\eta(s_1, \cdot)$ are supported in different sets. Should η' have the corresponding fibre structure? That is, for all $y \in S'$, should $\eta'(y, \cdot)$ be supported only in $\pi^{-1}(y)$? Since the fibre structure is important in observer theory, we discuss this question in detail below.

Consider the following motivation for the fibre structure. Perceptual systems, as a rule, propagate noise from premises to interpretations. In structure from motion, for example, any noticeable positional jitter in the 2D premise is also evident as noticeable positional jitter in the 3D interpretation. Similarly, the warped 2D lines of the object in Fig. 3 appear as warped edges of a 3D Necker cube. In both cases, noise that corrupts the premises also corrupts the interpretation. This behavior can be summarized by stating that up to the limit of resolution, an interpretation given to a distinguished premise $s \in S$ should be "perceptually distinct" from that given to a nondistinguished premise y that is formed by adding noise to s. We express this mathematically by stating the following condition on η'.

Condition 1. For a distinguished premise $y \in S' \cap S$ and a nondistinguished premise $y_0 \in S' - S$, $\eta'(y, \cdot)$ and $\eta'(y_0, \cdot)$ should be supported in different sets.

Condition 1 would, of course, be true if η' were to have the fibre structure, for then $\eta'(y, \cdot)$ and $\eta'(y_0, \cdot)$ would lie respectively on the disjoint fibres $\pi^{-1}(y)$ and $\pi^{-1}(y_0)$. We show below that in fact the fibre structure guarantees a much more general relationship between premises and conclusions.

To introduce this relationship, we require a more general representation for premises and conclusions. We have thus far treated premises as elements of Y and conclusions as measures on X. A more natural approach would be to use probability measures for premises as well (Bennett *et al.*, 1989, pp. 29–30).

DEFINITION 2. A *sensory premise* λ is a probability measure on $(Y, \mathcal{Y})$.

This representation for sensory premises is especially useful when there is uncertainty, resulting from noise or measurement error, about the "true" premise. For each $A \in \mathcal{Y}$, $\lambda(A)$ gives the probability that the true premise is in A. If punctual premises—premises that are single elements of Y—are available, then they can be described by Dirac measures $\varepsilon_y(\cdot)$ on points $y \in Y$.

Given a premise λ, the corresponding conclusion given by η' is a subprobability measure ν on $(X, \mathscr{X})$, where for $B \in \mathscr{X}$,

$$\nu(B) = \int_Y \lambda(dy)\, \eta'(y, B). \tag{1}$$

The probability of no response, given the sensory premise λ, is $1 - \nu(X)$. We henceforth write $\nu = \lambda\eta'$ to mean Eq. (1).

We can now make things more rigorous. If two premises μ and λ are mutually singular as probability measures, then they represent logically mutually exclusive input propositions to the perceptual inference process: μ and λ have disjoint supports, and hence the uncertainty associated with μ is exclusive of that of λ and vice versa. Now in general we view an inference process (such as a perceptual capacity) as a mapping from a set of premise propositions to a set of conclusion propositions, a mapping which respects logical relationships. In particular, mutually exclusive premises imply mutually exclusive conclusions. Thus we are led to the following condition on η'.

Condition 3. η' maps mutually singular premises to mutually singular conclusions, i.e., if $\mu \perp \lambda$ then $\mu\eta' \perp \lambda\eta'$.

(Note that for the purposes of Condition 3 we are *not* assuming that all pairs of premise measures are mutually singular, only that *if* a pair is mutually singular, then η' yields conclusion measures that are mutually singular.)

We now establish that for Condition 3 to hold it is sufficient that η' respect the fibres of π.

THEOREM 4. *For Condition* 3 *to hold for all premises* μ, λ, *it suffices that for all* $y \in Y$, $\eta'(y, X) = \eta'(y, \pi^{-1}(y))$, *i.e., that* η' *respect the fibre structure.*

