Toward a computational model of constraint-driven exploration and haptic object identification

Robert L Klatzky  
University of California, Santa Barbara, CA 93106, USA

Susan J Lederman  
Queen's University, Kingston, Ontario K7L 3N6, Canada  
Received 3 August 1992

Abstract. A conceptual model of the human haptic system in relation to object identification is presented. The model encompasses major architectural elements including representations of haptically accessible object properties and exploratory procedures (EPs)—dedicated movement patterns that are specialized to extract particular properties. These architectural units are related in processing-specific ways. Properties are associated with exploratory procedures in keeping with the extent to which a given procedure delivers information about a given property. The EPs are associated with one another in keeping with their compatibility, as determined by parameters of motor execution and interactions with the object and the workspace. The resulting architecture is treated as a system of constraints which guide the exploration of an object during the course of identification. The selection of the next step in a sequence of exploration requires that constraints be optimally satisfied. A network approach to constraint satisfaction is implemented and shown to account for a number of previous empirical results concerning the time course of exploration, object classification speed, and incidental learning about object properties. This system has potential applications for robotic haptic exploration.

1 Nature of haptic object identification

Object identification consists, in essence, of establishing the link between sensory information and categorical knowledge—the Hofding step, as it was called by Gestalt psychologists. To understand object identification, then, one must deal with a broad range of problems, including how the stimulus is codified and how objects are represented in memory. These issues have received substantial attention for such stimulus domains as natural and fabricated objects, letters, words, and faces. A common assumption in this work, however, has been that vision is the sensory modality by which objects are recognized and categorized. In contrast, our work focuses on the recognition of objects through haptic perception.

Haptics is a perceptual system that is based on sensory receptors embedded in the skin, muscles, tendons, and joints. The system comprises a number of distinct sensory mechanisms: mechanoreceptors that innervate the skin and primarily sense pressure and vibrations; thermoreceptors within the skin that sense steady and changing skin temperature; and a variety of mechanoreceptors in muscles, tendons, and joints that encode limb position and movement. Whereas visual receptors code the effects that the surface of an object imposes on the light reaching the eyes, the haptic receptors code the effects that an object imposes directly on the body, such as the ensuing deformation of the skin, the heat flow between skin and object, and the joint angles between contacting end-effectors.

Ultimately, this proximal sensory information is converted to information about distal properties of the touched object. We use the term property here to denote a variable that can be assigned a value for a given object. The property value is usually enduring in any one object but variable across the object pool. We will use the terms 'attribute' and 'dimension' interchangeably with 'property', while recognizing that a dimension of variation can be simply nominal. It should also be noted that a property
can be described at different levels of specificity or precision and potentially with many measures. An example is roughness. Although we might say that the surface of an object has a ‘high’ level on the roughness dimension, alternative descriptions of surface texture might use other parameters such as the spatial density of raised surface elements, their sharpness, or the coefficient of friction [which Taylor and Lederman (1975) have found to be independent of perceived roughness].

We began our study of haptic object identification with two provocative observations. The first documented that there was a phenomenon worth studying: people are remarkably good at recognizing familiar objects by touch alone (Klatzky et al. 1985). We presented subjects with 100 common objects and measured their naming latency and accuracy. With a strict criterion, performance was at 96%, with a modal response latency of under 2 s. Allowing false starts and near misses brought the accuracy rate to 99%.

This level of performance was somewhat surprising, given the prevalent belief that haptics was a poor substitute for vision in tasks involving stimulus discrimination and identification (eg Bryant and Raz 1975; Cashdan 1968). However, in experiments supporting this view, generally fully three-dimensional, multiattribute objects had not been used. Often, shape was the only discriminating attribute. In other studies, the size of the objects was not necessarily scaled to accommodate peripheral limitations, such as the spatial resolution of the skin (see Loomis 1990), or, at more central levels, limits on memory or integration processes needed to encode contours over time (Loomis et al. 1991).

We have argued that tasks involving matching or categorization of two-dimensional contours by touch underestimate the considerable capacities of the haptic system for processing information about objects. Such tasks promote a strategy whereby subjects attempt to form and use a visuospatial image of haptically explored contours. For example, when subjects attempted to name raised-line drawings of objects, an image-mediation strategy was indicated by positive correlations between their accuracy and both (a) independent ratings of the imageability or vividness of the pictures and (b) measures of the subjects’ general imagery ability (Lederman et al. 1990). It is not surprising that subjects chose to use imagery, nor is it surprising that they performed worse with imagery-based haptic tasks than they did when vision itself was used for the same tasks. Although image mediation is of substantial interest in its own right and may play some role in object recognition by touch, especially by sighted individuals, our work suggests that it is not the principal means by which we perform haptic recognition of ecologically valid common objects.

The second observation initially motivating our work was that during the course of object identification, the haptic system provides visible cues to internal processes, in the form of specialized patterns of exploration that we call ‘exploratory procedures’ (EPs) (Lederman and Klatzky 1987). An EP is a stereotyped hand movement that is spontaneously used when extracting a particular object property and is generally optimal (if not necessary) for extracting that property. The principal EPs and their associated properties are depicted in figure 1. For example, the EP called Lateral Motion—typically, repetitive tangential movement between the skin and the object—is used to extract information about the roughness of the surface of an object. Some of these patterns had previously been noted by others as well (Gibson 1966; Katz 1925; Lotz 1856/1885). A more detailed description will be given below.

The occurrence of specific patterns of hand/object (or more generally, end-effector/object) interaction has important theoretical and methodological implications for the study of haptic information processing. On the theoretical side, there are a number of powerful processing constraints associated with the EPs. First, as we will describe further below, the EPs play a ‘gateway’ role in perception. That is, the movements
that are performed determine what information is acquired about a touched object. Conversely, the anticipation of certain attributes of an object acts top–down to determine which EPs are executed. Furthermore, the potential for performing two EPs together determines which properties can be extracted in tandem and thus influences higher-order processes of integration and decision. In short, a description of the interplay between exploration and information extraction is central to any model of object identification involving touch.

Our research also capitalizes on the methodological usefulness of EPs. Exploration, we argue, provides essential data for understanding object recognition by touch. The EPs are reliably codeable from videotaped records, and their frequencies of occurrence serve to illuminate the processing that occurs during object apprehension, identification, and categorization.

Our purpose in this paper is to present a general model of haptic object identification—one that deals with the 'macrostructure' (Schneider and Detweiler 1987) of haptics as an information-processing system. The model specifies essential structures and the potential for flow of information between them. Processing interactions between these structures can be represented as activation and inhibition within a network with weighted connections. In essence, then, the model takes a connectionist approach. However, formal connectionist models often focus on the model-based development of weights. Our approach is somewhat different: we empirically derive weights with at least ordinal relations from a substantial body of data on haptic object recognition. The associative architecture is then used to explore further implications of this weight pattern. The resulting model accounts for a number of effects, including the frequency and order of EP execution during haptic object identification, response times for object categorization, and limitations on incidental learning about object properties.

