Inferring the depth of 3-D objects from tactile spatial information

Hiroyuki Mitsudo

Published online: 12 March 2015
© The Psychonomic Society, Inc. 2015

Abstract Four psychophysical experiments were conducted to examine the relation between tactile spatial information and the estimated depth of partially touched 3-D objects. Human participants touched unseen, tactile grating patterns with their hand while keeping the hand shape flat. Experiment 1, by means of a production task, showed that the estimated depth of the concave part below the touched grating was well correlated with the separation between the elements of the grating, but not with the overall size of the grating, nor with the local structure of the touched parts. Experiments 2 and 3, by means of a haptic working memory task, showed that the remembered depth of a target surface was biased toward the estimated bottom position of a tactile grating distractor. Experiment 4, by means of a discrimination task, revealed that tactile grating patterns influenced speeded judgments about visual 3-D shapes. These results suggest that the haptic system uses heuristics based on spatial information to infer the depth of an untouched part of a 3-D object.

Keywords Cutaneous input · Somatotopic information · Depth · Haptic working memory · Speeded discrimination

Haptic object recognition involves analyzing tactile spatial and material information through a set of hand movements called *exploratory procedures* (Lederman & Klatzky, 1987, 2009). Examples of spatial information¹ include the layout, shape, and size of touched objects and their parts, whereas examples of material information include the roughness and hardness of touched surfaces (Loomis & Lederman, 1986). For sighted and blindfolded participants, haptic object recognition is easier for familiar 3-D objects than for their 2-D versions such as raised-line drawings (e.g., Klatzky, Loomis, Lederman, Wake, & Fujita, 1993). Lawson and Bracken (2011) found that this was the case even when material information was controlled. Furthermore, these studies reported that recognition performance was better when participants used the hand than when they used only one finger. When participants touch an object with the hand, the haptic system receives much spatial information across a relatively wide area of the skin. How does the haptic system use spatial information in object recognition?

A straightforward idea is that the haptic system uses spatial information to recognize the properties of the touched object parts (e.g., shape, size, and relative location), which may be distinctive enough to identify familiar 3-D objects. Indeed, the haptic system is sensitive to several local properties of objects such as length (Green, 1982), curvature (Goodwin, John, & Marcegaglia, 1991), and orientation (Frisoli, Solazzi, Reiner, & Bergamasco, 2011; Levy, Bourgeon, & Chapman, 2007). Given that familiar 3-D objects contain more distinctive parts than

H. Mitsudo
Faculty of Human-Environment Studies, Kyushu University,
Fukuoka, Japan

H. Mitsudo (✉)
Department of Human Sciences, Faculty of Human-Environment
Studies, Kyushu University, 6-19-1 Hakozaki,
Higashiku, Fukuoka 812-8581, Japan
e-mail: hmitsudo@lit.kyushu-u.ac.jp

¹ Throughout this article, I use the term “spatial information” to refer to 2-D tactile patterns of the physical force applied to the skin, whereas tactile spatial information (or somatotopic information) also includes the magnitude and direction of the force applied. In other studies, the term “space” has also been used to mean the 3-D environment recognized through kinesthetic and proprioceptive inputs (e.g., Haggard, Newman, Blundell, & Andrew, 2000).

2-D patterns do, haptic recognition performance would be better for 3-D objects than for 2-D patterns.

A related but unresolved issue is whether the haptic system uses spatial information to infer or estimate *untouched*, rather than *touched*, parts of a 3-D object. For example, when the hand touches a 3-D object with a concave part in real-life situations (Fig. 1a), the outer edges or elements of the concave part can simultaneously stimulate different skin locations separated by a certain distance. At this moment, unless the hand directly touches the bottom of the concave part in a subsequent movement, the haptic system in principle cannot know how deep the concave part is. In vision, an analogous issue has been investigated extensively.² It is known that the visual system quickly completes partially occluded shapes and edges by using pictorial image cues such as T-junctions (e.g., Rensink & Enns, 1998). In haptics, Kennedy, Gabias, and Nicholls (1991) mentioned that no study had systematically examined whether humans can recognize raised-line drawings that contain T-junctions. Because raised-line drawings are generally difficult to recognize by touch, as was noted above, I will not consider raised-line drawings here.

To gain insight into the issue of whether spatial information is used to infer the depth of an untouched part of a 3-D object, I investigated whether the physical depth of concave object parts is correlated with interelement separation in real-life situations. Eighty-seven sample objects were collected in indoor environments, according to the following two criteria: In ordinary situations, (a) the object had a concave part of a fixed size defined by obviously touchable outer parts (or edges and a rim), and (b) both the separation between the parts and the depth of the concave part were measurable and approximately within a range of 1 to 15 cm. To reduce sampling bias, objects were collected from a variety of categories of common objects, including personal and entertainment items, kitchen supplies, and office supplies (Klatzky, Lederman, & Metzger, 1985). A digital caliper was used to measure (a) the separation between the outer parts and (b) the depth of the concave part(s) of each object.³ Figure 1b shows the depth of the concave part as a function of the interelement separation ($n = 87$).

² As in the case of vision, here I consider haptic “inference” to be an aspect of object recognition. Note that haptic experience itself is not restricted to within the body, although the receptors that process kinesthetic and cutaneous inputs are within the body (for a discussion, see Loomis, 1992).

³ The separation measured here was the shortest distance between, or the width of empty space bounded by, the adjacent outer parts. For objects with repetitive outer parts that defined concave parts (e.g., a container with partitions), the mean of the separations was calculated. For objects with a circle rim (e.g., a cup), the diameter was measured. For objects whose outer shape was not a circle (e.g., an oval rim or a rectangular hole), the shortest separation between the outer parts facing each other was measured. The depth measured here was the shortest distance from the imaginary plane that intersected the touched parts to the deepest position of the bottom part. For containers such as a cup or a box, their contents were excluded during the measurements.