Proof. Suppose $\mu \perp \lambda$. Then there exists $B \subset Y$ such that $\mu(B) = \lambda(B^c) = 0$. If η' has the fibre structure, then

$$\begin{aligned}
\mu\eta'(\pi^{-1}(B)) &= \int_Y \mu(dy)\, \eta'(y, \pi^{-1}(B)),\\
&= \int_Y \mu(dy)\, \eta'(y, \pi^{-1}(B) \cap \pi^{-1}(y)),\\
&= \int_B \mu(dy)\, \eta'(y, \pi^{-1}(y)),\\
&\leqslant \mu(B) = 0,
\end{aligned}$$

where we derive the inequality by noting that $\eta'(y, \pi^{-1}(y)) \leqslant 1$. Similarly, we can prove that $\lambda\eta'(\pi^{-1}(B^c)) = 0$. Since $\pi: X \to Y$ is surjective, $\pi^{-1}(B)$ and $\pi^{-1}(B^c)$ partition X, and hence $\mu\eta' \perp \lambda\eta'$. ∎

Thus Condition 1 and its mathematical generalization, Condition 3, are satisfied if η' has the fibre structure. We note that this fibre structure is not *necessary* for Condition 3 to hold. However, one of the basic ideas of observer theory is that the fibre structure is the "universally efficient" way for Condition 3 to hold; this means that in a situation where Condition 3 holds but the fibre structure fails, the syntax is somehow inefficient, and modifications of the definitions of X, Y, π, η', etc., would yield a more efficient representation of the inference in which η' has the fibre structure. We do not pursue this idea of "efficiency" further here; however, we emphasize that the fibres of π are to be viewed as basic syntactic discriminants of perceptual interpretations.

For these reasons we require that η' have the fibre structure; i.e., we require that for all $y \in S'$, $\eta'(y, \cdot)$ is supported in $\pi^{-1}(y)$. That Conditions 1 and 3 hold in concrete perceptual situations is thus a prediction of our model. In order to test the fibre structure experimenally, one must choose a response variable that is logically dependent on the interpretation kernel η', and not on the identification kernel $\eta' R$. In fact, for premises outside of S, $\eta' R$ never has the fibre structure, since it is supported on $E \subset \pi^{-1}(S)$.

We summarize the principles discussed above in the following definition (compare Figs. 1 and 2).

DEFINITION 5. A *performance observer* $\mathcal{O}'$ is the collection (X, Y, π, η') satisfying the following conditions:

(a) X and Y have measurable structures $(X, \mathscr{X})$ and $(Y, \mathscr{Y})$, respectively. The points of X and Y are measurable.

(b) $\pi: X \to Y$ is a measurable surjective map.

(c) $\eta': Y \times \mathscr{X} \to [0, 1]$ is a submarkovian kernel. η' respects the fibres of π: for all $y \in Y$, $\eta'(y, \pi^{-1}(y)) = \eta'(y, X)$.

DEFINITION 6. Let $\mathcal{O} = (X, Y, E, S, \pi, \eta)$ be an observer, with $(E, \mathscr{E})$ the measurable structure on E. A *retraction kernel* for $\mathcal{O}$ is a submarkovian kernel R on $X \times \mathscr{E}$, such that for all $x \in X$, $R(x, E)$ is either zero or one.

DEFINITION 7. A pair $(\mathcal{O}', R)$ consisting of a performance observer $\mathcal{O}'$ and a retraction kernel R is a *performance extension* of a competence observer $O = (X, Y, E, S, \pi, \eta)$ if the following conditions are satisfied:

(a) The measurable structures $(X, \mathscr{X})$, $(Y, \mathscr{Y})$, and the map π are the same for $\mathcal{O}'$ and O.

(b) Compatibility conditions. Let

$$E' = \{x: R(x, E) > 0\}, \qquad S' = \{y: \eta'(y, X) > 0\};$$

then:

(i) $\pi(E')=S'$.

(ii) $\forall y \in Y: \eta'(y, E')=\eta'(y, X)$.

(iii) $R(x, \cdot)=\varepsilon_x(\cdot)$ for $x \in E \cap E'$, where $\varepsilon_x(\cdot)$ is the Dirac measure on the point x.

(iv) $\eta' R(s, \cdot)=\eta(s, \cdot)$ for $s \in S \cap S'$.

3.1. *Discussion*

We begin with a couple of reminders. First recall that, as discussed earlier, the performance interpretation kernel η' is defined to be submarkovian in order to allow for a variable probability of giving distinguished interpretations: a performance observer, given a premise $y \in Y$, may not be confident that y should be assigned a distinguished interpretation. The number $0 \leqslant \eta'(y, \pi^{-1}(y)) \leqslant 1$ gives the degree of confidence that the premise should be assigned a distinguished interpretation.