As indicated in figure 2, our general approach to haptic object processing can be understood by analogy to the computational model of reading developed by Just and Carpenter (1980, 1987), which spans processes from eye fixations to comprehension. This model is particularly relevant to present concerns because the hand movements of a haptic explorer, like eye movements, provide direct information about internal processing. Carpenter and Just divide text comprehension into several stages, including moving the eye to a new location, local (i.e., lexical) interpretation, and building a

Figure 1. Exploratory procedures (EPs) and associated object properties (adapted from Lederman and Klatzky 1987 for Lederman 1991).
global representation. Stages are executed as fully as possible within one fixation and then recur over successive fixations, with inputs from each new fixation being integrated into the ongoing representation.

The fixation period in reading is analogous to an exploratory period in our model. What we call the selection-extraction loop occurs during this period. An EP (or potentially, more than one) is selected and executed at some object location, on the basis of activation flow. The data resulting from EP execution are used to assign a local interpretation to an object region, which is used in turn to build a global representation of the object that can be matched with stored categorical representations. Local object processing and global object processing occur as fully as possible within an exploratory period, and these same processes recur over successive exploratory periods and object locations. The outcome of each period is either selection of the next exploratory procedure(s) to execute or termination in the form of a recognition decision.

**Figure 2.** Analogous stages between the reading model of Just and Carpenter (1980, 1987) and haptic object recognition.

### 2 Haptic representations and their associative structure

At this point we will describe the structure of the haptic object-recognition system, in terms of the types of data represented and the linkages among representations that we assume to exist. The essential representations within our system and the links between them are identified in figure 3. An object component represents specific objects and their particular property values. A property component represents the attributes along which objects potentially vary. An EP component represents exploratory procedures. Links between the various components reflect relations, the nature of which depends on the linked elements themselves. For example, the link between a property and an EP reflects the precision with which executing that EP enables the system to discriminate between values of the property. The neurophysiological underpinnings of these components include mechanisms that control and execute the EPs and that process ensuing information extracted from sensory receptors. We refer to these mechanisms as the sensorimotor component, but we do not attempt to model this component of the system.
Figure 3. Model of the macrostructure of haptic object identification.

2.1 Object component

The object component comprises object identities and their specific property values. A substantial body of work in cognitive psychology deals with the representations of objects. The most relevant aspects of this research for our purposes concern the associative connections between different object categories and between objects and property values. For example, in his seminal work Rosch (eg 1978) has described hierarchical nesting of object categories and has developed the concept of a privileged ‘basic’ level, which designates the common name of the object and serves as the usual ‘entry level’ (Jolicoeur et al 1984) in perceptual recognition. The basic level is distinguished from more inclusive superordinate levels and more specific subordinate levels. We adopt the conventions of semantic-network modeling (eg Collins and Loftus 1975), which assume that weighted links connect objects to their superordinate and subordinate categories and connect objects at all levels to particular property values. For example, associated with the object ‘rubber eraser’ is a particular value (or range) on the property of hardness.
2.2 Property component

This component consists of representations of object properties, as the term is defined above. Note that properties per se (e.g., hardness) are represented, not the values associated with specific objects (e.g., the hardness of a rubber eraser). We are interested, in particular, in those properties that are accessible to the sense of touch. (This excludes electric conductivity, magnetic properties, and purely optical properties.) To date, our property descriptions have been based on classical work in the visual and tactual sciences, which have tended to make distinctions that are also made in language—that is, conceptually distinct properties have distinct labels. The principal properties we have studied are called—in lay terminology—texture, hardness, weight, size, shape, and temperature. (Part motion and geometry-determined function are not considered here but were included in our initial set of properties and related EPs.) In fact, as we will indicate, each of these properties potentially comprises a group of subproperties.

We divide properties into two general groups: material and geometric. (In previous papers, we have used the terms ‘substance-related’ and ‘structural’, but we have adopted the current terms because of their use in materials science and robotics.) A material property is defined as a factor affecting the response of a given material to imposed stimuli and constraints, independent of the shape and size of a particular sample (Rosenthal and Asimow 1971). What we commonly call hardness is such a property; materials scientists measure it under more technical names such as elasticity and plasticity, by using the stress/strain profile (i.e., how a specific sample deforms in response to a load). The apparent temperature of an object is a property that is related to characteristics of material such as thermal conductivity.

Texture is also a material property. However, it should be noted that texture can depend on microgeometric characteristics of the material of an object. For example, if an object is fabricated so that its surface has small projections, then judgments of its surface texture may lead to calling it rough. Note too that the same object could be used for judgments of geometric properties. If people are asked to judge the spatial density of the microelements, they become objects in their own right, and a geometric property is being judged (see Lederman et al. 1986).

Properties related to the geometry of a particular material sample (i.e., an object) are sometimes called technological properties; we prefer the term geometric. Size and shape are geometric properties. For our purposes, shape refers to the macrogeometry of an object, that is, at a scale where contour changes are not treated as textures. We use the term hybrid to refer to properties that directly reflect both geometry and material. Weight (or, more properly, mass) is a hybrid that is jointly determined by geometry (size; hollow/filled) and material (density). The evidence suggests, however, that the contribution of geometry to perceived weight is substantially greater than is that of material (Ellis and Lederman 1993).

This description of object properties makes clear that the properties we have considered are relatively abstract categories that can potentially be instantiated by multiple variants. Texture, for example, can refer to roughness or slipperiness. There are a number of strain-related properties that we commonly call ‘hardness’, including compliance, elasticity, and brittleness. The number of shape metrics is virtually limitless. We have considered, for example, a nominal shape variable describing the two-dimensional or three-dimensional object envelope (e.g., round versus oval; sphere versus cube) and a shape-complexity variable (i.e., number of protrusions from the envelope). In computational vision a number of metrics have also been used, such as superquadrics (for application to robotic perception, see Bajcsy and Solina (1987)). Ultimately, we wish to decompose these relatively inclusive property descriptions.
However, the current property descriptors may constitute a basic-level categorization of haptically perceived properties, especially in that the properties tend to be distinguished by visible characteristics of the EPs used to extract them (see descriptions of EPs below).

2.3 EP component
Exploratory procedures are stereotypical patterns of contact and movement between an exploring effector and an object. The several exploratory procedures described in our previous work (see figure 1) are differentiated by their gross kinematic and dynamic properties, the geometry of the end-effector relative to the object, and the position of the effector and object in the workspace.

On the basis of our considerable corpus of hand-movement data, we have developed a set of visible parameters that are necessary and sufficient to differentiate among the EPs. We have observed that the values of these parameters tend to be consistent and stereotypic for a given EP over a wide range of objects (common and custom designed). In differentiating EPs, we consider what the characteristic value of the EP is with respect to each parameter, as indicated in table 1. In essence, a parameter captures some constraint inherent in an EP, a constraint on its execution that is mandated by the need to extract certain kinds of information.

The parameters are: (i) movement (whether the end-effector is static or dynamic), (ii) the direction of the principal force applied relative to the surface(s) of the object contacted by the end-effector (normal versus tangential), (iii) the region of the object that is contacted by the end-effector (surface, edge, or surfaces and edges—where usually occurs when the hand adopts a prehensile posture), and (iv) whether there are workspace constraints (i.e., support-surface requirements) on the position of end-effector and object. Note that the parameters are not independent; for example, if Movement = static, then Direction of Force = normal (because normal force is applied in order to maintain contact).

We make two points of clarification with regard to these parameters. One concerns the Region parameter, which takes on values of surface, edges, or surface-and-edges. We define a surface as a region of the object where contours do not change abruptly. An edge is a region of abrupt local change in contour, or alternatively a path along the object envelope where the contour changes are greatest. In some cases, surface and edge information occur in the same region; for example, on a sphere.