The correlation coefficient was .59 and was significantly different from 0 ($p < .0001$). The results suggest that, at least for manmade objects, the physical depth of a concave part is, to some extent, positively correlated with the separation between touched locations in real-life situations.

The primary purpose of this study was to examine whether, as in the case of real-life situations, the *estimated* depth of untouched object parts would be correlated with interelement separation. By examining this issue experimentally, I tried to provide insights into the issue of whether the haptic system analyzes spatial information to infer the depth of an untouched part of a 3-D object. In this study, I used tactile grating patterns consisting of bar elements (e.g., Johnson & Phillips, 1981), the top of which were touched by participants (Fig. 2). Touching a periodic pattern with an interelement separation of less than 1.25 mm can produce a sense of roughness, even without a lateral hand movement (e.g., Hollins & Risner, 2000; Lederman & Taylor, 1972). In the present study, I used much greater interelement separations (20–80 mm) and focused on the depth of trenches perceived between the tactile bar elements. In the present experiments, participants were asked to touch unseen grating patterns with their hands and to keep the hand shape flat, so that a tactile pattern was created on the palm and fingers (i.e., static contact: Lederman & Klatzky, 1987). The actual depth of the trenches between the bars was always 10 mm. Throughout the experiments, the participants did not touch the bottom position of the trenches. If the haptic system uses spatial information to infer the depth of an untouched part of a 3-D object, the estimated depth of the concave part would increase as a function of interelement separation (Exp. 1), would influence positional judgments in a haptic working memory task (Exps. 2 and 3), and would influence speeded judgments in a visual object discrimination task (Exp. 4).

Experiment 1

In Experiment 1, participants were asked to estimate the bottom position (i.e., depth) of trenches perceived between unseen tactile bar elements and to produce the estimated depth, by pointing at a response panel with a tablet pen held in the left hand (Fig. 2b). Because the response panel and the pen were clearly visible to the participants, this task involved transforming the haptically estimated depth into a visual judgment. One might think that this production task is unconventional, because (a) no correct response was expected, because of the nature of the task, and (b) no direct haptic judgment was required. However, if estimated size and position are assumed to be more accurate in vision than in haptics (e.g., Gentaz & Hatwell, 2003; Green, 1982), participants’ responses in this task could be regarded as accurately reflecting the haptically estimated depth. I varied the spatial configuration of the bars

Fig. 1 Examples of objects with concave part(s) in real-life situations. (a) Pictures of objects collected in indoor environments. (b) Correlations between the interelement separations and depths of the sampled objects.

(including interelement separation and overall size, 20–80 mm) and the shape of the bars (round, square) across trials.

Method

Participants Nine sighted volunteers (between 20 and 36 years of age; seven female, two male) participated in the experiment. No participant reported a sensorimotor deficit. All were right-handed and were naive as to the purpose of the experiment. The experiment was approved by the ethics committee of the Faculty of Human-Environment Studies of Kyushu University. Written consent was obtained from all participants.

Apparatus and stimuli Figure 2 shows the tactile stimulus objects, consisting of plastic bars (length, 80 mm) placed parallel to each other on a horizontal plastic plate (80 × 100 mm).

Two types of bar shapes were used, in which the vertical cross-section was either round (a cylinder with a diameter of 5 mm) or square (5 × 5 mm). The bars were attached to square-shaped ribs (80 × 5 × 5 mm) on a plate, resulting in *actual* depths of 10 mm for the trenches between the bars (Fig. 2c). Four types of bar configurations were used. In three conditions in which the overall size was 80 mm (the right three panels in Fig. 2a), five, three, and two bars were arranged to form three separations between the bars (20, 40, and 80 mm, respectively). In another condition, in which the overall size was 20 mm, two bars were arranged at the center of the plate (the left panel in Fig. 2a).

The stimulus object was placed horizontally with respect to the tabletop at a slant of 55° about a vertical axis, so that the participant touched the object in a comfortable position (Fig. 2b). The top of the object was fixed at a height of 25 cm from the tabletop and was surrounded by two opaque

Fig. 2 Schematic illustration of the tactile stimulus objects and setup for Experiment 1. (a) Top view of the tactile gratings. (b) Overview of the experimental setup. (c) Side view of the gratings and the response apparatus. The upper half represents round- and square-shaped 80-mm-size gratings with an interelement separation of 20 mm. The bottom half represents a part of the response apparatus, indicated by the dashed oval in panel b. The right hand and the stimulus, drawn in dashed lines, were masked from the participant's view.

large boards (45 × 60 cm). The two boards were separated by 15 cm and placed in parallel to each other.

The participant's responses were made using a pen tablet (Wacom Intuos4) attached to the large board that was close to the participant's body (Fig. 2b). This board visually masked both the stimulus and the right hand from the participant's view. A white plastic panel (32 × 18 × 0.3 cm), visible to the participant, was attached to the tablet surface. On the white panel, a strip of black paper (80 × 2 mm) was stuck at the same height as the top of the object and served as a reference line for reporting the depth of the trenches.

Procedure In a quiet, well-lit room, the participant sat in a chair in front of the table. The participant was instructed to keep his/her right hand flat throughout each trial. At the beginning of each trial, the participant placed the right hand in the starting position, 30 cm above the tabletop. The position was indicated by a marker attached to the large board that was placed away from the participant's body (Fig. 2b). The participant moved the hand forward until the fingertips touched a small end-panel orthogonal to the fingers. The separation between the end panel and the front-end of the tactile stimulus was 50 mm (Fig. 2c). Then, the participant lowered the hand in order to firmly touch the top of all the bars while keeping the hand as flat as possible. The participant did not touch the bottom position of the trenches. The participant was also instructed to keep the contact force as constant as possible across trials. (I was not able to measure the contact force in the experiment.) The experimenter visually checked the participant's hand movement. In practice trials, the participant received verbal feedback from the experimenter in order to follow the instructions. A red light-emitting diode (LED) turned on when the hand touched the stimulus, and turned off 5 s after. The participant's task was to estimate and then produce the bottom position of the trenches perceived between the bars, by pointing at the response panel with the tablet pen held in the left hand. The participant was instructed to regard the horizontal reference line on the response panel as representing the height of the touched position. Both the reference line and the tablet pen were clearly visible to the participant. The participant was required to respond before or when the LED turned off. A touch on the response panel was indicated to the participant by a click. The participant was allowed to correct the reported position by immediately touching the screen if necessary. At the end of each trial, both hands returned to a resting position.