Second recall that, as discussed above, the retraction kernel R is defined to be submarkovian in order to distinguish between two kinds of configurations, viz., (1) those which are given positive measure by η' and (2) those which are not. For configurations of type (2), the retraction kernel gives the zero measure. For configurations e' of type (1) the retraction kernel gives a probability measure on E, describing the probability for each $e \in E$ that e' is a perturbation of e. We note that $R(e', E)$ does not take all values between 0 and 1. Why? For a set $A \in \mathscr{E}$, $R(e', A)$ is the probability, according to the performance observer, that e' is a perturbation of a distinguished configuration $e \in A$, *given that* the performance observer accepts e' as a perturbation of some distinguished configuration. The number $\eta'(y, \{e'\})$ is, intuitively, the probability that e' will be so accepted. (The reason that this statement is only intuitive is that in the non-discrete case, $\eta'(y, \{e'\})$ might be zero for all points e' even when $\eta'(y, \pi^{-1}(y)) \neq 0$.)

In Definition 7, $\mathcal{O}'$ is *a* performance extension of O—not *the* performance extension of O—because E', S', η', and R are meant to be specific to a single instantiation. As discussed earlier, it may be possible to determine experimentally an "average" E', S', η', and R, but that is not required here.

The definition reflects, with a minor addition, the principles discussed earlier in this section. The addition is condition (b)(iv), which requires that the distribution $\eta' R(s, \cdot)$ agree with $\eta(s, \cdot)$ on $S \cap S'$. This condition provides a definition of η' that is consistent with that of η. To see this, recall that the definition of the competence observer O leaves undetermined the probability distributions $\eta(s, \cdot)$. These distributions model the relative frequency of interpretations given in E. However, since the relative frequency varies with the instantiation, η can only be determined from experimental measurements of $\eta' R(s, \cdot)$. Condition (b)(iv) then provides a means of determining η from a determination of $\eta' R$.

4. Applications of Performance Observers

Having introduced the formalism of performance observers, it is now appropriate to examine the usefulness of this formalism in analyzing real perceptual tasks. The formalism is, after all, somewhat complex and one might wonder if this complexity is necessary. Is there anything new and useful here that we do not already have in other theories, such as statistical decision theory, that study performance in the presence of noise? In short, does the formalism of performance observers give analytical power to the perceptual theorist, or provide practical tools to the experimental psychologist?

Indeed it does. The essential difference between observer theory and statistical decision theory in the context of perception is worth mentioning. In those cases which typify statistical decision theory *all* uncertainty is due to corruption of the sensorial data (e.g., by noise in the generation, transmission, and registration of sensory information). Hence, assuming a noise-free channel and perfect resolution at the sensory receptors, the decision to be made is, in principle, trivial. By contrast, in perception as modeled by observer theory, even when there is no corruption in the sensorial data there remains nonetheless a fundamental uncertainty and hence a nontrivial decision to be made. This fundamental uncertainty is due to the fact that the mapping from the environment onto the sensorium is, in general, infinite-to-one (e.g., in the visual mapping a 3D environment *projects* onto a 2D retinal image). Thus, even when given perfect sensory information, the state of the environment is infinitely underdetermined (e.g., a perfectly noise free retinal image, if such were possible, would still have infinitely many 3D interpretations). The raison d'être of observer theory is to model the inductive (i.e., nondemonstrative) inferences made by perceptual systems in consequence of this fundamental uncertainty. In short, *the inferences which typify statistical decision theory trivialize in the absence of noise, whereas those of observer theory do not.* The entire competence-observer structure (X, Y, E, S, π, η) characterizes the perceptual inference that is required *in the absence of all noise.* The performance observer then extends this competence-observer structure to deal with noise and to model tolerances in the process of making perceptual interpretations.

Statistical decision theory is not a theory of perception. However, it potentially applies to one aspect of the perceptual process, viz., to decision tasks necessitated by noise in the data received at a sensorium. The structure of the performance observer is essential for the modeling and the analysis of perceptual tasks which are more than simply this kind of decision task.