The second point concerns the workspace parameter, which indicates whether there is a constraint on the position of the hand and the object in the workspace. All exploratory procedures require that the object be stabilized against the force of the exploring effectors by opposing forces: the other hand, other surfaces of the exploring hand, or an external support such as a table. Unsupported Holding requires specifically that countergravitational force be exerted by the hand, without external support. It is this constraint that distinguishes Unsupported Holding from Enclosure.

<table>
<thead>
<tr>
<th>EP</th>
<th>Movement</th>
<th>Direction</th>
<th>Region</th>
<th>Workspace constraint?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Contact</td>
<td>static</td>
<td>normal</td>
<td>surface</td>
<td>no</td>
</tr>
<tr>
<td>Pressure</td>
<td>dynamic</td>
<td>normal</td>
<td>surface</td>
<td>no</td>
</tr>
<tr>
<td>Lateral Motion</td>
<td>dynamic</td>
<td>tangential</td>
<td>surface</td>
<td>no</td>
</tr>
<tr>
<td>Enclosure</td>
<td>static</td>
<td>normal</td>
<td>surface and edges</td>
<td>no</td>
</tr>
<tr>
<td>Contour Following</td>
<td>dynamic</td>
<td>tangential</td>
<td>edges</td>
<td>no</td>
</tr>
<tr>
<td>Unsupported Holding</td>
<td>static</td>
<td>normal</td>
<td>surface and edges</td>
<td>yes</td>
</tr>
</tbody>
</table>
Although in principle it is possible to balance an object in the palm without enclosing it, most instances of Unsupported Holding also involve Enclosure, in the form of grasping and/or molding (even for small objects, which tend to be cupped in the palm). (Later we will discuss the extent to which certain EPs intrinsically include others, as is the case with Unsupported Holding incorporating Enclosure.)

There are other parameters that would be informative as well. One is whether a prehensile hand shape is involved (Klatzky et al 1990); however, prehension is redundant with the surface-and-edges value of the Region parameter and therefore we do not include it. Two others that we have considered are the size of the hand surface contacting the object and the degree of force imposed by the hand. However, the values on these properties are likely to vary considerably with the properties of the particular object being explored, rather than being reasonably stable attributes of the exploratory pattern (Klatzky et al 1987b). Another potential parameter is whether movement occurs continuously in some direction or whether its direction varies (eg Contour Following is relatively continuous; Lateral Motion often occurs with back-and-forth motion). In the future, we intend to obtain kinematic and force measures to determine if there are systematic and object-independent differences between EPs on such parameters. The present parameter set is, however, sufficient to distinguish among the EPs at the level we have defined.

2.4 Sensorimotor component

The sensorimotor component underlies the execution of EPs and the extraction of associated property values. The mechanisms of this component are, at best, only partially understood. The neural coding that occurs during the given pattern of movement provides information about the property values of a given object. At the periphery of the haptic system are mechanoreceptors and thermal receptors that are driven by stimulation from contact and movement of the end-effector—that is, by the execution of EPs. This information must pass through successive levels up to and within the cortex, although very little is known about these subsequent processing stages.

Insights into the nature of certain property computations are made possible by several bodies of research encompassing psychophysical and/or neurophysiological approaches. The reader is referred to theoretical/empirical work, for example, by Taylor and Lederman (1975) and Connor et al (1990) on surface roughness; LaMotte and Srinivasan (1987; Srinivasan and LaMotte 1987), Loomis (1990), Phillips et al (1990), and Bankman et al (1990) on microgeometry; Turvey, Solomon, and associates (eg Solomon et al 1989) on the dynamic parameters underlying the spatial attributes of a wielded rod; and Sumino and Dubner (1981) and Keshalo (1984) on thermal coding. On the motor side, control of complex arm and hand movements like those involved in the EPs has been studied in areas 4 and 6 of the monkey cortex by Gentilucci and Rizzolatti (1990).

2.5 Links between object and property components

We now turn to the linkages between and within system components, beginning with object-to-property associations. Recall that objects are assumed to be associated with specific property values within the object component of the architecture; for example, 'cotton ball' would be associated with a low value of hardness. Such links are assumed to exist entirely within the object component. Here we consider associative links between the object and property components, reflecting the importance of a given property for categorizing a specific object. This link does not carry the specific property value, but rather the predictive worth of the property for identifying the object. For example, the property of surface roughness is particularly relevant to identifying a piece of sandpaper. In such cases the property is said to be 'diagnostic'
of the particular object; how much so is represented by the weight on an object-to-property link. Presumably, diagnosticity results when the value of some property for a particular object is extreme, relative to the distribution of property values across the universe of objects. Conjunct distributions may be even more important in determining diagnosticity, for example, a balloon is light for its size.

Diagnosticity can be considered in two respects, corresponding to top-down and bottom-up processing. Top-down diagnosticity refers to the extent to which a property is sought out when an object is recognized. Bottom-up diagnosticity refers to the extent to which a property actually contributes to object categorization when the stimulus is physically present and encoded by the sensory system. The links representing diagnosticity may therefore be unidirectional, with different weights on the link from object to property and from property to object.

We have empirically derived top-down diagnosticity weights for a set of 114 common objects at the basic level and 114 objects at the more specific subordinate level, by having subjects select diagnostic properties from a closed list (Lederman and Klatzky 1990b). Subjects were given the name of an object and were asked to select from the list, in the order they were thought of, the properties they would seek in trying to identify the object. The list included: hardness, texture, weight, temperature, size, shape, part (defined as a section of the object independent of any of its perceptual attributes such as shape), and motion of a part.

While the resulting norms indicate the relative weighting between objects and properties for specific objects, another issue is whether some properties are generally more diagnostic than others over some population of objects. Such general diagnosticity of a property may vary, depending on the level at which categorization occurs. When objects are categorized at the basic level, there is substantial evidence that geometric properties—particularly shape—are generally most diagnostic. Tversky and Hemenway (1984) found that when subjects were asked to list object properties, distinctions among objects at the basic level were largely based on their part structure. Material properties appear to be of greater importance in differentiating objects at more specific levels of categorization. Whereas Tversky and Hemenway dealt with top-down norms, Biederman (1987) has demonstrated the critical role of geometric information (ie volumetric descriptions) for visually driven object recognition at the basic level.

An important issue is whether the weights between objects and properties vary with the sensory modality being used. Aside from Paivio’s dual-code theory (1971), there has been relatively little consideration of how object representations might vary between perceptual/representational systems, and particularly between vision and touch. The work of Biederman and that of Tversky and Hemenway suggests that shape would be highly weighted for object identification at the basic level, regardless of modality. Yet this does not mean there is no effect of modality. We are not concerned here with distinct, modality-specific representations, but rather with the possibility that the weights between objects and properties have some degree of modality dependence. In support of this idea, consider that the extent to which a given object property is activated by the stimulus input—bottom-up—is likely to vary with the input modality. For example, thermal variations are far more available to haptics than to vision, as was noted by Gibson (1966). Accordingly, top-down hypotheses may depend on the input modality, so that thermal properties are more likely to be anticipated when an object is touched.

