This is the peer reviewed version of the following article:
Mitsudo, H. (2015). Inferring the depth of 3-D objects from tactile spatial information. *Attention, Perception, & Psychophysics, 77(4),* 1411-1422, http://dx.doi.org/10.3758/s13414-015-0878-5.
The final publication is available at http://link.springer.com.

Running head: TACTILE SPATIAL INFORMATION AND 3-D SHAPE

Inferring the Depth of 3-D Objects from Tactile Spatial Information

Hiroyuki Mitsudo
Kyushu University

Author Note

Hiroyuki Mitsudo, Faculty of Human-Environment Studies, Kyushu University.

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

Correspondence concerning this article should be addressed to Hiroyuki Mitsudo, Department of Human Sciences, Faculty of Human-Environment Studies, Kyushu University, 6-19-1 Hakozaki, Higasiku, Fukuoka 812-8581, Japan.
E-mail: hmitsudo@lit.kyushu-u.ac.jp

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 the hand, while keeping the hand shape flat. Experiment 1, by using 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 using 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 using a discrimination task, found 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

Inferring the Depth of 3-D Objects from Tactile Spatial Information

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 information[1] 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 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, & Marceglia, 1991), and orientation (Frisoli, Solazzi, Reiner, & Bergamasco, 2011; Levy, Bourgeon, & Chapman, 2007). Given that familiar 3-D objects contain more distinctive parts than 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 that 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 (Figure 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

---

[1] Throughout this manuscript, 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" can also be used to mean the 3-D environment recognized through kinesthetic and proprioceptive inputs (e.g., Haggard, Newman, Blundell, & Andrew, 2000).

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

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 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 inter-element 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 in a fixed size defined by obviously touchable outer parts (or edges and rim), and (b) both of 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, entertainment, 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 separation ($n = 87$). The correlation coefficient was .59 and was significantly different from 0 ($p < .0001$). The results suggest that, at least for man-made objects, the physical depth of a concave part is, to some extent, positively correlated with the separation between touched locations in real-life situations.

Insert Figure 1 about here

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 is correlated with inter-element 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 (Figure 2). Touching a periodic pattern with an inter-element separation of less than 1.25 mm can produce a sense of roughness, even without a lateral hand movement (e.g., Lederman & Taylor, 1972; Hollins & Risner, 2000). The present study used much greater inter-element separations (20-80 mm) and focused on the depth of trenches perceived between the tactile

[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 define 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 circle (e.g., an oval rim or a rectangular hole), the shortest separation between outer parts facing each other was measured. The depth measured here was the shortest distance from the imaginary plane that intersects the touched parts to the deepest position of the bottom part. For containers such as a cup and a box, their contents were excluded during the measurements.