To see this, it is best to examine a concrete example. We will consider a psychophysical experiment reported by Saidpour, Braunstein, and Hoffman (1992) which investigates the perception of subjective surfaces in displays of structure from motion. The motivation for the experiment derives from a remarkable feature of certain motion displays, e.g., displays which simulate the motion of dots rigidly attached to a rotating invisible cylinder. Each frame of such a display consists of but a few white dots against an otherwise black background. But when the display is put in motion, one perceives the dots to be moving in three dimensions, e.g., as though attached to a cylinder. Moreover, and here is the remarkable feature, one not only sees the dots in three dimensions, one also sees (perhaps "hallucinates" is better) a *surface*, e.g., a cylinder surface, filling the empty space between the dots. The surface is not real, in the sense that there is no physical surface passing through the dots, but is instead a *subjective surface*. Saidpour *et al.* investigated the perceived shape of this subjective surface in regions between visible dots, and used their results to evaluate the psychological plausibility of one theory of surface interpolation, viz., Grimson's (1981) "quadratic variation" theory.

To do this, Saidpour *et al.* showed subjects displays which simulated the motion of dots rigidly attached to an invisible cylinder. The cylinder rotated back and forth through a total angle of 38° about a vertical axis. A narrow vertical strip of the cylinder (the "gap") was devoid of dots. A single dot, a so-called structure-from-motion "probe dot," was placed in the center of this gap and could be moved by the subject under joystick control either inward, towards the major axis of the cylinder, or outward, towards the subject. The subject's task was to place the dot as accurately as possible on the perceived subjective surface.

Subjects were quite good at this task, in the sense that if the dot density was sufficiently high and the gap was not too large, then the standard deviation in their placements of the probe dot was quite small. The result of the experiment was unexpected. Apparently subjects perceived a surface that bulged out slightly in the region of the gap: their mean placement of the probe dot was well outside of an interpolating cylinder, and also well outside of the position predicted by Grimson's quadratic variation functional. The difference between Grimson's prediction and the subjects' performance was statistically significant, leading Saidpour *et al.* to suggest that the quadratic variation functional does not properly model human perception.

Before we analyze this task using performance observers, note that the task would be nontrivial even if the displays were completely noise free: the central feature of this task is that subjects are inferring a 3D interpretation from 2D images, and the fundamental uncertainty inherent in such an inference is in no way reduced by reducing noise in the 2D images.

An observer-theoretic analysis of this task begins with the construction of two competence observers. The first observer models an inference whose premise is the 2D motions of the dots and whose conclusion is a 3D interpretation of the positions and motions of the dots. The second observer models an inference whose premise is the conclusion of the previous inference, viz., the 3D positions of the dots, and whose conclusion is a smooth surface passing through these 3D positions.

For the first observer we can use Ullman's structure-from-motion observer discussed in Section 2; the displays used by Saidpour *et al.* consist of at least three views of four or more noncoplanar points in rigid motion, so the conditions for Ullman's theorem are satisfied.

To construct the second observer, we use the following theorem proved by Grimson (1981):

Given four noncoplanar points in space, there exists a unique surface $S=\{(x, y, z) \mid z=f(x, y)\}$ (where f is some function $f: R^2 \to R$) which passes through the four points and which minimizes the quadratic variation functional $\Theta(f)$ defined by

$$\Theta(f)=\left\{\iint (f_{xx}^2+2f_{xy}^2+f_{yy}^2)\,dx\,dy\right\}^{1/2}$$

(where the subscripts indicate partial derivatives).

The premise space Y for this observer consists of all sets of four points in three space, i.e., $Y=R^{12}$. The distinguished premises S is that subset of Y consisting of all sets of four *noncoplanar* points in three space. For each element $y \in Y$, representing a particular set of four points in three space, denote the set of all smooth surfaces which pass through those points by the symbol $\pi^{-1}(\{y\})$. Then the interpretation space X is $X=\bigcup_{y \in Y} \pi^{-1}(\{y\})$. The perspective map $\pi: X \to Y$ is the map whose fibres just are the sets $\pi^{-1}(\{y\})$. For each $s \in S$, Grimson's theorem assures us that there is precisely one surface, E_s, in $\pi^{-1}(\{s\})$ that minimizes the quadratic variation functional Θ. We let the distinguished interpretations E be the subset of X defined by $E=\bigcup_{s \in S} E_s$. Finally, the interpretation kernel η is defined by $\eta(s, \cdot)=\varepsilon_{E_s}$, where ε_{E_s} denotes the Dirac measure at the point E_s.