From our own studies, there is some reason to suspect that the context of perception by touch would increase the weight strengths between objects and material properties. To assess the cognitive importance of various properties, we devised a free-sorting task, in which subjects were asked to place similar objects together.
(Klatzky et al 1987a; Summers et al, in preparation). The objects were constructed by factorially manipulating properties; for example, in one study there were small and large spheres and cubes. This means that when objects were segregated by one property, they had to be aggregated on another (eg if small objects were treated as similar and sorted separately from large objects, small spheres would be lumped with small cubes). Properties that were used as the basis for segregation were, by definition, the most highly weighted. Our work with this paradigm indicated that property weightings varied substantially with the task context. When subjects sorted using vision and touch, or when they used touch alone but under instructions to form visual images of the objects, they tended to give shape the highest weight. In contrast, when subjects used touch without visual-imagery instructions, material properties became relatively important.

Our previously described norming study, in which subjects indicated which properties were diagnostic of object categories, is relevant to the issue of whether object–property weights depend on perceptual context. In that study, subjects were specifically told to think about identifying the object by touch. To determine the general importance of an object property, we assigned it a diagnosticity score, based on the frequency of it being selected over the pool of objects, weighted by the rank order in which it was listed.

The pool of objects was constructed so that at the subordinate level some were likely to be identified by texture, some by thermal properties, and so on. For example, subjects were asked how they would determine if a fork was further a dessert fork, a subordinate-level distinction for which the property of size is particularly important. To test predictions about exploration, the distribution of property ratings at the subordinate level was constrained by the experimenters’ selection of object categories. However, the same objects were also tested at the basic level (eg what makes an eating implement further a fork), and here property diagnosticity was free to vary, so that scores could be compared across properties. As might be expected, the diagnosticity scores confirmed the importance of shape at the basic level, even in this haptically biased context. Size and texture were also quite important at the basic level. (The finding for texture should not be taken as indicating a high weight for material properties, however, because texture regions may be used to isolate object parts.) There was also some indication that material properties become more important in categorization below the basic level, in that thermal properties tended to be rated as more diagnostic at the subordinate level than at the basic level of categorization.

2.6 Links between EPs and properties

Weighted links between the EP and property components represent the precision with which the EP delivers information about the given property. On this basis, there appears to be a readily specifiable mapping between the EPs in figure 1 and our coarse, or basic level of, property descriptions. We identified this mapping in a study where subjects performed a match-to-sample task with targeted object properties (Lederman and Klatzky 1987). On each trial they received a sample object and three potential matches. They were to pick the best match on the specified property, for example ‘hardness’. Our interest was in the hand movements that subjects used to explore the first object. Explicit scoring criteria were used to divide the exploratory period into a sequence of EPs.

As indicated in figure 1, we found that there was a tendency for certain EPs to dominate, depending on the targeted property. Subjects seem to select an EP spontaneously when a property is to be extracted. We proposed that this occurs because EPs differentially deliver information about properties.
Using a variant of the match-to-sample task, we empirically derived a set of weight values for EP-to-property links. Subjects in this condition were instructed to use a particular EP as well as to make judgments about a specified property. All EPs were paired with all properties. This experiment allowed us to distinguish between four levels of performance, based on the subject’s match-to-sample accuracy relative to chance. (Here, accuracy was assessed either by an objectively correct value or by adopting the consensus response from the first experiment as correct.) An EP can lead to chance-level performance in discriminations on the given property; it can be sufficient to produce performance above chance but not to provide the best performance level; it can be optimal (ie provide the best performance); or it can be necessary (ie be the only means of performing above chance). These outcomes were used to assign weights of 0–3, respectively, as shown in table 2.

In general, those EPs that were used spontaneously in the original match-to-sample task also tended to prove optimal in the constrained version. There is one exception, in that rather than the anticipated EP—Enclosure—Static Contact was found optimal for extracting global shape in the constrained task. However, subjects might have been able to mold the palm to the upper surface of the object in producing Static Contact, and the version of Enclosure we used was quite restrictive (the object was placed in the open palm, while the hand remained on the table, to preclude information from Unsupported Holding). Because other tasks consistently produce Enclosure in response to conditions promoting processing of shape information (eg Reed et al 1990), we feel that the link from Enclosure to global shape should be considered optimal (and the table so indicates).

Table 2 also shows, for each EP, a generality score. The entry is simply the sum of the nonzero entries for that EP and thus a count of the number of properties that an EP is at least sufficient to extract. Contour Following is found to be the most broadly sufficient, with Enclosure next. Lateral Motion and Pressure, in contrast, are relatively narrow in their applicability. However, the breadth of information provided by Contour Following must be pitted against its slowness. This can also be seen from the values of average durations (from Lederman and Klatzky 1987) of each EP, when subjects were judging the property(ies) for which the EP is optimal (as defined by table 2).

Table 2. EP-to-property weightings (adapted from Lederman and Klatzky 1990a), generality, and average duration for each EP.

<table>
<thead>
<tr>
<th>Property</th>
<th>EP</th>
<th>Lateral Motion</th>
<th>Pressure</th>
<th>Static Contact</th>
<th>Unsupported Holding</th>
<th>Enclosure</th>
<th>Contour Following</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texture</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hardness</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Temperature</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Weight</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Volume</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Global shape</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Exact shape</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Generality</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Duration/s</td>
<td>3.46</td>
<td>2.24</td>
<td>0.06</td>
<td>2.12</td>
<td>1.81</td>
<td>11.20</td>
<td></td>
</tr>
</tbody>
</table>
2.7 Links between exploratory procedures

These links reflect the extent to which two exploratory procedures can be executed at the same time, each still extracting the property for which it is optimal with at least enough precision to guarantee sufficiency (in the types of discrimination required in our experiments). For example, an object can be explored with Lateral Motion and Pressure at the same time, delivering information about texture and hardness. Another example is the pair Enclosure and Static Contact; here both can be instantiated by simply enclosing the object. In the first example, two distinct EPs were executed together; in the second example, executing one EP effectively executed the other simultaneously. We say, in either of these cases, that the two EPs are 'compatible'. Essentially, we say that two EPs are compatible if there is some manner of exploration that can be used to execute both at the same time without too great a loss of information.

We assume that compatibility exists between two EPs only if the constraints inherent in their parameter values can be satisfied at the same time—by some means of exploration. Recall that the parameter values pertain to physical aspects of exploration: whether the EP is static or dynamic, the direction of principal force, the object region explored, and position relative to the workspace. Note that our definition of compatibility therefore transcends that of simple motoric compatibility, since it also includes constraints imposed by the object and the workspace. It further requires that information content not be too greatly reduced, which we assume would occur if a constraint on one of the EPs were violated.

To determine compatibility, then, we must consider the individual parameters of table 1. Obviously, two EPs would be compatible if they had the same values on every parameter, but then the two EPs would not be differentiable by our analysis. Any pair of EPs must differ with respect to their values on at least one parameter. However, the two EPs might still become compatible if there were some means of exploration that resolved the incompatibility by satisfying the constraints inherent in the two distinct parameter values. For example, consider the parameter Region, which takes values surface, edges, and surface-and-edges. If one EP has the constraint that the surface be contacted and another has the more restrictive constraint that both surface and edges be contacted, the first constraint can be satisfied when the second EP is performed.