Mitsudo Fig. 1

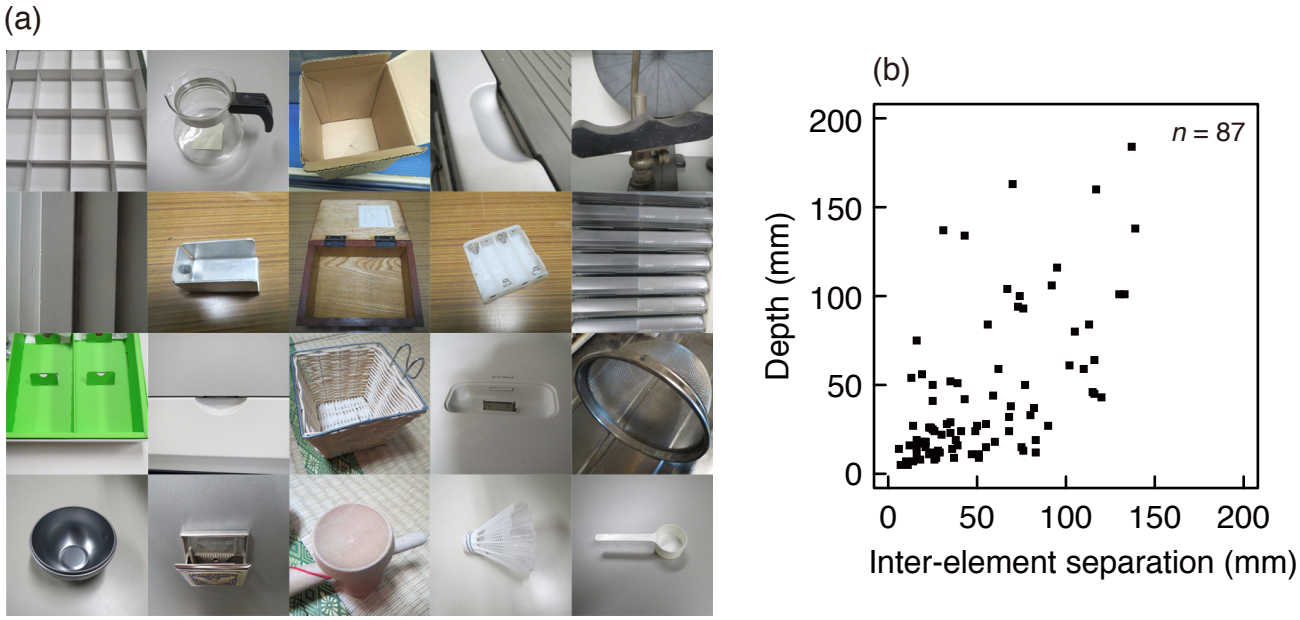

Figure 1. Examples of objects with concave part(s) in real-life situations. (a) Pictures of the objects collected in indoor environments. (b) Correlation between the inter-element separation and the depth of the sampled objects.

bar elements. In the present experiments, participants were asked to touch unseen grating patterns with the hand 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. The participants did not touch the bottom position of the trenches throughout the experiments. 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 will increase as a function of inter-element separation (Experiment 1), influence positional judgments in a haptic working-memory task (Experiments 2 and 3), and influence speeded judgments in a visual object discrimination task (Experiment 4).

Insert Figure 2 about here

Experiment 1

In Experiment 1, participants were asked to estimate the bottom position (i.e., depth) of trenches perceived between the 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 (Figure 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., Green, 1982; Gentaz & Hatwell, 2003), participants' response in this task can be regarded as accurately reflecting the haptically estimated depth. I varied the spatial configuration of the bars (including inter-element separation and overall size, 20-80 mm) and the shape of the bars (round, square) across trials.

Method

Participants. Nine sighted volunteers (aged between 20 and 36 years; seven female, two male) participated in the experiment. No participant reported a sensorimotor deficit. All were right-handed. All 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 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 shape 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 the plate, resulting in *actual* depths of 10 mm at trenches between the bars (Figure 2c). Four types of bar configuration were used. In three conditions where the overall size was 80 mm (the right three panels in Figure 2a), five, three, and two bars were arranged to form three separations between the bars (20, 40, and 80 mm, respectively). In another condition where the overall size was 20 mm, two bars were arranged at the center of the plate (the left panel in Figure 2a).