Each participant completed six blocks of the eight conditions. The first block was practice trials. The order of trials was randomized across blocks and participants. The number of repetitions and that of participants were determined on the basis of previous studies that used a similar procedure (e.g., Green, 1982).

Results and discussion

Figure 3 shows the mean depth estimates as a function of interelement separation ($n = 9$). A two-way repeated measures analysis of variance (ANOVA) was performed on the mean depth estimates, with the factors of Bar Configuration and Bar Shape. Only the main effect of bar configuration was significant [$F(3, 24) = 29.3, p < .0001, \eta^2 = .63$]. Multiple comparisons (Ryan's method) revealed that the depth estimates were significantly greater with an interelement separation of 80 mm than with the other three separations ($ps < .05$). In addition, a linear regression analysis for the individual data revealed a significant positive correlation between interelement separation and the depth estimates for all nine of the participants ($ps < .001$; mean slope = 0.83, mean adjusted $R^2 = .76$).

These results suggest that interelement separation plays an important role in estimating the depth of an untouched part of a 3-D object. The negligible effect of bar shape suggests that the local structure of the touched patterns, such as a gradient of the pressure from each bar element, had a minor role in the present experiment. The depth estimates were consistently small and similar for the configurations with an interelement separation of 20 mm, regardless of the number of the bars touched (two or five). This result indicates that the physical pressure from each bar element does not seem to be sufficient to explain the present results, because the participants were asked to touch the whole pattern at a constant physical force. In addition, this result suggests that the overall size of a touched pattern does not play a critical role in inferring the depth of the concave part.

Experiment 2

Another explanation for the results of Experiment 1 is that the participants might have adopted a response strategy for

Fig. 3 Results for Experiment 1. Mean depth estimates are shown as a function of interelement separation. Open circles and filled squares represent the depth estimates for the round- and square-shaped bar conditions, respectively. Error bars represent standard errors.

directly reporting the perceived *separation* between the bars, without estimating the 3-D structure of the partially touched gratings. Because no correct response was defined and no reliable information was provided in the production task, the participants might depend heavily on the most salient information—interelement separation. To examine this possibility, I used a haptic working memory task in which a correct response was defined. This task was a haptic-depth version of the Brown–Peterson paradigm (e.g., Kaas, Stoeckel, & Goebel, 2008). In the present experiment, participants haptically reproduced the remembered vertical position of a flat target object after touching a tactile grating distractor (Fig. 4a). I varied the vertical position of the target (i.e., depth: 26 and 36 mm below the top of the distractor) and the distractor type (two gratings with an interelement separation of 20 or 80 mm) across trials. I also used a flat-panel distractor to establish a baseline for the positional judgment. When a target and a distractor share a certain feature dimension in working

memory (e.g., temporal frequency), judgments about the target feature can be influenced by the feature value of the distractor (through *feature overwriting*; Bancroft & Servos, 2011; Mercer & McKeown, 2010). If the bottom surface of tactile gratings is maintained in working memory, the reproduced vertical position for a target will be affected by the estimated position of the bottom surface of a tactile grating distractor.

Method

The methods were identical to those used in Experiment 1, except for the following.

Participants Nine sighted volunteers (between 20 and 36 years of age; seven female, two male) participated in the experiment. All were right-handed, except for one. All were naive as to the purpose of the experiment. One had also participated in Experiment 1. The data from one of the participants were excluded from the analysis because she could not reliably touch the response panel with the middle fingertip.

Apparatus and stimuli Figure 4a shows the main apparatus, consisting of three components: target and distractor objects and a response panel. These components and both hands were masked from the participant's view by a horizontally oriented black board (60 × 45 cm) placed in front of the participant at a height of 23 cm from the tabletop. The board had three small translucent windows (2 × 2 cm each, separated from each other by 18 cm) that indicated the approximate horizontal positions of the three components. Light from a red LED was visible to the participant through each window.

The top area of the target and the distractor was 150 × 80 mm. The target was a flat plastic plate (thickness, 2 mm) for which the depth (vertical position) was variable across trials: either 26 or 36 mm below the top of the distractor. Of the three distractor patterns, two were grating patterns consisting of round bars (diameter 5 mm, length 150 mm) attached to square-shaped ribs (150 × 5 × 5 mm), similar to those used in Experiment 1, and one was a flat plate identical to the target. The numbers of bars were five and two, yielding interelement separations of 20 and 80 mm, respectively. As in Experiment 1, the actual depths of the trenches between the bars were always 10 mm. To control the touched position on the hand, small end-panels orthogonal to the fingers (not shown in Fig. 4a) were placed 5 cm ahead of the front ends of the target and the distractor.

The participants' responses were recorded through a 7-in. touch panel (Quixun QT701AV, 800 × 480 pixels, 152 × 90 mm). The response panel was placed to the left of the end panel for the distractor. The top of the response panel was 19 mm above that of the distractor.

Fig. 4 Setup and results for Experiment 2. (a) Schematic illustration of the experimental setup. In this panel, the distractor object has an interelement separation of 80 mm. Dashed lines approximately represent the required hand movements of the right hand during a trial. See the text for details. (b) Mean vertical position shifts for the four combinations of target depth and distractor type. Positive values indicate that the reproduced position was lower for a grating distractor than for the baseline, flat-panel distractor. Error bars represent standard errors.