On the basis of our substantial body of data to date, which indicate when EPs are executed alone or in tandem with a large array of multidimensional objects, we have developed a set of five rules that specify when two different values on a given parameter can be reconciled. Each rule operates only within certain limiting conditions, corresponding to values of other parameters. These rules are given here, followed by brief rationales.

(i) With respect to the Movement parameter, the value static is satisfied by the value dynamic for any two EPs having identical values on the Direction of Force parameter.  
(Rationale: When sufficient force is applied tangentially or normally to produce movement, information that would be provided by static force in the direction of motion is also available.)

(ii) With respect to the Region parameter, the value surface or the value edges is satisfied by the value surface-and-edges for any two EPs having identical values on the Direction of Force parameter.  
(Rationale: If the object is contacted on homogeneous regions as well as lines of contour, the need to obtain information from either region by itself with the same direction of force is satisfied.)
(iii) With respect to the Region parameter, the value *surface* is satisfied by the value *edges* for any two EPs having identical values on the Direction of Force parameter.

(Rationale: If two EPs differ in that one typically explores the surface and the other the edge, then applying both EPs at the edge would generally provide sufficient, if nonoptimal, information about the surface of the object.)

(iv) With respect to the Workspace Constraints parameter, the value *no* is satisfied by the value *yes* for any two EPs.

(Rationale: Exploring in a particular position relative to the workspace satisfies EPs that require no particular position.)

(v) With respect to the Direction of Force parameter, values *normal* and *tangential* can be satisfied together for any two EPs having the value *surface* on the Region parameter.

(Rationale: It is possible to execute a force vector with both of these direction components to extract surface information, as we have frequently observed in our data. In contrast, if one EP typically explores at the edge, and hence uses tangential force, and the other exerts normal force, the edge of a compliant object would deform, grossly distorting the information.)

The five rules reconcile specific differences between parameter values. We now return to compatibility of the EPs themselves. Compatibility of two EPs requires that the constraints that are inherent in their parameter values be capable of being simultaneously satisfied. If we consider each parameter separately, compatibility exists (a) if the values match, or (b) if selection of one parameter value satisfies the other, as per rules (i)–(v) above, or (c) if both parameter values can be present simultaneously, as per rule (v).

We express such compatibilities or incompatibilities as plus or minus entries in a matrix of EP-to-EP weights, as shown in table 3. Consider, for example, the relation between Pressure and Unsupported Holding. Examination of their parameter values reveals mismatches on the Movement, Region, and Workspace parameters. These are reconciled by rules (i), (ii), and (iv), respectively, so that a + is entered. In contrast, Lateral Motion and Static Contact mismatch only on Movement and Direction of Force, but rule (i) cannot be used to reconcile the Movement mismatch; hence a − is entered in this cell.

### Table 3. Compatibility relations between EPs: + indicates compatible; − indicates incompatible.

<table>
<thead>
<tr>
<th>Pressure</th>
<th>Lateral Motion</th>
<th>Enclosure</th>
<th>Contour Following</th>
<th>Unsupported Holding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static Contact</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Pressure</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Lateral Motion</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Enclosure</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Contour Following</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

2.8 **Links between properties**

In addition to object-to-property, property-to-EP, and EP-to-EP weights, we have also considered weightings between properties themselves. In particular, weights might be used to instantiate the distinction between geometric and material properties, so that

---

(1) Order of application does not matter in practice. However, the rules should logically be applied so that (ii) and (iii) precede (v). Rule (v) cannot be used to justify (i), (ii), or (iii) because it allows for coexecution but not substitution of one force direction for another.
there would be a positive weight between properties of the same type. Such a property–property weighting scheme is motivated by evidence that properties of the same type tend to be treated together. For example, in the free-sorting task described above, subjects who explored objects by touch tended to find material properties more important, or 'salient', than geometric properties, whereas the reverse held for subjects exploring with vision or under visual-imagery instructions. Other data suggesting interdependence of properties of the same type come from an integration paradigm to be described further below. With planar objects (ie those having in two-dimensional shape with a third dimension of constant thickness), we have found evidence for substantial integration of texture and hardness, both material properties, and of size and shape, both geometric properties, with less coprocessing across the two types of properties. This pattern reflects information availability: geometric information in a planar object is carried at the edge, whereas material properties are optimally available in the homogeneous planar region.

Finally, our norms for the diagnosticity of object properties, described above, support links among material and among geometric properties. We derived a measure of interproperty similarity from the frequencies with which pairs of properties were held to be diagnostic of the same object. A cluster analysis on the similarities found shape and size, both geometric properties, to cluster first. A second cluster then emerged linking the material properties of texture and hardness. Weight then appeared first joining the material and then the geometric properties, as seems appropriate given its hybrid status. Temperature, however, clustered relatively late, suggesting that it is not jointly diagnostic with the other material properties.

3 The selection–extraction loop and the process of EP selection
In keeping with the interactive activation approach (eg McClelland and Rumelhart 1981; Seidenberg and McClelland 1989), we view haptic identification as a parallel, interactive process. As shown in figure 4, the process is assumed to proceed in a sequence of selection–extraction loops: at each step, an EP is selected and executed, potentially along with other compatible EPs. This leads to the extraction of information about associated object properties, the quality and/or quantity of that information depending on the EP-to-property weights. Over the sequence of such loops, an object representation is constructed and used as a probe for matching with object representations. This search process proceeds until a match criterion is met, at which point the object is said to be recognized.

```
Input object
     ↓
Select and execute compatible EPs
   Extract property information
     ↓
Build an object representation
     ↓
Compare representation with memory and output decision when possible
```

Figure 4. The EP selection–extraction loop.
The main goal during the selection–extraction loop is to select an EP for execution. This goal must be met under a number of competing constraints. The EPs considerably constrain what information is available about a touched object, hence the EP that is selected must provide information about desired object properties, if any. These properties can be dictated either by object-specific hypotheses (I expect an egg, which has a certain shape) or by more general biases that reflect the pool of possible objects and their diagnostic dimensions (shape is generally useful in basic-level categorization and may be so here). Object properties may become desirable, too, because of initial data-driven processing. For example, initial information suggesting an unusual texture value may drive the system to execute Lateral Motion, which provides greater texture discrimination.

Although there may be particular interest in some property, a general aim is to learn about as many properties as possible in the shortest time. There are two ways to accomplish this. One is to seek breadth of application (within the level of precision demanded by the task), that is, to execute EPs that are broadly sufficient to deliver information about many properties. Contour Following and Enclosure provide at least coarse information about multiple object properties, suggesting by this strategy that these be favored for selection. An alternative approach is to execute as many EPs as possible at the same time. Here one must consider another set of constraints which concern EP compatibility. Enclosure, Unsupported Holding, and Static Contact constitute a group of compatible EPs that might be simultaneously executed under this strategy, for example.

Some constraints may lead to intrinsic biases governing the use of EPs. Table 2 indicates that Contour Following tends to be slow, at least with unfamiliar objects. Table 3 indicates that it is also motorically exclusive (ie incompatible with most other EPs), and it is relatively complex, typically involving rotation and stabilization of the object with one hand while exploring edges with the other. There is also substantial evidence from the shape-matching literature, and our own studies (Klatzky et al 1991a; Lederman et al 1991a), that the level of precision that can be obtained from Contour Following is still not sufficient to make shape discriminations that are obvious to the eye. These factors would lead to a bias against Contour Following. On the other hand, there might be an intrinsic bias in favor of Enclosure, owing to its general sufficiency for encoding properties, or Pressure, for its ability to combine with other EPs and the brief duration of execution.