Mitsudo Fig. 2

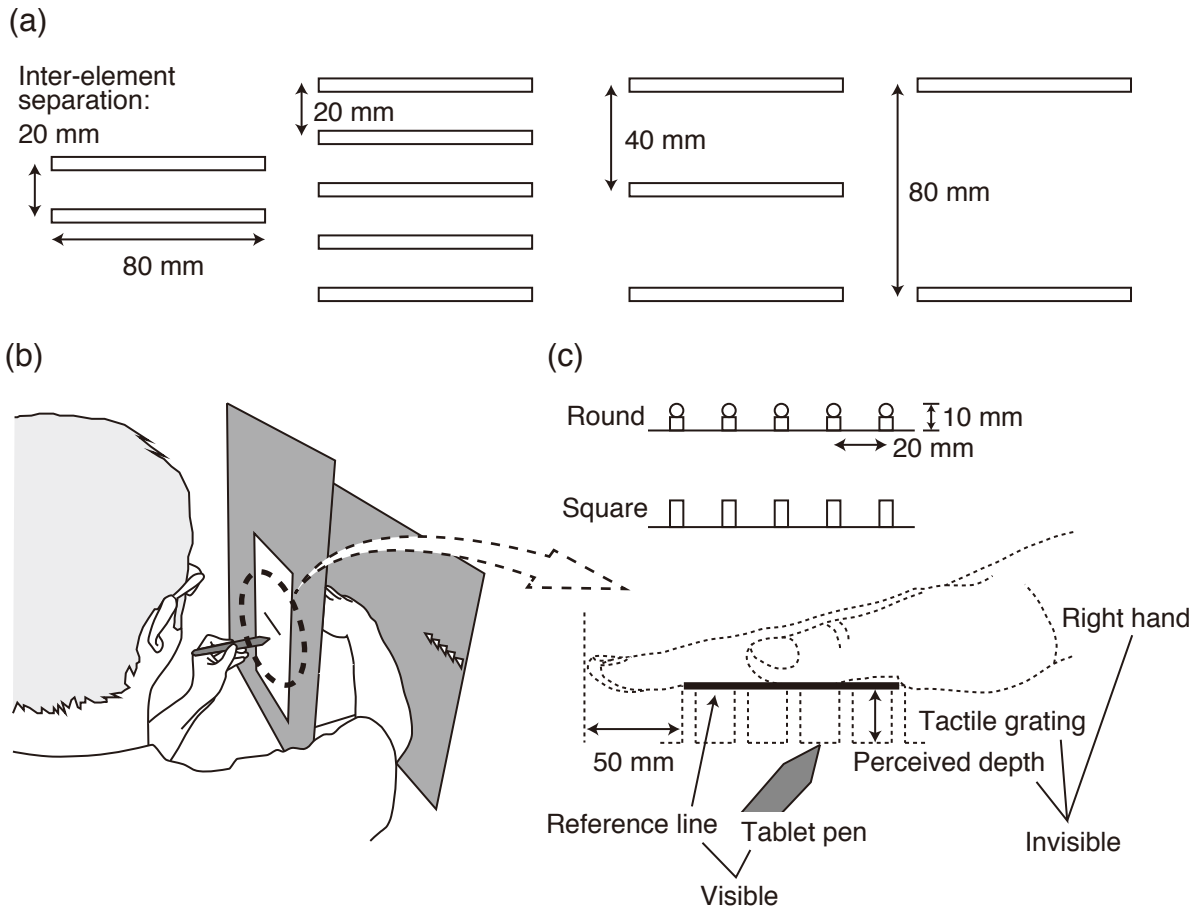

Figure 2. Schematic illustration of tactile stimulus objects and setup for Experiment 1. (a) Top view of the tactile gratings. (b) Overview of the experimental setup. (c) Side view of the gratings and the response apparatus. The upper half represents round- and square-shaped 80-mm-size gratings with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

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

Mitsudo Fig. 3

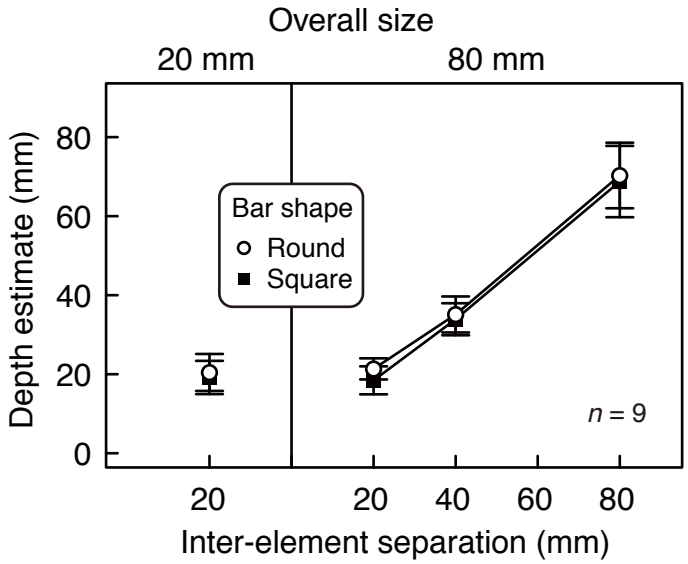

Figure 3. Results for Experiment 1. Mean depth estimates are shown as a function of inter-element separation. Open circles and filled squares represent the depth estimates for the round- and square-shaped bar conditions, respectively. Error bars represent standard error.