Procedure The participant was instructed to keep the right hand as flat as possible and the left hand in a resting position during each trial. At the beginning of each trial, the participant set the right hand to the starting position, which was in front of the target and 11 cm above the tabletop, indicated by a plastic strip (15 × 2 cm) attached to a vertical board placed on the right side of the target. The participant was asked to touch the three components sequentially in the following three steps, guided by the three LEDs, while keeping the direction of the axis of the right hand parallel to the midsagittal plane and the palm parallel to the transverse plane as accurately as possible. First, the participant moved the right hand forward until the middle fingertip touched the end panel for the target, and then lowered the hand to firmly touch the target surface. The participant touched the target for 5 s. During this period, the participant was asked to remember the vertical position, while keeping his/her hand stationary. Second, the participant moved the hand approximately 7 cm upward and 18 cm left, and lowered the hand to firmly touch the distractor. The participant touched the distractor for 5 s. When touching the target and distractor, the participant was instructed to keep the contact force as constant as possible across the objects and trials. Third, the participant moved the hand approximately 7 cm upward and 18 cm left again, and lowered the hand to reproduce the remembered vertical position of the hand when touching the target surface. Then the participant moved the hand forward in order to touch the response panel with the middle fingertip. The participant was instructed to touch the panel 2 s after touching the distractor. The participant's contact with the response panel was indicated to the participant by a click.

Each participant completed 11 blocks of the six conditions. The first two blocks were practice trials.

Results and discussion

Figure 4b shows the mean position shifts, calculated by subtracting the position obtained for the flat-panel distractor from the positions for the grating distractors ($n = 8$). By using this procedure, I aimed to remove the effect of touching the distractor itself from the data. The mean reproduced positions for the flat-panel distractor were 22.0 ($SE = 8.2$) and 26.2 ($SE = 7.5$) mm at the target depths of 26 and 36 mm, respectively. A two-way repeated measures ANOVA was performed on the position shifts, with the factors of Target Depth (26, 36 mm) and Distractor Type (interelement separations of 20, 80 mm). Neither the main effect of target depth nor that of distractor type was significant [$F(1, 7) = 0.062, p = .81, \eta^2 < .001$, and $F(1, 7) = 1.2, p = .31, \eta^2 = .01$, respectively]. The two-way interaction was significant [$F(1, 7) = 6.8, p = .035, \eta^2 = .03$]. At a target depth of 36 mm, the reproduced position was significantly lower for the 80-mm-separation distractor than for the 20-mm-separation distractor [$F(1, 14) = 5.8, p = .03$];

at a target depth of 26 mm, similar magnitudes of the downward position shift were observed for the two distractors. Furthermore, I examined whether this result was seen in the data from individual trials. In an attempt to remove fluctuations across trials and participants in the manual responses, I matched the position shifts for the two distractors on experimental block. I performed two-sided paired t tests on the pooled position data and again found that distractor type had a significant effect on the reproduced positions for a target depth of 36 mm [$t(71) = 2.5, p = .015, r = .29$], but not for 26 mm [$t(71) = 0.71, p = .48, r = .08$].

The observed downward shifts are consistent with the idea that the estimated depth of the distractor gratings is maintained in working memory, if a few assumptions are made. In terms of the principle of feature overwriting, I assume that a downward shift occurs when the estimated bottom surface of the distractor is *lower* than the retained target depth. To obtain a prediction, it is necessary to assume specific values for both the retained target depth and the bottom surface of the distractor. With respect to the target, on the basis of the findings of Baud-Bovy and Viviani (1998), I assume that the retained target positions are higher than the actual ones—for example, approximately 10 and 20 mm, instead of actual depths of 26 and 36 mm, respectively (i.e., a 16-mm upward shift relative to the actual position). With respect to the distractor, according to the results of Experiment 1, I assume that the estimated bottom positions of the distractors are 20 and 65 mm for interelement separations of 20 and 80 mm, respectively. Given these assumptions, at a target depth of 36 mm, a downward position shift would occur selectively for the 80-mm-separation distractor because the estimated bottom surface of the distractor (65 mm) is lower than the retained target depth (20 mm), but that of the 20-mm-separation distractor (20 mm) is not. At a target depth of 26 mm, overwriting would produce downward position shifts for both of the distractors because the bottom surfaces of both distractors (65 and 20 mm) are lower than the retained target depth (10 mm). These predictions are consistent with the observed downward shifts.

The present results are not explained by local skin indentation produced by touch, because the amount of skin indentation was generally small on the skin of the hand, less than 2 mm (Greenspan, 1984).

Experiment 3

To examine whether the assumptions made in Experiment 2 are valid, I conducted a similar working memory experiment with different target depths. In Experiment 3, I used target depths of 16 and 46 mm (instead of 26 and 36 mm in Exp. 2). On the basis of the assumptions made above (i.e., a 16-mm upward shift relative to the actual position), the retained target

depths were approximately 0 and 30 mm, respectively. If the other experimental conditions were identical to those of Experiment 2, downward shifts would be obtained again, except for a target depth of 46 mm with the 20-mm-separation distractor.

Method

The methods were identical to those used in Experiment 2, except for the following.

Participants Ten sighted volunteers (between 20 and 36 years of age; three female, seven male) participated in the experiment. All were right-handed except for one; all participants were naive as to the purpose of the experiment, and no one had participated in Experiment 1 or 2. The data from two participants were excluded from the analysis because they could not touch the three distractor objects in a consistent manner.

Apparatus, stimuli, and procedure The apparatus, stimuli, and procedure were the same as those used in Experiment 2, except that the depths of the target were 16 and 46 mm.

