Extracting object properties through haptic exploration

Susan J. Lederman *, a, Roberta L. Klatzky b

a Dept. of Psychology, Queen's University, Kingston, Ontario, Canada K7L 3N6
b Dept. of Psychology, University of California, Santa Barbara, CA 93106, USA

This paper reviews some of our recent research on haptic exploration, perception and recognition of multidimensional objects. We begin by considering the nature of manual exploration in terms of the characteristics of various exploratory procedures (EPs) or stereotypical patterns of hand movements. Next, we explore their consequences for the sequence of EPs selected, for the relative cognitive salience of material versus geometric properties, and for dimensional integration. Finally, we discuss several applications of our research programme to the development of tangible graphics displays for the blind, autonomous and teleoperated haptic robotic systems, and food evaluation in the food industry.

1. Introduction

Some years ago now, we demonstrated that people are very good at recognizing common objects on the basis of touch alone. Klatzky et al. (1985) found that blindfolded adults were able to identify each of 100 common objects (e.g., hammer) with almost perfect accuracy and typically within only a few seconds.

As a result of this initial study, we have developed a systematic research programme to uncover the bases of this impressive performance. Our programme focusses on the haptic perception and recognition of multidimensional objects. By the term 'haptic', we are referring to an information-processing perceptual system that uses inputs from receptors embedded in the skin, as well as in muscles, tendons and joints (Loomis and Lederman, 1986). The varying sensory input patterns created as people move their hands over an object during perceptual exploration and manipulation tell us about its properties – for example, that it is smooth, hard, cold, round, and weighs very little. Subjects in the study above mentioned these and other such properties when asked to explain the bases of their object recognition.

* Corresponding author. Fax: (0613) 545 2499, E-mail: susan.lederman@queensu.ca
judgments, thus emphasizing the value of multidimensional as well as three-dimensional sources of information.

2. Manual exploration: Exploratory Procedures and associated property extraction

We began to suspect that the explorer's hand movements might provide the secret to the success with which touch alone can identify common objects. Accordingly, Lederman and Klatsky (1987) examined what people actually did with their hands when asked to make haptic evaluations of various object properties. More specifically, on each trial subjects were initially shown a multidimensional object (the standard stimulus), and asked to examine a specific property, e.g., its hardness. Next, they were presented in sequence with a set of three other multidimensional objects (the comparison stimuli). Subjects had to select the one comparison object that best matched the standard in terms of the named dimension (hardness). Hand movements were videotaped and subsequently analyzed.

One thing was immediately obvious. Although subjects were usually unaware of what they did with their hands, the movements themselves were both purposive and systematic. Subjects performed highly stereotypical movement patterns, which possessed both necessary and typical features. These patterns, which we call 'Exploratory Procedures' or 'EP's', are shown in Fig. 1 as they have been typically observed. We also found that during free manual exploration for a targeted property, subjects tended to select specific EPs (also shown in Fig. 1).

Thus, Lateral Motion, a repetitive and lateral rubbing motion, was most often associated with texture. Pressure, or opposing forces applied normally to a surface or a torque about some object axis, was used to extract information about hardness. Static Contact, stationary contact on a surface without molding, was used primarily to determine thermal properties. Unsupported Holding, used to lift an object away from a supporting surface, was used to extract information about weight. Enclosure, involving dynamic molding of the palm and/or finger(s) to the contours of an object, was used to extract both volumetric cues and envelope information about an object's shape. Contour Following, or dynamic edge following, was used to obtain the more precise spatial details concerning an object's shape. Other EPs shown here concern the extraction of information pertaining to the motion of an object part and function, as determined by the object's structure.

At the level of analysis we found to be most informative with respect to associations between EPs and object properties, neither the particular end effector(s) nor the area of skin contacted was relevant. We return to the significance of this point in the Applications section at the end of the paper.

Fig. 1. Exploratory Procedures and associated properties (reprinted with permission from Lederman, 1991, as revised from Lederman and Klatsky, 1987).