These two competence observers model the perceptual inference in the task of Saidpour *et al.* under the assumption that there are no corrupting influences such as noise or discretization errors. The output of the first observer is the input to the second. The conclusion of Ullman's observer is a description of the 3D positions of four noncoplanar points (generalizations to more points are straightforward). This description of the 3D positions of the four points feeds in as a distinguished premise to Grimson's observer. Given such a premise, the conclusion of Grimson's observer is a (unique) surface which passes through the four points and minimizes the quadratic variation functional. For future reference, we will call this surface the *predicted Grimson surface.*

Now we allow corrupting influences. In this case the two competence observers alone do not adequately model the subjects' perceptual inferences; we must now consider performance extensions to these observers. In particular, the interpretation kernel η of Grimson's observer, which for each distinguished premise s gives a Dirac measure (i.e., a measure having no dispersion) at the point representing the predicted Grimson surface E_s, must be replaced with a performance interpretation kernel η' having nonzero dispersion. And we must introduce a retraction kernel R relating the events charged by η' to the E of Grimson's competence observer. (Similar comments hold *mutatis mutandis* for Ullman's competence observer as well.)

Given these changes we can discuss the findings and conclusions of Saidpour *et al.* The (expected) finding that there is nontrivial dispersion in the placements of the probe dot is modeled in the performance observer by the nontrivial dispersions in the probability measures $\eta'(y, \cdot)$. (It seems reasonable to assume that the dispersion in the subjects' responses is primarily perceptual, not motor. The motor task was straightforward, and subjects were given as much time as they needed to place the probe dot on the surface they saw.) But here we must be careful. The probability measure $\eta'(y, \cdot)$ is by definition a probability measure on interpolating *surfaces*, whereas in the experiment each probe point placement gives data on the fibre $\pi^{-1}(y) \subset X$ for our second observer; i.e., it is a probability measure on only *one point* of the interpolating surface, viz., the point at the probe dot. However, on the hypothesis that Grimson's theory is a valid psychological model, we can take each probe point placement to represent that surface which minimizes quadratic variation and interpolates the probe point and all the other visible points; we will call these surfaces the *probe point surfaces.* Under this hypothesis we can view $\eta'(y, \cdot)$ as a measure on the set of possible probe point placements. Note that there is a distinct probe point surface for each probe point position. Moreover, only one possible probe point position lies on the predicted Grimson surface, which then coincides with the corresponding probe point surface.

Saidpour *et al.* reject the hypothesis that Grimson's theory is a valid psychological model. Their argument, as we have mentioned, is that the mean probe point placement is well away from the position predicted by Grimson's theory, and that the difference is statistically significant. Here we see an argument which goes from data on perceptual judgments to a conclusion about a competence theory. Within the formalism of performance observers, the mathematical structure which relates perceptual judgments to competence theories is the retraction kernel $R: X \times \mathscr{E} \to [0, 1]$. For distinguished premises $y \in S$, as in the present case, we must have by definition of R (see Definition 7) that $\eta' R = \eta$. In particular we may view the experiment of Saidpour *et al.* as revealing a performance observer with interpretation kernel η' which codifies the statistics of the subjects' probe point placements. To say that this experiment supports Grimson's theory is to say, in our language, that this performance observer is a performance extension of Grimson's competence observer. In particular, by definition of a performance extension, if η denotes the interpretation kernel for Grimson's competence theory this condition obtains only if $\eta = \eta' R$ for some reasonable retraction kernel R. For our purposes we may consider $\eta(y, \cdot)$ to be Dirac measure centered at the unique probe point placement P_0 representing the Grimson surface through the original display points. In the experiment, however, the distribution of probe point placements is centered about a point well away from P_0. For it to be the case that $\eta' R = \eta$ under these conditions, the retraction kernel R must be biased, i.e., it must map the measure $\eta'(y, \cdot)$ to a point well away from its mean. Since the required bias was statistically significant, and since Saidpour *et al.* saw no reason that corrupting influences such as noise and quantization should induce such a bias into their data, they concluded that the E of Grimson's competence theory was not plausible. Whatever competence-observer theory is proposed, they concluded, must incorporate an E which passes much nearer to the mean of the probe point placements.