Results and discussion

Figure 5 shows the mean position shifts, calculated according to the same procedure that was used in Experiment 2 ($n = 8$). The mean reproduced positions for the flat-panel distractor were 28.5 ($SE = 6.6$) and 47.7 ($SE = 5.5$) mm at target depths of 16 and 46 mm, respectively. As in Experiment 2, a two-way repeated measures ANOVA was performed on the position

shifts, with the factors of Target Depth (16, 46 mm) and Distractor Type (interelement separations of 20, 80 mm). Neither the main effect of target depth nor that of distractor type was significant [$F(1, 7) = 2.8, p = .14, \eta^2 = .14$, and $F(1, 7) = 0.7, p = .43, \eta^2 = .02$, respectively]. The two-way interaction was significant [$F(1, 7) = 13.1, p = .0086, \eta^2 = .07$]. At a target depth of 46 mm, the reproduced position was significantly lower for the 80-mm-separation distractor than for the 20-mm-separation distractor [$F(1, 14) = 5.8, p = .03$]; at a target depth of 16 mm, similar magnitudes of the downward position shift were observed for the two distractors. As in Experiment 2, I performed two-sided paired t tests on the pooled position data from the individual trials, and again found that distractor type had a significant effect on the reproduced position for a target depth of 46 mm [$t(71) = 2.2, p = .035, r = .25$], but not for 16 mm [$t(71) = 0.87, p = .39, r = .10$]. These results replicated those of Experiment 2, suggesting that (a) the assumptions of the feature-overwriting account are valid, and therefore (b) the untouched bottom surface of the distractor is maintained in haptic working memory.

Experiment 4

The results of Experiments 1–3 suggested that interelement separation is associated with the estimated depths of untouched parts of 3-D objects. Experiment 4 was designed to examine this association by using a speeded discrimination task with visual test stimuli. Speeded judgments such as discrimination and classification have been widely used to investigate automatic (and probably perceptual) crossmodal processing (see, e.g., Evans & Treisman, 2010; Spence, 2011). In the present experiment, while touching two objects with both hands, participants viewed a pair of 3-D visual test objects (Fig. 6). I used two different object pairs: one consisted of shapes whose interelement separation and trench depth were *correlated* within each object (as is the case in real-life situations), and the other consisted of shapes whose interelement separation and depth were *anticorrelated*. For each pair, the participants' task was to judge, as quickly as possible, which visual object had a deeper/shallower trench. A pair of touched objects, placed where the two visual test objects appeared (Fig. 6a), were task-irrelevant and differed between the following two conditions: In a crossmodally congruent condition, the shapes of the touched parts were identical to the square-shaped gratings used in Experiment 1, which was consistent with the visual display; in a crossmodally neutral condition, the touched objects were two hemispheres that were not consistent with the visual display. If tactile interelement separation is used to judge the depths of untouched parts of 3-D objects in an automatic manner, the crossmodally congruent tactile patterns would facilitate reaction times (RTs) to the visual correlation pair.

Fig. 5 Mean vertical position shifts for the four combinations of target depth and distractor type in Experiment 3. The target depths were different from those in Experiment 2. Positive values indicate that the reproduced position was lower for a grating distractor than for the baseline, flat-panel distractor. Error bars represent standard errors.

Fig. 6 Setup and visual stimuli for Experiment 4. Dashed lines schematically represent both hands, which were masked from the participant's view. (a) Schematic illustration of the side view of the apparatus in the crossmodally neutral condition (i.e., the haptic objects were hemispheres). (b–d) 3-D objects were the visual stimuli presented. Arrows represent required hand movements during a trial. (b) The starting positions of the hands and the visual display before touch. (c) The visual

correlation stimulus. (d) The visual anticorrelation stimulus. In panels c and d, arrows represent the required hand movements for participants who were asked to respond to objects with deeper trenches. The top parts of the two visual test objects were consistent with the touched objects in the crossmodally congruent condition, and not in the neutral condition. See the text for details.

Method

The methods were the same as those used in the previous experiments, except for the following.

Participants A total of 24 sighted volunteers (between 19 and 45 years of age; 15 female, nine male) participated in the experiment. All were naive as to the purpose of the experiment, but eight had participated in the previous experiments. All but two were right-handed. Six were assigned to each of the four combinations of crossmodal congruency (congruent, neutral) and response type (deeper, shallower).

Apparatus and stimuli Visual stimuli were presented on a 20-in. LCD screen (Apple Cinema Display, $1,680 \times 1,050$ pixels, frame rate of 60 Hz). Stimulus presentation and data collection were controlled by a personal computer (Apple Macbook) using MATLAB with the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997).

A visual display consisted of two simulated cube-like 3-D objects placed side by side on a simulated rectangle viewed from above (Fig. 6c–d). The visual display was presented in the bottom part of the screen, as if the simulated objects were on the table. Each object subtended approximately 10° of visual angle, and the two objects were separated by approximately 10° of visual angle. All elements had zero binocular

disparity on the screen. All of the stimuli were achromatic; the luminance of the bottom rectangle was 3.9 cd/m^2 , and the brightest part of the 3-D objects had a luminance of 14.9 cd/m^2 . Each object had a constant simulated height of 8 cm and a flat-bottom trench defined by two 5-mm-width edge plates on the top. The simulated separations between the two edge plates were 20 and 80 mm, and the simulated depths of the trench were 40 and 60 mm. To make these shapes easy to recognize for participants, the two edge plates of each object were rendered translucent.

There were two types of visual test display, which differed in terms of the combination of interelement separation and depth. In a visual correlation stimulus (Fig. 6c), one object had an interelement separation of 80 mm and a depth of 60 mm, and the other had a separation of 20 mm and a depth of 40 mm. In a visual anticorrelation stimulus (Fig. 6d), one object had a separation of 80 mm and a depth of 40 mm, and the other had a separation of 20 mm and a depth of 60 mm.