An additional experiment by Lederman and Klatzky (1987: Expt. 2) provided us with important information concerning the relative precision and speed with which each EP extracts various properties; it also enabled us to assess an EP’s breadth of sufficiency in simultaneously extracting multiple properties. We used the same methodology as before in this experiment, with one major change. Rather than permitting free exploration, subjects were required to perform a property-matching task on each trial, by executing a specified EP in conjunction with a designated property (e.g., match for weight using Lateral Motion). Across trials, all EPs and property-matching instructions were paired.

Accuracy scores (and response times, to break accuracy ties) were first used to compare EPs in terms of their relative precision in extracting a designated property. The results are presented in terms of an EP–Property weight matrix (Table 1a). A cell is assigned a ‘0’ when the EP cannot extract the property with above-chance accuracy. It is given a ‘1’ when the EP can perform the task with above-chance accuracy: the EP is said to be ‘sufficient’ for extracting that property. It is assigned a ‘2’ when the EP can perform the task better than any other sufficient EP; in this case, the EP is further considered ‘optimal’ for extracting the designated property. Finally, the cell is given a ‘3’ when the EP is shown to be ‘necessary’ for extracting the designated property, as Contour Following is for precise shape.

In Table 1b, we show the average duration (in seconds) for each EP, as determined from the first experiment, which used free exploration. Clearly, when full contour information is required, Contour Following is very much slower than the other EPs. Such temporal differences in exploration might bias a haptic observer generally against using Contour Following.

Table 1b
Average duration and breadth of sufficiency for each EP (from Klatzky and Lederman, 1991, as adapted from Lederman and Klatzky, 1990a)

<table>
<thead>
<tr>
<th>Duration (s)</th>
<th>Breadth of sufficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lateral Motion</td>
<td>3.46</td>
</tr>
<tr>
<td>Pressure</td>
<td>2.24</td>
</tr>
<tr>
<td>Static Contact</td>
<td>0.06</td>
</tr>
<tr>
<td>Unsupported Holding</td>
<td>2.12</td>
</tr>
<tr>
<td>Enclosure</td>
<td>1.81</td>
</tr>
<tr>
<td>Contour Following</td>
<td>11.20</td>
</tr>
</tbody>
</table>

We also show the relative breadth of sufficiency for each EP, based on the results of the constrained-exploration experiment (Table 1a). The numeric entry for each EP represents the number of non-zero cells in the appropriate row, that is, the number of properties for which that EP was sufficient. Thus, Lateral Motion and Pressure are sufficient for extracting relatively few properties, whereas Enclosure (grasp) and Contour Following are both broadly sufficient procedures, and each can provide fairly coarse information about almost all of the object properties shown here.

Finally, we have also considered a somewhat different characteristic of our exploratory procedures, which further influences haptic perception and recognition. This relates to the extent to which EPs are compatible with one another, that is, the extent to which they may be performed in tandem or very close in time. To determine EP–EP compatibility, Klatzky and Lederman (1993) have determined four parameters to be necessary and sufficient for differentiating EPs from one another. These include movement (static: dynamic), direction of force applied (normal: tangential), the object region explored (edges and/or central surface), and position relative to the workspace (whether or not the object must be removed from a supporting surface). Because the parameters pertain not only to aspects of movement but also to the geometry of the object and the object/workspace relation, our definition of compatibility extends beyond the notion of simple motoric compatibility.

We assume that compatibility exists between a pair of EPs only if the constraints inherent in their parameter values can be satisfied simultaneously by means of exploration. Any pair of EPs must differ on one parameter or we could not distinguish them. However, there may be some form of exploration that is still able to resolve the incompatibility by satisfying the constraints inherent in the two different parameter values. For example, if one EP requires exploration of surfaces (e.g., Static Contact), while another EP requires that both surfaces and edges be contacted (e.g., Enclosure), the first...
Table 2
Compatibility relations between Exploratory Procedures (+ means compatible; − means incompatible) (from Klatsky and Lederman, 1993)

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

will be satisfied when the more restrictive second EP is performed. From a substantial body of data that documents when EPs are performed alone or together on a large set of multifaceted objects, we developed a set of rules that indicate when two different values on a parameter may be satisfied (for details, see original paper). Based on these rules, we have determined a matrix of EP-EP compatibilities that is presented in Table 2 in terms of binary values. A ‘+’ indicates compatibility; a ‘−’ indicates incompatibility.