These constraints and biases can potentially be represented by the weights between and within components of our system. Object-specific associations and property expectations for specific objects can be represented by associative weighting within the object component and by links between the object and property components. The constraints of breadth and precision are represented by the EP-to-property links in table 2; and the constraints inherent in the EP parameters determine compatibility, which is the basis for EP-to-EP links in table 3. Intrinsic preferences based, for example, on motoric effort or the duration of execution, can be represented by item-specific bias terms.

4 Constraint satisfaction
Collectively, these constraints operate to determine the selection of an EP, given an object and some a priori expectations. From a connectionist perspective, the selection process can be viewed as a constraint satisfaction algorithm, in which the weightings act as constraints to be relaxed until some elements are maximally activated (see Rumelhart et al 1986). A system with symmetric weights and asynchronous updating minimizes a cost function over the set of constraints (weights), ultimately selecting an optimum state (activation pattern) over the connected units (Hopfield 1982). Here we
consider constraint satisfaction as a means of selecting the next procedure in a sequence [see Kintsch (1988) for a similar use]. The weights are theoretically and/or empirically derived relations among EPs and properties (and, potentially, objects), and we are interested in selecting the EPs with the highest terminal activation. The emergence of a stable activation state, over progressive relaxation of the system, becomes a prediction as to which EP(s) will be executed in some exploratory condition.

In order to investigate the implications of EP-to-property associations and the compatibility between EPs, we implemented the weights presented in tables 2 and 3 as a constraint-satisfaction system. Nodes in the system represented the EPs and properties; this is equivalent to looking at a single generic object. No bias terms were incorporated.

Initially, the system is started by putting a small external input into one EP. (The particular node selected for this purpose has essentially no effect on the behavior of the system, given the small value involved.) The system can also be started with one or more property values ‘clamped’, which sets their activation values to 1 (the maximum possible) throughout. The constraints are satisfied iteratively and asynchronously, by means of an algorithm of McClelland and Rumelhart (1986). At each step, the net input to a randomly selected node is calculated, as is typical, by summing weights x input activations from the other nodes. The activation of the node is then incremented by the product of the input and the distance between the current activation and +1 (if the input is positive) or -1 (if the input is negative). This product is scaled by an excitation parameter, and from it is subtracted the effect of a decay parameter, which pulls the activation back toward zero in proportion to its distance from zero. After 52 updates (an average of 4 per node), the system is checked to determine whether the activation values are changing. If the average activation over the nodes has changed less than a criterion (here, 0.0001), the system is considered stable. At this point, the system goodness (a measure of the extent of constraint satisfaction) is calculated by summing, across all pairs of nodes, the product of their activations.

Table 4. EP activation-level after constraint satisfaction when properties have been activated externally. Also shown is the goodness level at the point of relaxation, and the time (in multiples of 52 updates) to relax. The highest activation for a given property is shown with an asterisk.

<table>
<thead>
<tr>
<th>Activated property</th>
<th>Activated EP</th>
<th>Time to relax</th>
<th>Goodness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lateral Motion</td>
<td>Pressure</td>
<td>Contour Following</td>
</tr>
<tr>
<td><strong>Full weight matrix</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Texture</td>
<td>0.45</td>
<td>0.58</td>
<td>0.48</td>
</tr>
<tr>
<td>Hardness</td>
<td>0.40</td>
<td>0.60</td>
<td>0.47</td>
</tr>
<tr>
<td>Exact shape</td>
<td>0.36</td>
<td>0.54</td>
<td>0.58</td>
</tr>
<tr>
<td>Global shape</td>
<td>0.33</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>Size</td>
<td>0.33</td>
<td>0.55</td>
<td>0.46</td>
</tr>
<tr>
<td>Weight</td>
<td>0.39</td>
<td>0.58</td>
<td>0.46</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.53*</td>
<td>-0.34</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>-0.36</td>
<td>0.52*</td>
<td>-0.51</td>
</tr>
<tr>
<td>Exact shape</td>
<td>0.39</td>
<td>-0.38</td>
<td>0.62*</td>
</tr>
<tr>
<td>Global shape</td>
<td>-0.39</td>
<td>0.37</td>
<td>-0.50</td>
</tr>
<tr>
<td>Size</td>
<td>-0.39</td>
<td>0.37</td>
<td>-0.50</td>
</tr>
<tr>
<td>Weight</td>
<td>-0.40</td>
<td>0.38</td>
<td>-0.50</td>
</tr>
<tr>
<td>Temperature</td>
<td>-0.40</td>
<td>0.37</td>
<td>-0.50</td>
</tr>
</tbody>
</table>
activation values and the weight between them. Also added to goodness is the activation of any clamped property.

Table 4 gives the results of this model with two sets of weights. Results are shown when each property is clamped. In one case, the full-weight matrix indicated by tables 2 and 3 is used. Because the relative scaling between the EP-property and EP-EP sections of the matrix is arbitrary, we have set the means of the former to 1 and the latter to zero (cf the means of the entries in tables 2 and 3, which are 0.86 and 0.07, respectively).

We use the second set of weights to represent the situation which might occur after exploration has progressed for some period of time. The coarse information that could be extracted by sufficient but not optimal patterns of exploration is likely to have been exhausted at this point. If this information does not allow for adequate task performance, further exploration with optimal EPs must be used to extract information about the necessary object properties. Hence in this version of the model we use only optimal weights in the EP-Property section of the matrix. (No further rescaling of means is done.)

The upper portion of table 4 describes the results obtained with the use of the full weight matrix, in which positive weights exist between EPs and properties for which they are merely sufficient. The stable state of this system is one in which the highest activation level accrues to the Enclosure EP, regardless of which property is clamped. This is consistent with the assumptions in the model that Enclosure is broadly sufficient and tends to be compatible with other EPs. The EP that is next most active is frequently Unsupported Holding. Because these two EPs are compatible, they could be selected in the same step and executed together.\(^{(2)}\)

The lower portion of table 4 indicates that the system behaves quite differently when only optimal EP-property pairings are weighted above zero. Now, biasing a property generally results in selection of the optimal EP. This result is consistent with the idea that the emergence of specialized EPs reflects a demand for increased precision in property information, which can only be obtained by optimal exploration.

5 Empirical studies of EP selection

These outcomes agree with data we have collected (Lederman and Klatzky 1990b) on exploration during object classification, in which a two-stage exploratory sequence was observed. On each trial, the subject was given an object, prefaced by a question of the form, “Is this X further a Y?” For example, they might be asked, “Is this fork further a dessert fork?” The object sets we used were those for which we had ascertained diagnostic attributes, as described above.