Haptic objects were placed under the visual screen, 22 mm above the tabletop, and masked from the participant's view by a large horizontal board. Two types of haptic objects were used. In a crossmodally congruent condition, the top parts of the touched objects were the square-shaped gratings used in Experiment 1 and were placed at the corresponding position for the two visual objects. The interelement separations were 20 and 80 mm. The two grating objects were separated by 180 mm from center to center and placed side by side. In a

crossmodally neutral condition, two plastic hemispheres with a radius of 20 mm were placed at the same center positions as those in the crossmodally congruent condition. The purpose of using two hemisphere objects was to reduce crossmodal congruency effectively, because their shape differed from the visual stimuli locally as well as globally. To measure the force from the hands and RTs, four force-sensing resistors (Interlink Electronics FSR-406; two for each object) were placed under the haptic objects.

Procedure The participant binocularly viewed the visual screen at a viewing distance of approximately 40 cm. Both hands, as well as the haptic objects, were masked from the participant's view during a trial. Throughout each trial, the participant was asked to keep both hands flat, aligned to each other, and parallel to the transverse plane. The detailed instructions on moving the hand were similar to those used in Experiment 1. At the beginning of each trial, the participant placed both hands in the starting positions, approximately 10 cm above the tabletop. On the visual screen, two identical, simulated cubes (edge length, 80 mm) were presented side by side (Fig. 6b), and a fixation cross (approximately $1^\circ \times 1^\circ$ of visual angle, not shown in figure) was presented in the middle of the two cubes. The participant was asked to touch the two haptic objects using both hands, while keeping both hands as flat as possible. When the force from each hand exceeded 3.5 N, a blank display appeared, followed by a visual test display. For the experimental trials, the mean force averaged over all participants was 10.4 N ($SD = 3.4$) while touching. The stimulus onset asynchrony (SOA) between the blank display (triggered by touch) and the visual test display was 200 and 1,000 ms. Each participant was asked to judge which visual object had a deeper/shallower trench and to respond by releasing the hand on the corresponding side of the two touched objects. An RT was recorded when the force from one hand became less than 3.5 N. Half of the participants responded to the deeper side, and the other half responded to the shallower side. Each

participant was asked to respond as quickly as possible while maintaining accuracy. When a response was made, the visual display disappeared and was followed by a feedback display presented for 500 ms. When the response was correct, the “+” fixation was presented; when the response was incorrect, an “x” mark was presented. After that, the participant returned both hands to the starting position.

One block consisted of 32 trials. The horizontal position of the two visual objects (and therefore of the haptic grating objects used in the crossmodally congruent condition), trench depth, and SOA were randomized across trials. Each participant completed four blocks. The first block was practice trials.

Results and discussion

The median correct RTs and error rates were calculated for each condition and each participant. Calculations were carried out after excluding trials in which an incorrect haptic stimulus was presented accidentally or the participant responded before the visual test display was presented (ten trials in total, 0.35 % of all trials). Data were collapsed across the response positions (right, left). Because the main concern of Experiment 4 was to see whether crossmodally congruent tactile patterns influenced the RTs, it was unnecessary to directly compare the visual correlation and anticorrelation pairs. Therefore, in the following analyses, a three-way mixed-design ANOVA was conducted on the median RTs separately for each visual test pair, with the factors of Crossmodal Congruency (congruent, neutral), Response Type (deeper, shallower), and SOA (200, 1,000 ms). Crossmodal Congruency and Response Type were between-participants factors, and SOA was a within-participants factor. Figure 7 shows the mean values of the median RTs and error rates, averaged over the 12 participants in each condition and collapsed across the response types.

Reaction time For the visual correlation stimulus, the main effect of SOA was significant [$F(1, 20) = 21.1, p = .0002, \eta^2$

Fig. 7 Mean values of median reaction times and error rates as a function of stimulus onset asynchrony in Experiment 4. (a) The visual correlation stimulus. (b) The visual anticorrelation stimulus. Asterisks indicate $ps < .05$. Error bars represent standard errors.

= .10]. The two-way interaction between crossmodal congruency and SOA was also significant [$F(1, 20) = 5.3, p = .032, \eta^2 = .03$]. At an SOA of 200 ms, RTs were shorter in the crossmodally congruent condition than in the neutral condition [$F(1, 40) = 4.3, p = .045$]. At an SOA of 1,000 ms, RTs did not differ significantly between the crossmodally congruent and neutral conditions [$F(1, 40) = 0.05, p = .82$]. Furthermore, RTs were shorter at an SOA of 1,000 ms than at that of 200 ms in the neutral condition [$F(1, 20) = 23.8, p = .0001$], but not in the crossmodally congruent condition [$F(1, 20) = 2.6, p = .12$]. The two-way interaction between crossmodal congruency and response type was also significant [$F(1, 20) = 4.9, p = .039, \eta^2 = .12$].

For the visual anticorrelation stimulus, the main effect of SOA was again significant [$F(1, 20) = 35.9, p < .0001, \eta^2 = .07$]. As for the visual correlation stimulus, the two-way interaction between crossmodal congruency and SOA was significant [$F(1, 20) = 4.7, p = .042, \eta^2 = .01$]. Unlike for the visual correlation stimulus, RTs did not differ significantly between the crossmodally congruent and neutral conditions at an SOA of either 200 or 1,000 ms [$F(1, 40) = 2.8, p = .10$, and $F(1, 40) = 0.48, p = .49$, respectively]. Furthermore, RTs were shorter at an SOA of 1,000 ms than an SOA of 200 ms in both the crossmodally congruent and neutral conditions [$F(1, 20) = 7.3, p = .014$, and $F(1, 20) = 33.3, p < .0001$, respectively]. The two-way interaction between crossmodal congruency and response type was significant [$F(1, 20) = 4.7, p = .043, \eta^2 = .16$].