In short, using the language of performance observers, Saidpour *et al.* reject Grimson's competence observer because the only way to relate the experimental data to the E of Grimson's observer is via a biased retraction kernel, and there is no reason to expect in this experiment that the corrupting influences should introduce any bias (they should only increase dispersion). Although Saidpour *et al.* do not mention this, the same reasoning can be used to reject interpolation models based on cubic splines, hermite polynomials, and bezier curves. This raises the question: What interpolation models might still be acceptable?

We mention one, briefly, and only in the two-dimensional case (the extension to three dimensions is straightforward but lengthy to describe). Suppose one is given the visible points $P_i = (x_i, y_i)$, $i = 1, 2, ..., n$, ordered so that $x_i \leqslant x_j$ iff $i \leqslant j$. One computes the difference vectors $\mathbf{v}_i = P_{i+1} - P_i$. One then assigns to point P_i a tangent vector $\mathbf{t}_i$ which is a weighted sum of the difference vectors $\mathbf{v}_{i-1}$ and $\mathbf{v}_i$, the shorter of these two difference vectors getting the greater weight, i.e., $\mathbf{t}_i = (\|\mathbf{v}_i\| / \|\mathbf{v}_{i-1}\|^2)\, \mathbf{v}_{i-1} + (1/\|\mathbf{v}_i\|)\, \mathbf{v}_i$. One then computes, for all i, the intersection point Q_i of the line $P_i + \alpha \mathbf{t}_i$ with the line $P_{i+1} + \beta \mathbf{t}_{i+1}$, where α and β vary in $\mathbf{R}$. The points Q_i are then inserted into the list of points P_i and the entire procedure repeated. One iterates this process to the desired resolution. Alternately, one can stop at any iteration after the first and interpolate the updated list of points using Grimson's procedure.

Returning to the discussion of performance observers, note that our analysis of the experiment by Saidpour *et al.* cannot be done in the language of statistical decision theory alone, for the analysis first requires a detailed understanding of Ullman's competence theory and Grimson's competence theory, and both these theories describe the nontrivial perceptual inferences required to perform the task *when there are no corrupting influences such as noise and quantization.* One needs the language of competence observers (to describe the perceptual inference in the absence of noise) and of performance observers (to model the addition of corrupting factors) to give an adequate analysis of this task.

Note that the theory of performance observers gives only general conditions on the acceptability of retractions (e.g., they must be representable by submarkovian kernels). And this is appropriate. The acceptability of a particular retraction depends on the nature of the corrupting influences that one expects to obtain in a particular experimental situation. If there is reason to expect that these influences are systematically biased, as when one wears prism lenses, then there is reason to find acceptable a correspondingly biased retraction kernel. If, as in the experiment just discussed, there is no bias expected, then there is no reason to accept a biased retraction kernel. In this fashion the theory of performance observers provides the link between psychophysical data and competence-observer theories that allows one to use such data to confirm or disconfirm competence observers.

5. Decision Strategy

We now explore a connection between the performance extension $\mathcal{O}'$ and statistical decision theory. The connection concerns the following interesting fact: a performance extension contains enough information in its components to formulate a rule—analogous to the well-known likelihood ratio rule—to decide whether or not a point premise y is compatible with a distinguished configuration, i.e., whether or not y is compatible with an e in E. The rule that we are about to establish is of the same *form* as the likelihood ratio rule, but *need not* be equivalent to the likelihood ratio unless $\mathcal{O}'$, or some observer "higher up" with respect to $\mathcal{O}'$, knows the correct probability densities. (Observer theory provides a precise definition of how one observer can be "cognitive," or higher up, with respect to another, see Bennett *et al.* (1989, pp. 41–52).)

Let us first discuss the standard decision problem in the language of statistical decision theory (Van Trees, 1968). A decision maker receives an observation ω in an observation space Ω. From this, it must decide between two hypotheses, denoted H_1 and H_0, about ω. Let $p(\omega \mid H_1)$ and $p(\omega \mid H_0)$ denote respectively the conditional probability densities of ω under H_1 and H_0. Furthermore, let P_1 and P_0 represent respectively the a priori probabilities of H_1 and H_0. Finally, let C_{ij} represent the cost associated with deciding H_i given that H_j actually occurred. Then the decision rule that minimizes expected cost is to decide H_1 iff (Van Trees, p. 26)

$$\frac{p(\omega \mid H_1)}{p(\omega \mid H_0)} > \frac{P_0}{P_1}\kappa, \tag{2}$$

where $\kappa = (C_{10} - C_{00})/(C_{01} - C_{11})$ represents a ratio of costs.