\(^{(2)}\) The system behaves in much the same way when the entries of tables 2 and 3 are not rescaled to have means of 1 and 0. The one exception is that when the full-weight matrix is used (ie including nonoptimal EP-to-property weights) and exact shape is clamped, the system activates not Enclosure, but Contour Following—the optimal EP. In this case goodness is relatively low and the system is slow to relax. When a negative bias term is used for Contour Following, a reasonable step in terms of its longer duration, the system reverts to selecting Enclosure when exact shape is clamped. In another version, we have added to the original-weight matrix a portion reflecting property-to-property relations, where a weight of 1 is given between two properties if they are both geometric and are both related to material substance. (For these purposes the property weight is treated as geometric.) In this case the system produces results like those in table 4. Thus in a number of variations in which we use the full-weight matrix, Enclosure best satisfies the constraints.
We analyzed subjects' exploration as a sequence of EPs. This analysis revealed two distinct stages, as shown in figure 5. In the first stage, Enclosure of the body of the object occurred. It was very frequently accompanied or followed by Unsupported Holding. This pair of EPs constitutes what we call a 'grasp/lift' routine. Subsequently the other EPs emerged. At this point, the occurrence of a particular EP could be predicted from the properties that were most critical to identification—that is, from the most diagnostic attributes of the object. For example, if hardness was particularly important in identifying an object, as it would be for determining that a noodle was further a cooked noodle, one would expect to see use of the Pressure EP after the initial grasp/lift routine. Such predictions were confirmed, supporting our assumption that the second phase of exploration is a more directly targeted one that uses EPs capable of extracting desired properties with greatest precision (ie optimal EPs).

In short, our data on exploration during object identification support the view of exploratory control as a constraint-satisfaction process, in which constraints change over the course of exploration. The critical constraints are connections between EPs and properties, reflecting the degree to which exploration delivers information; and connections between EPs themselves, reflecting their compatibility with respect to parameters of the hand, object, workspace, and outcome. Early on, when little is known about an object, the sufficiency of EPs predicts exploration. Later, when more precision is required, constraints appear to be tuned so that merely sufficient EPs are excluded and optimal EPs are selected.

In its current form the network of constraints does not incorporate multiple objects. However, as we have indicated, a pool of objects could be connected to the property component of the matrix by weights representing diagnostic properties. If the distribution of diagnostic properties in the pool reflected real-world constraints, shape should be most diagnostic at the basic level, which would tend to activate Enclosure and Contour Following when no particular object was expected and objects were simply activated at random. Activation of a particular object would in turn activate most strongly its particular diagnostic attribute(s), which we directly express in the present form of the model by clamping property nodes. We could also add to

![Figure 5. Cumulative percent occurrence of each EP as a function of position in the exploratory sequence (from Lederman and Klatzky 1990b). The solid lines indicate the grasp/lift combination. Static Contact is not included because it occurred very infrequently.](image-url)
the network nodes representing pairs of compatible EPs, with their connections to properties adjusted to reflect any reduction in precision when one EP is combined with another. This mechanism would require further empirical data, however.

6 The selection–extraction loop and the extraction of object properties
Our model deals not only with how EPs are selected, but with how their selection determines the precision with which information is extracted about a felt object. In particular, the EP-to-property links in table 2 directly indicate the extent to which the properties of an object are apprehended or learned through a particular means of exploration. A straightforward assumption is that each execution of an EP serves to sample properties of the explored object, and the amount learned about a given property from the sample (i.e., the precision of discrimination that the sample affords) is directly related to the weight between the EP and the property. This assumption has several implications. First, learning about a property should be fastest when exploration uses the optimal EP, because fewer samples are then required to extract a given amount of information. Furthermore, if exploration consists of an EP that is optimal for a given property, incidental learning about other properties will be limited to those properties for which the EP is sufficient. Once an EP is selected for exploration, it is assumed to block incompatible EPs from being executed at that time, forcing sequential output when incompatible EPs are selected. If the eliminated EPs are necessary for a given property, no learning about that property will ensue.

These predictions make clear that the EPs serve as a ‘gateway’ to object apprehension. Their role in determining access to object properties has been investigated in a series of studies in which categorization tasks were used. For use in these tasks, we designed sets of objects that factorially varied in a number of properties, for example, hardness, shape, and texture. One or more of these properties were then used to define a partitioning of the objects into categories. For example, the hardest objects might be called category A, the next hardest B, and so on. On each of a series of trials an object was presented, and the subject indicated to which category it belonged. The time required for the classification (from object contact, measured by a sensor, to response) was recorded.

This categorization task required that the subject extract information about one or more diagnostic properties of the presented object. The speed of identification was expected to increase when EPs having high weights for those properties were executed early in exploration. In support of this idea, we found that subjects who repeatedly classified objects showed a decrease in response time, which was correlated with an increasing probability that the optimal EP was executed (Klatzky et al 1989). (Note that other factors undoubtedly contributed to the decrease in reaction time, which generally agreed with the power law of practice.)

If more than one property is diagnostic of an object class, classification should be speeded (redundancy gain, e.g., Ashby and Townsend 1986; Garner 1974). This effect should be limited, however, by the extent to which the EPs that are highly weighted for the redundant properties can be executed in tandem. We found empirically that there were limits on redundancy gain, reflecting the compatibility of EPs (Klatzky et al 1989; Reed et al 1990). For example, if texture and hardness redundantly defined an object category (e.g., all As were smooth and soft), categorization time was faster than if just one property was relevant. Texture and hardness could be extracted together, because the corresponding EPs—Lateral Motion and Pressure—are compatible. However, if two-dimensional shape was added as a category cue, there was no further speedup, nor did exploration change substantially from the case in which texture and hardness were the only category cues. This is predicted from our
assumption that neither of the shape-extraction procedures, Contour Following and Enclosure, is compatible with both Pressure and Lateral Motion.

When an object is explored by a single EP that is most strongly weighted for some target property, knowledge about any other property should accrue in direct relation to the weight between it and that EP. We assessed such incidental knowledge acquisition with a 'withdrawal' paradigm, in which one property of an object, $P_1$, was targeted for classification while another, $P_2$, implicitly varied in a redundant fashion. Targeting of $P_1$ was done by explicitly defining the categories in terms of its levels; no mention was made of the correlation between $P_1$ and $P_2$ values. Subjects then generally chose to explore with the optimal EP for $P_1$. After performance stabilized, the value of the implicit property, $P_2$, was changed from redundant with $P_1$ to constant over the set of objects, without informing the subject. We refer to this as 'the withdrawal of property $P_2$'. For example, some subjects began classifying by texture, with hardness varying as a redundant cue. They tended to adopt Lateral Motion as the principal means of exploration. At the point of withdrawal, a new set of objects was introduced with the same texture variations as before, but having a constant level of hardness. Any increase in response time at the point of withdrawal was used as a measure of knowledge about the implicit property ($P_2$). We found that the effect of $P_2$ withdrawal increased with the weighting between $P_2$ and the EP that was optimal for $P_1$ (Klatzky et al 1989). Continuing the above example, because Lateral Motion is sufficient to extract hardness, a large effect of withdrawing hardness was expected—and obtained.

When we used the withdrawal paradigm with planar objects, we found large and symmetric withdrawal effects between texture and hardness. A similar pattern was found for planar shape and size. In contrast, when either texture or hardness was targeted and shape was withdrawn, there were only minimal effects. Similarly, there was little effect on reaction time when size was the basis for categorization and texture or hardness was withdrawn. These results suggest substantial coprocessing between material properties and between geometric properties with planar objects, with relatively little interaction between the material and geometric properties. However, in another study in which we used fully three-dimensional ellipsoids varying in curvature (as determined by the ratio between the axes) and punctate texture (as determined by the size of sharp projections from the surface), substantial withdrawal effects were found between texture and shape (Lederman et al 1993). This is to be expected, because shape and texture can both be adequately extracted under these conditions with Enclosure.