Error rate Three-way ANOVAs were performed on the arcsine-transformed error rates, with the same three factors that were used for the RT data. For the visual correlation stimulus, no main effect or interaction was significant. For the visual anticorrelation stimulus, the two-way interaction between crossmodal congruency and SOA was significant [$F(1, 20) = 5.6, p = .028, \eta^2 = .06$]: In the crossmodally neutral condition, error rates were marginally lower at an SOA of 200 ms than at an SOA of 1,000 ms [$F(1, 20) = 4.0, p = .060$].

Summary The main results can be summarized as follows: (a) The crossmodally congruent tactile patterns facilitated RTs to the visual correlation pair at an SOA of 200 ms, and (b) the crossmodal congruency effect did not accompany a speed–accuracy trade-off. Because a facilitation effect obtained at short SOAs (approximately less than 200 ms) is thought to reflect automatic processing (e.g., Neely, 1991), the present results suggest that tactile spatial patterns can be used to infer the depths of untouched parts of 3-D objects in an automatic (and probably perceptual) manner.

Even for the visual anticorrelation pair, RTs were somewhat shorter in the crossmodally congruent condition than in the neutral condition (Fig. 7b). One might think that this contradicts the idea above, because tactile interelement separation

did not “predict” the depth of the objects in the visual anticorrelation pair. Note that, to perform the task, the participants would be required to accurately recognize the overall 3-D shapes defined by both interelement separation and depth. Therefore, although uninformative, the crossmodally congruent tactile patterns may have acted as a “preview” of the top part of any visual test object. If so, the crossmodally congruent tactile patterns would not necessarily interfere with the discrimination of the visual anticorrelation pair. Crossmodally congruent tactile patterns seem to facilitate the recognition of complex 3-D shapes, especially when the judged shapes are likely to appear in real-life situations, as in the case of the visual correlation pair.

In almost all of the conditions, RTs were shorter at an SOA of 1,000 ms than at an SOA of 200 ms. Given the results, it seems that the temporal proximity between the onsets of tactile and visual stimuli, rather than the entire duration of the task-irrelevant tactile stimuli, increased RTs in the discrimination task. The present results are consistent with the finding that it is generally easy to ignore task-irrelevant information (i.e., the tactile patterns in this experiment) when the onset of a target is temporally separated from that of the task-irrelevant primes and distractors (e.g., de Groot, Thomassen, & Hudson, 1986; Shore, Barnes, & Spence, 2006).

The tactile patterns differed between the two hands in the crossmodally congruent condition (i.e., two grating patterns with different interelement separations), but not in the neutral condition (i.e., two hemispheres of the same size). Participants might therefore have formed some location bias toward either pattern in the crossmodally congruent condition, but not in the neutral condition. Such a bias, if present, does not seem to explain the present data, because the crossmodal congruency effect was specific to the visual stimuli with a correlation between interelement separation and depth.

General discussion

The results of the four experiments are consistent with the idea that interelement separation biases participants’ interpretations of an untouched concave part of a 3-D object. The dependency of the estimated depth magnitude on interelement separation found in Experiment 1 was somewhat similar to those that had been reported in studies on perceived roughness, whereas the spatial scale was much greater in the present study (interelement separations of 20–80 mm) than in the previous studies (e.g., less than 1.25 mm; Lederman & Taylor, 1972). Although interelement separation is unlikely to be the sole cause of haptic depth recognition, the present results suggest that humans use heuristics based on spatial information to infer the depths of untouched parts of 3-D objects.

Several psychophysical studies have investigated the haptic perception of objects near the hand. The present finding is

consistent with studies showing that humans can localize *untouched* objects by using temporally changing cutaneous inputs in the near absence of active body movements (Békésy, 1959; Miyazaki, Hirashima, & Nozaki, 2010). Furthermore, the present finding is also in line with studies that have shown that humans can pick up information about the spatial layout of nearby objects (that are not touched directly) by using tools with exploratory movements (e.g., a rod and string: Barac-Cikoja & Turvey, 1993; Cabe & Hofman, 2012), because tool use involves tactile processing as well as kinesthetic processing.

This study was based on somewhat artificial psychophysical tasks with restricted hand movements. What aspects of daily haptic exploration do these tasks reveal? In real-life situations, to haptically recognize a concave part of a 3-D object, it seems common to (a) touch the outer edges or elements of the object first, and then (b) move the hand or fingers into the concave part. With respect to the first movement, tactile spatial patterns are formed on the skin. I believe that a depth estimation process like that considered here occurs in the first case and is useful for efficiently preparing a subsequent hand movement in the second.

In the present experiments, although the participants were instructed to keep their hand as flat as possible when touching the gratings, the hand might have been slightly flexed toward the untouched bottom plate, especially for the gratings with an interelement separation of 80 mm. Therefore, one could think that hand flexure might be related to the present results. According to informal measurements, the value of possible hand flexure was at most 5 mm for the configurations with an interelement separation of 80 mm, and was close to 0 mm for the other configurations. Possible differences in hand flexure are unlikely to explain the present results for the following two reasons. First, in Experiment 1, the differences in the depth estimates between the gratings with interelement separations of 20 and 80 mm were evidently greater than 5 mm (>40 mm, on average; Fig. 3). Second, in Experiment 2 or 3, no main effect of distractor type (i.e., interelement separation) was found. Nevertheless, in an attempt to keep the hand flat when touching the different gratings, the participants might have made reactive kinematic responses that differed across the configurations. This study does not rule out the possibility that such reactive responses might contribute to the estimated depths of tactile gratings. To examine this issue, it will be helpful to use dynamic touch or more natural hand movements in further studies.

Author note I thank Takako Mitsudo and Kyoshiro Sasaki for comments on an earlier version of this manuscript, and Daiichiro Kuroki for technical assistance.