comparisons (Ryan's method) revealed that depth estimates were significantly greater with an inter-element 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 inter-element separation and the depth estimates for all the nine participants (ps < .001; mean slope = 0.83; mean adjusted R^2 = .76).

Insert Figure 3 about here

These results suggest that inter-element 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 touched patterns, like a gradient of the pressure from each bar element, has a minor role in the present experiment. The depth estimates were consistently small and similar for the configurations with an inter-element separation of 20 mm, regardless of the number of the bars touched (2 and 5). 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 adopt a response strategy for 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—inter-element separation. To examine this possibility, Experiment 2 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 (Figure 4a). I varied the vertical position of the target (i.e., depth; 26 and 36 mm below the top of the distractor) and distractor type (two gratings with inter-element separations of 20 and 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 (feature overwriting, Mercer & McKeown, 2010; Bancroft & Servos, 2011). 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.

Insert Figure 4 about here

Method

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

Mitsudo Fig. 4

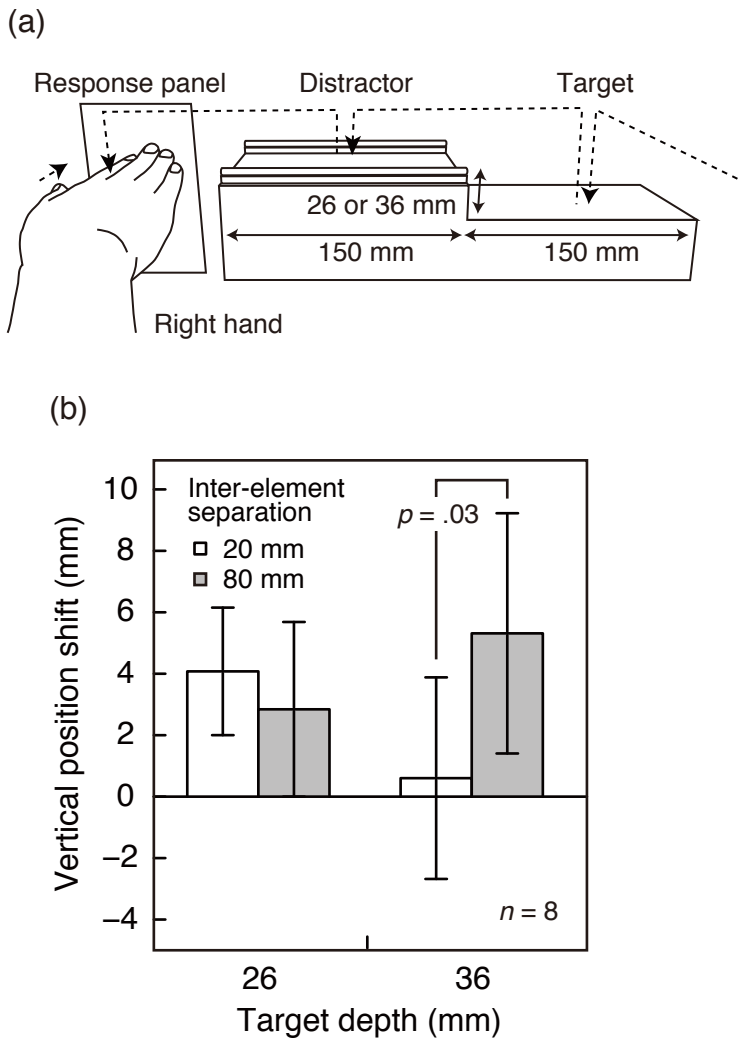

Figure 4. Setup and results for Experiment 2. (a) Schematic illustration of the experimental setup. In this panel, the distractor object has an inter-element separation of 80 mm. Dashed lines approximately represent required hand movements of the right hand during a trial. See text for details. (b) Mean 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 error.