4. Modeling haptic object identification and manual exploration

Briefly, we treat haptic object identification as a parallel interactive process (for further details, see Klatsky and Lederman, 1993). We assume the process proceeds in a sequence of EP selection/property extraction loops. In each loop, the haptic system selects and executes an EP, possibly with others with which it is compatible. This provides information about various properties, the precision and amount of this information depending upon the EP-property strengths (Table 1a). Over a sequence of these steps, the system builds a representation and uses it as a probe for matching with existing object representations in memory. This process continues until some match criterion is achieved. At this point, the object is said to be recognized.

The primary goal of the selection-extraction loop above is to select an EP for execution. However, the selection will be subject to a number of competing constraints related to the relative precision and breadth of sufficiency with which EPs may extract properties (Table 1) and EP-EP compatibility considerations (Table 2). We have also noted that EP selection may be further biased by the length of time it usually takes to perform them; for example, there may be an intrinsic bias to avoid Contour Following because it can be so slow. From a connectionist perspective, we have treated the selection process as a constraint satisfaction algorithm in which the EP-prop-

5. Empirical implications for manual exploration and haptic processing of object properties

As EPs constrain the type and quality of information available, they must be selected with the particular goals of the task in mind. One might wish to learn quickly as much as possible about an object, in which case an Enclosure might be selected; this EP is both broadly sufficient and fast (as compared to Contour Following, which is broadly sufficient but relatively slow). Alternatively, the observer might immediately test some hypothesis about a particular property (e.g., Lederman and Klatsky, 1987). To explore these issues, we have investigated whether the sequence of hand movements is dictated by property diagnosticity (i.e., the extent to which a specific property can be used to identify an object) in a constrained object identification task.

In the first of these studies, Lederman and Klatsky (1990b) required subjects to decide whether or not an object placed in the hands was a member of a named category, where the name was either at a relatively inclusive or specific level (e.g., a wine class might be called a ‘drinking vessel’ or a ‘glass’) (for details of object classification taxonomy, see Rosch, 1978). Subjects were permitted to explore the objects freely; hand movements were videotaped on each trial.

Each trial was analyzed as a sequence of EPs. The analysis indicated that people consistently selected a two-stage sequence of exploratory procedures, as evident in Fig. 2. The graph shows the percent cumulative occurrence of each EP as a function of its position in the EP sequence. Stage 1 (represented by the thick, dark lines) consisted of a general grasp and lift (Enclosure and Unsupported Holding) routine, regardless of the class targeted in the question. These two EPs, both of which are broadly sufficient and relatively fast to perform, can provide coarse information about almost all object properties. In contrast, Stage 2 (represented by the thinner, lighter lines) consisted of the more precise EPs that are optimal for extracting the property most diagnostic of the targeted object class. We had previously obtained the requisite information about the property diagnosticity for our current object classes in a separate study.

A second study in this series (Klatsky and Lederman, 1992) constrained subjects initially to perform only the grasp-and-lift sequence of Stage 1; following that, they were allowed to explore further. Accuracy was above chance after Stage 1, confirming our assumption that the grasp/lift routine was broadly useful. Stage 2 was executed to increase the subjects’ accuracy and confidence. It primarily elicited exploratory procedures associated with
object geometry, although exploration was also influenced by diagnostic object properties.

Next we turn to circumstances that do not explicitly or implicitly target a specific property for further examination. Given the relative accuracy and speed with which material as opposed to geometric properties can be extracted with haptic EPs, we predicted in this unconstrained situation that subjects would attend more to material than geometric properties when touch alone was used. In contrast, we argued that, when vision was also allowed, subjects would emphasize the geometric properties more strongly. Here, vision would be considerably more efficient than any haptic EP at extracting geometry but somewhat less effective than touch at extracting material properties.