References

- Bancroft, T., & Servos, P. (2011). Distractor frequency influences performance in vibrotactile working memory. *Experimental Brain Research*, 208, 529–532. doi:10.1007/s00221-010-2501-2
- Barac-Cikoja, D., & Turvey, M. T. (1993). Haptically perceiving size at a distance. *Journal of Experimental Psychology: General*, 122, 347–370. doi:10.1037/0096-3445.122.3.347
- Baud-Bovy, G., & Viviani, P. (1998). Pointing to kinesthetic targets in space. *Journal of Neuroscience*, 18, 1528–1545.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 443–446. doi:10.1163/156856897X00357
- Cabe, P. A., & Hofman, L. L. (2012). Haptic distal spatial perception mediated by strings: Point of closest approach and bypass distance. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 1328–1340. doi:10.1037/a0027738
- de Groot, A. M. B., Thomassen, A. J. W. M., & Hudson, P. T. W. (1986). Primed-lexical decision: The effect of varying the stimulus-onset asynchrony of prime and target. *Acta Psychologica*, 61, 17–36.
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10(1), 6.1–12. doi:10.1167/10.1.6
- Frisoli, A., Solazzi, M., Reiner, M., & Bergamasco, M. (2011). The contribution of cutaneous and kinesthetic sensory modalities in haptic perception of orientation. *Brain Research Bulletin*, 85, 260–266. doi:10.1016/j.brainresbull.2010.11.011
- Gentaz, E., & Hatwell, Y. (2003). Haptic processing of spatial and material object properties. In Y. Hatwell, A. Streri, & E. Gentaz (Eds.), *Touching for knowing* (pp. 123–159). Amsterdam, The Netherlands: Benjamins.
- Goodwin, A. W., John, K. T., & Marceglia, A. H. (1991). Tactile discrimination of curvature by humans using only cutaneous information from the fingerpads. *Experimental Brain Research*, 86, 663–672.
- Green, B. G. (1982). The perception of distance and location for dual tactile pressures. *Perception & Psychophysics*, 31, 315–323. doi:10.3758/BF03202654
- Greenspan, J. D. (1984). A comparison of force and depth of skin indentation upon psychophysical functions of tactile intensity. *Somatosensory Research*, 2, 33–48.
- Haggard, P., Newman, C., Blundell, J., & Andrew, H. (2000). The perceived position of the hand in space. *Perception & Psychophysics*, 62, 363–377. doi:10.3758/BF03205556
- Hollins, M., & Risner, S. R. (2000). Evidence for the duplex theory of tactile texture perception. *Perception & Psychophysics*, 62, 695–705. doi:10.3758/BF03206916
- Johnson, K. O., & Phillips, J. R. (1981). Tactile spatial resolution. I. Two-point discrimination, gap detection, grating resolution, and letter recognition. *Journal of Neurophysiology*, 46, 1177–1192.
- Kaas, A. L., Stoeckel, M. C., & Goebel, R. (2008). The neural bases of haptic working memory. In M. Grunwald (Ed.), *Human haptic perception: Basics and applications* (pp. 113–129). Basel, Switzerland: Birkhäuser Verlag.
- Kennedy, J. M., Gabbias, P., & Nicholls, A. (1991). Tactile pictures. In M. A. Heller & W. Schiff (Eds.), *The psychology of touch* (pp. 263–299). New York, NY: Psychology Press.
- Klatzky, R. L., Lederman, S. J., & Metzger, V. A. (1985). Identifying objects by touch: An “expert system.”. *Perception & Psychophysics*, 37, 299–302.
- Klatzky, R. L., Loomis, J. M., Lederman, S. J., Wake, H., & Fujita, N. (1993). Haptic identification of objects and their depictions. *Perception & Psychophysics*, 54, 170–178. doi:10.3758/BF03211752
- Kleiner, M., Brainard, D., & Pelli, D. (2007). What’s new in Psychtoolbox-3? *Perception*, 36(ECVP Abstract Supplement).

- Lawson, R., & Bracken, S. (2011). Haptic object recognition: How important are depth cues and plane orientation? *Perception*, 40, 576–597.
- Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: A window into haptic object recognition. *Cognitive Psychology*, 19, 342–368.
- Lederman, S. J., & Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, 71, 1439–1459. doi:10.3758/APP.71.7.1439
- Lederman, S. J., & Taylor, M. M. (1972). Fingertip force, surface geometry, and the perception of roughness by active touch. *Perception & Psychophysics*, 12, 401–408.
- Levy, M., Bourgeon, S., & Chapman, C. E. (2007). Haptic discrimination of two-dimensional angles: Influence of exploratory strategy. *Experimental Brain Research*, 178, 240–251.
- Loomis, J. M. (1992). Distal attribution and presence. *Presence: Teleoperators and Virtual Environments*, 1, 113–119.
- Loomis, J. M., & Lederman, S. J. (1986). Tactual perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Cognitive processes and performance, Vol. 2, pp. 31.1–31.41). New York, NY: Wiley-Interscience.
- Mercer, T., & McKeown, D. (2010). Updating and feature overwriting in short-term memory for timbre. *Attention, Perception, & Psychophysics*, 72, 2289–2303. doi:10.3758/BF03196702
- Miyazaki, M., Hirashima, M., & Nozaki, D. (2010). The “cutaneous rabbit” hopping out of the body. *Journal of Neuroscience*, 30, 1856–1860. doi:10.1523/JNEUROSCI.3887-09.2010
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, 10, 437–442. doi:10.1163/156856897X00366
- Rensink, R. A., & Enns, J. T. (1998). Early completion of occluded objects. *Vision Research*, 38, 2489–2505.
- Shore, D. I., Barnes, M. E., & Spence, C. (2006). Temporal aspects of the visuotactile congruency effect. *Neuroscience Letters*, 392, 96–100.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995. doi:10.3758/s13414-010-0073-7
- von Békésy, G. (1959). Neural funneling along the skin and between the inner and outer hair cells of the cochlea. *Journal of the Acoustical Society of America*, 31, 1236–1249.