When no property is targeted by instruction or by how object categories are defined, knowledge about a haptically explored object should depend on which EPs are spontaneously executed. This is the case in the free-sorting task described previously, in which objects are placed into bins according to similarity. (The task is unspeeded, and reaction time is not recorded.) The execution of an EP should reflect, in this unbiased case, the motoric and/or processing costs associated with it. Since Contour Following imposes considerable cost in terms of long duration and low accuracy, exact shape should be relatively difficult to process. In accordance with this prediction, the shape of the two-dimensional object envelope was given relatively low weight by subjects who aggregated objects into self-defined classes. Only when shape was targeted, by adding vision or stressing visual imagery, was it primarily used to construct object classes (Klatzky et al 1987a). In a version of the task where three-dimensional objects in the shapes of spheres and cubes were used, EPs other than Contour Following (ie Enclosure) were clearly sufficient for the easy shape discrimination, and shape was accordingly given a high weight by all groups in sorting (Summers et al, in preparation).
7 Extension of the model to vision/haptics interactions

Vision can potentially be incorporated into our model as an additional EP. The use of vision, along with haptic EPs should then be governed by the same factors that have been used to make previous predictions, including its sufficiency to extract information about a given property, its compatibility with other EPs, and intrinsic properties such as the duration necessary for encoding at a given level of precision.

With respect to sufficiency, vision should have high weights for geometric properties—size and shape—and relatively low weights for material properties such as temperature, weight, and roughness. Gibson (1966) suggested that temperature was virtually exclusive to touch. The weight of a static object would also be difficult to judge by sight alone. It appears that sight of active lifting conveys weight information (Bingham 1987; Runeson and Frykholm 1981), but minimally so for objects that are easily lifted (ie within the weight range of most hand-held objects). Vision can be used to judge roughness (eg Lederman et al 1986), but Heller (1989) found that it was inadequate for discrimination of relatively fine textures.

These various findings indicate that vision will be relatively weakly connected to material dimensions. We have found further support for this idea from a task where subjects judged named objects with regard to a targeted property (eg which is heavier, a baseball or a can of soup?). People tended to imagine themselves looking at the objects if size and shape were judged, whereas they often imagined themselves touching the objects if material properties were judged (Klatzky et al 1991b).

With respect to compatibility, the extraction and use of visual information would appear to be compatible with any manual exploratory activity. Finally, in terms of intrinsic biases, vision should have a substantial advantage in processing duration, leading to a positive bias toward its selection. Indeed, we observed very little haptic exploration in studies in which objects were freely sorted with vision and touch (Klatzky et al 1987a; Summers et al, in preparation).

Taking these factors together, we would expect little use of touch for judging geometric properties of objects. For material properties, the probability of using haptics should be greater, the less the sufficiency of vision to extract the desired information. This would depend both on the nature of the property and on the level of precision required.

In a task similar to our property-judgment paradigm but with real objects, these predictions were generally confirmed (Klatzky et al, in press). Subjects used vision virtually exclusively for difficult judgments involving size and shape and for very easy judgments on material properties as well. When making difficult judgments about material properties, people showed a strong tendency to reach out and touch the objects, the more so the more difficult the task had been found to be when they used vision alone.

8 Applications to development of dextrous robot-hand systems

For tasks that must be performed in unstructured environments such as outer space, the deep sea, and certain industrial workspaces, or for those tasks in which a human cannot be involved directly (eg disposal of radioactive waste), flexible action by autonomous or teleoperated robotic systems equipped with sensitive dextrous hands is essential. The continual extraction and updating of perceptual information is necessary not only for object identification, but also for stably grasping, lifting, and subsequently manipulating an object, alone or in conjunction with others. Often visual information is inadequate owing to low levels of illumination or to object occlusion by environmental or task-required features (including the robotic effector itself). Haptic sensing must be incorporated into the effector system, not only to compensate for
poor visual information but in its own right [see also Westling and Johansson (1984) for the contribution of cutaneous information to grasping].

The use of haptic sensing has been advocated within the robotics field by Bajcsy (eg Bajcsy 1985; Bajcsy et al 1989; Ulrich et al 1987), who has pioneered the development of an emerging field, ‘exploratory robotics’. Our documentation of EPs and the associations of EPs with object properties has been used to develop a number of robotic systems with haptic exploratory capabilities. Stansfield (1988) implemented an object-recognition system that used haptics as the principal exploratory modality. In this system a single ‘finger’ equipped with sensors for force and position was used, along with a local gray-scale array sensor. It directly modeled versions of Contour Following, Pressure, and local Enclosure, which rendered it capable of identifying a range of objects both at the basic and at the subordinate levels of categorization. The scope of the treated objects and the level of discrimination required went well beyond earlier robotic object-recognition systems, which typically modeled only the geometric attributes of objects. Koutsou (1988) subsequently extended Stansfield’s algorithms to the case of a two-fingered gripper, and still more recently Allen and Michelman (1990) have implemented a system for three-dimensional shape recovery and object recognition in which haptic exploration was used to complement vision. It incorporates a version of the Enclosure EP to obtain gross shape, and two types of more specialized Contour Following procedures which are used to test surface planarity and to extract detailed edge information.

Our approach suggests a paradigm for selecting the patterns of robotic exploration for specific tasks in a way that optimizes performance. Following the implementation of a set of robotic EPs, our constraint-satisfaction approach would be used to guide the selection-extraction sequence. The EP-to-property weights would be determined by experimental tests to indicate whether a given EP was necessary, optimal, or sufficient for a given property. The EP-to-EP weights would be determined by an analysis of parameters related to the hardware configuration, the workspace, and the level of discrimination required by the task. The level of analysis that we have used to parameterize EPs makes it possible to consider the same variables in machine systems. The obvious differences between various effector systems do not preclude this possibility.

Note that the application of such work on biological haptic sensing does not call for the blind adoption of an anthropomorphic approach to robotics. Rather, it offers new conceptualizations and scientific methodologies, along with specific empirical findings about how living haptic systems approach tasks that the roboticist has not yet solved. In other sources we consider in more detail how the study of human haptics can potentially contribute to the development of sensor-based robotic systems (Lederman et al 1992).

Acknowledgments. We wish to acknowledge the support from several agencies at various phases of this work: the Information Technology Research Centre, Institute for Robotics and Intelligent Systems, Manufacturing Research Corporation of Ontario, National Science Foundation, Natural Sciences and Engineering Research Council of Canada, and Office of Naval Research. We thank Ruzena Bajcsy for her energy and faith in applying our work to robotics. We also thank our collaborators and coauthors in this research program.

References


Srinivasan M, LaMotte R H. 1987 “Tactile discrimination of shape: Responses of slowly and rapidly adapting mechanoreceptive afferents to a step indented into the monkey fingerpad” Journal of Neuroscience 7 1682–1697

Stansfield S. 1988 “A robotic perceptual system utilizing passive vision and active touch” International Journal of Robotics Research 7 138–161


Summers C, Lederman S J, Klatzky R L. “Salience of three-dimensional object properties under haptic exploration with and without vision” (in preparation)