If we now replace Ω by the premise space Y of a performance extension $\mathcal{O}'$, then $\mathcal{O}'$ has an important detection problem, which is as follows. Given a premise y, $\mathcal{O}'$ must decide between hypothesis H_1 that y is compatible with an $e \in E$ (the distinguished configurations), and the null hypothesis H_0 that y is not compatible with an $e \in E$. The optimal (minimal cost) rule for this problem is given in Eq. (2), where the relevant probability densities on Y are $p(y \mid H_1)$ and $p(y \mid H_0)$. However, these densities are not explicitly encoded into any of the components in $\mathcal{O}'$. In other words, $\mathcal{O}'$ may not "know" $p(y \mid H_1)$ and $p(y \mid H_0)$.

For formulating a detection rule, $\mathcal{O}'$ has available the function $\alpha(y) = \eta' R(y, E)$. We call $\alpha(y)$ the *cathexis* for a premise y. It represents the *subjective a posteriori* conditional probability that a premise y is compatible with E, i.e., the subjective $p(H_1 \mid y)$. The cathexis α need not have any formal relation to the "true" probability $p(H_1 \mid y)$. However, if $\mathcal{O}'$ were to synthesize the best possible rule *solely from the information encoded in its components*, then it can do no better than to treat its cathexis $\alpha(y)$ as the real $p(H_1 \mid y)$. The optimal detection rule using α takes the following form.

THEOREM 8. *For a performance extension $\mathcal{O}'$, the detection rule based on α that minimizes expected cost is to decide H_1, i.e., that a premise y is compatible with a distinguished configuration in E, if and only if*

$$\frac{\alpha(y)}{1 - \alpha(y)} > \kappa. \tag{3}$$

Proof. The likelihood ratio rule of Eq. (2) can be written (replacing ω by y):

$$\frac{p(y \mid H_1)\, P_1}{p(y \mid H_0)\, P_0} > \kappa.$$

Let $p(y) = p(y \mid H_1)\, P_1 + p(y \mid H_0)\, P_0$ represent the probability density of y. For detection purposes, we can ignore the set of y where $p(y) = 0$, because this set will not affect the probabilities of error. Outside of this set, we can divide both top and bottom of the left-hand side by $p(y)$ to obtain

$$\frac{p(y \mid H_1)\, P_1/p(y)}{p(y \mid H_0)\, P_0/p(y)} > \kappa,$$

or, using Bayes' rule,

$$\frac{p(H_1 \mid y)}{p(H_0 \mid y)} > \kappa.$$

Replacing $p(H_1 \mid y)$ by $\alpha(y)$ and $p(H_0 \mid y)$ by $1 - \alpha(y)$, we have

$$\frac{\alpha(y)}{1 - \alpha(y)} > \kappa. \quad \blacksquare$$

Theorem 8 shows that the minimal cost rule that $\mathcal{O}'$ can synthesize based solely on its subjective probabilities is Eq. (3). We reiterate that the rule in Eq. (3) need not have any formal relation to the rule in Eq. (2), simply because the subjective probabilities need not be related to the "true" probabilities. However, in the special case that $\mathcal{O}'$ learns the true probabilities, presumably by the result of experience and feedback, then the rule in Eq. (3) is optimal, because the corresponding likelihood ratio rule in Eq. (2) is optimal as well.

6. Conclusions

Let us now review the benefits to both the perceptual theorist and the experimental psychologist of the performance extension framework. First, the theorist should find that by using the performance extension, the analysis of performance issues benefits from the same precision that the observer affords to theories of competence. The mathematical analysis of noise, for example, is couched almost exclusively in the language of stochastic processes. Here observer theory and the proposed performance extension are "at home," as their mathematical formulation is in the language of abstract probability theory, and they are well suited for using probabilistic representations of premises and perceptual conclusions. Second, the experimentalist should find that the performance extension provides a basis for principled study of the question of when a competence theory is confirmed or disconfirmed by experimental measurements. In our language, the question can be stated in the following more precise manner: when does a competence observer have a performance extension that accounts for the experimental data? The example given in Section 4 of surface interpolation with Grimson's theory shows how a competence observer can fail to have a performance extension that accounts for the experimental data. On the other hand, the alternative competence theory of surface interpolation proposed in that section does have a performance extension that accounts for the experimental data. Finally, another interesting feature of performance extension analysis is that it incorporates a subjective signal detection rule that is analogous to the likelihood-ratio rule.