We confirmed these predictions in a series of studies (Klatzky et al., 1987b; Summers et al., submitted) in which subjects were asked to sort multidimensional objects that varied on material (e.g., texture, hardness, thermal, and weight) and geometric (shape, size) dimensions. The objects were factorially manipulated by property; in one study, for example, three values each were used for texture, hardness, shape and size variations. Subjects were asked to sort the 81 objects into three piles by object similarity. Thus for example, if subjects sorted by texture, they necessarily aggregated objects on the other three dimensions (that is, objects with different hardness, shapes and sizes were placed in the same similarity pile). In this case, we would conclude that subjects judged texture to be more cognitively salient than the other dimensions.

The results of both studies collectively confirmed that the relative cognitive salience of material versus geometric properties is influenced by a modality encoding bias. When the instructions biased subjects toward haptic encoding, subjects emphasized material more strongly than geometric properties. When instructions biased subjects toward visual encoding (either by requiring them to sort haptically by the similarity of their visual images, or by providing vision in addition to touch), subjects' judgments indicated that the geometric properties (particularly, shape) were more cognitively salient than material properties.

EPs serve a 'gateway' role in determining the kinds and precision of information available for simultaneous processing by the haptic system. This was strikingly shown by another series of studies (Klatzky et al., 1989; Reed et al., 1990; Lederman et al., in press) investigating the extent to which variation on more than one object dimension influences the speed with which subjects learn to classify objects. A variety of dimensions was used but the principal results will be demonstrated here by considering texture, hardness and shape.

An initial series of studies used the same planar objects described in the cognitive salience studies above. Subjects were required to learn different classification rules. A one-dimensional rule classified multidimensional objects by variation on only one dimension (texture, shape, or hardness). Two-dimensional (texture/hardness; texture/shape; shape/hardness) and three-dimensional (texture/hardness/shape) redundancy rules classified objects by redundant variation on two dimensions (e.g., soft and rough vs. hard and smooth; soft and rough and one-lobed vs. hard and smooth and two-lobed, respectively). Adding redundant information speeded classification, that is, response times for two-dimensional classification tasks were shorter than for one-dimensional tasks (a redundancy gain effect). However, adding a third redundant dimension had no additional positive effect on response times. We interpreted the differences in dimensional integration in terms of whether the subjects could haptically extract the dimensional variations simultaneously, which in turn depends on EP compatibility (see Table 2).

We used another experimental paradigm, redundancy withdrawal, to pursue this interpretation. Subjects first learned a two-dimensional redundancy rule in which one of the two dimensions was targeted; when performance reached asymptote, the non-targeted dimension was withdrawn (i.e., held constant). The extent to which response times increased following the withdrawal of the non-targeted dimension was used as a measure of the extent to which the two sources of information were integrated. We found, for example, that redundant variation in both texture and hardness was strongly integrated, regardless of which dimension was withdrawn. In contrast, there was little, if any, effect of redundancy withdrawal when either texture or hardness varied redundantly with shape. This is not surprising, given that
evaluation is concerned, the tongue and teeth now become the principal agents of perceptual exploration. Together these provide considerable information concerning a food's material and geometric properties, presumably using versions of the exploratory procedures discussed above. We know of no research that has applied our conceptual framework concerning haptic exploration and perception of object properties to food materials.

References


6. Applications

The theoretical approach and empirical results of our research programme on haptic perception and recognition of objects have implications for the solution of a number of challenging practical problems.

First, the work suggests reasons why the success of tangible graphics displays for the blind (graphs, picture, and especially raised line depictions of common objects; see Lederman et al., 1990) has been relatively limited to date.

Second, we have argued (e.g., Klatzky et al., 1987a; Lederman et al., 1992) that knowledge of biological tactile/haptic systems potentially provides many new avenues in the design of autonomous and teleoperated sensor-based robotic systems that incorporate touch (with or without vision). The work on haptic exploration is particularly relevant to the performance of complex robotic perceptual and manipulatory tasks in highly unstructured environments. All such circumstances demand perceptual exploration by an end effector, whether it be the multifingered hand of a robotic arm/hand system or the foot/probe of a mobile robot; recall that the particular end effector was not relevant to the EP classification system.

Finally and in keeping with many of the articles in this Issue, our approach may be further extended to research by the food industry. Where food