We now consider some issues for future study. The kernels η' and R have not been formally linked other than by the compatibility conditions of Definition 7. It is interesting to consider what their relation might be as the cathexis $\alpha(y)$ increases. As $\alpha(y)$ increases, a performance extension $\mathcal{O}'$ becomes increasingly "sure" that a premise y is compatible with a distinguished configuration. We expect that as this happens, the retraction mapping R becomes increasingly specific about the element of E that generates the cathexis. Informally, we are claiming that "the more sure you are that you see something, the more sure you are of what it is exactly that you see." We state this as an empirical conjecture: for premises $y \in Y$ and configurations $x \in \pi^{-1}(y)$, as the cathexis $\alpha(y)$ increases, the dispersions of the retraction measures $R(x, \cdot)$ decreases.

To verify this conjecture, one needs a measure of dispersion that decreases as the retraction $R(x, \cdot)$ gives higher probability to smaller sets. There are several ways of determining the dispersion of a probability measure, but each requires assumptions that may not always hold in abstract spaces such as E. Measures of dispersion can be constructed from the variance of $R(x, \cdot)$, the entropy of $R(x, \cdot)$, and the L_2 norm of the density of $R(x, \cdot)$ with respect to some underlying unbiased measure. These and other measures need to be tested to determine which, if any, properly encodes our notation of dispersion.

References

Bauer, H. (1981). *Probability theory and elements of measure theory* (2nd English ed.; R. B. Burkel, Trans.). New York: Academic Press.

Bennett, B. M., Hoffman, D. D., & Prakash, C. (1987). Perception and computation. In M. Brady (Ed.), *Proceedings of the IEEE First International Conference on Computer Vision* (pp. 356–364). New York: IEEE Press.

Bennett, B. M., Hoffman, D. D., & Prakash, C. (1989). *Observer mechanics*. New York: Academic Press.

Bennett, B. M., Hoffman, D. D., & Prakash, C. (1991). Unity of Perception. *Cognition*, **38**, 295–314.

Bennett, B. M., Hoffman, D. D., & Murthy, P. (1993). Lebesgue logic for probabilistic reasoning and some applications to perception, *Journal of Mathematical Psychology*, in press.

Berger, J. O. (1985). *Statistical decision theory and bayesian analysis* (2nd ed.). New York: Springer-Verlag.

Billingsley, P. (1986). *Probability and measure* (2nd ed.). New York: Wiley.

BRAUNSTEIN, M. L., HOFFMAN, D. D., SHAPIRO, L. R., ANDERSON, G. J., & BENNETT, B. M. (1987). Minimum points and views for the recovery of three-dimensional structure. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 335–343.

GEISLER, W. G. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, **96**, 267–314.

GREEN, D. G., & SWETS, J. A. (1974). *Signal detection theory and psychophysics.* New York: Krieger.

GRIMSON, W. E. L. (1981). *From images to surfaces: A computational study of the human early visual system.* Cambridge, MA: MIT Press.

HILDRETH, E. (1984). *The measurement of visual motion.* Cambridge: MIT Press.

HOFFMAN, D. D., & BENNETT, B. M. (1985). Inferring the relative three-dimensional positions of two moving points. *Journal of the Optical Society of America A: Optics and Image Science*, **2**, 350–353.

MARR, D. C., & POGGIO, T. (1977). From understanding computation to understanding neural circuitry. *Neurosciences Research Progress Bulletin*, **15**, 470–488.

LONGUET-HIGGINS, H. C. (1982). The role of the vertical dimension in stereopsis. *Perception*, **11**, 377–386.

RICHARDS, W., (Ed.). (1988). *Natural computation.* Cambridge: MIT Press.

SAIDPOUR, A., BRAUNSTEIN, M. L., & HOFFMAN, D. D. (1992). Interpolation in structure from motion. *Perception and Psychophysics*, **51**, 105–117.

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

ULLMAN, S. (1984). Visual routines. *Cognition*, **18**, 97–160.

ULLMAN, S. (1986). Competence, performance, and the rigidity assumption. *Perception*, **15**, 644–646.

RECEIVED: September 18, 1990