Participants. Nine sighted volunteers (aged between 20 and 36 years; 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 also participated in Experiment 1. Data from one participant were excluded from the analysis because she could not touch the response panel with the middle fingertip reliably.

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 by 18 cm from each other) which indicated the approximate horizontal position 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) in which the depth (vertical position) was variable across trials, 26 and 36 mm below the top of the distractor. Of three distractor patterns, two were grating patterns consisting of the 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 number of bars was five and two, yielding inter-element separations of 20 and 80 mm, respectively. As in Experiment 1, the actual depths of the trenches between the bars were 10 mm. To control the touched position on the hand, small end-panels orthogonal to the fingers (not shown in Figure 4a) were placed 5 cm ahead of the front-ends of the target and the distractor.

The participant's response was recorded through a 7-inch 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.

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 mid-sagittal 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 the 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 eleven 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 tried to remove the effect of touching the distractor itself from the data. Mean reproduced positions for the flat-panel distractor were 22.0 (SE = 8.2) and 26.2 (SE = 7.5) mm at 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 (inter-element 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, a similar magnitude of the downward position shift was 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 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 position for a target depth of 36 mm [$t(71) = 2.5$, $p = .015$, $r = .29$], but not for that of 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 I make a few assumptions. As 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, based on the findings of Baud-Bovy and Viviani (1998), I assume that the retained target positions are higher than the actual ones, e.g., 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 distractor are 20 and 65 mm for inter-element separations of 20 and 80 mm, respectively. Given these assumptions, at a target depth of 36 mm, a downward position shift will 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 will produce a downward position shift for the two 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 is 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. Experiment 3 used target depths of 16 and 46 mm (instead of 26 and 36 mm in Experiment 2). Based on the assumptions made above (i.e., a 16-mm upward shift relative to the actual position), retained target depths are approximately 0 and 30 mm, respectively. If 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 (aged between 20 and 36 years; three female, seven male) participated in the experiment. All were right-handed, except for one. All were naive as to the purpose of the experiment, and no one had participated in Experiment 1 or 2. 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 by the same procedure as that used in Experiment 2 (n = 8). 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 (inter-element 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, a similar magnitude of the downward position shift was observed for the two distractors. As in Experiment 2, I performed two-sided paired t-tests on the pooled position data from 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 that of 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.

Insert Figure 5 about here

Mitsudo Fig. 5

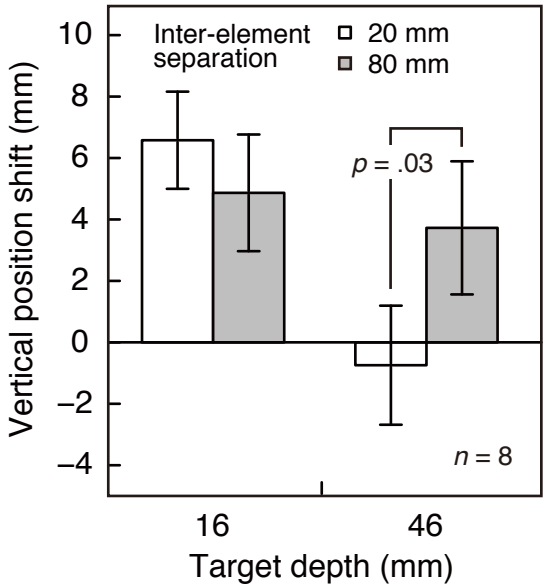

Figure 5. Mean vertical position shifts for the four combinations of target depth and distractor type in Experiment 3. The target depth was different from that of Experiment 2. Positive values indicate that the reproduced position was lower for a grating distractor than for the baseline, flat-panel distractor.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

Experiment 4

The results of Experiments 1-3 suggest that inter-element separation is associated with the estimated depth of untouched parts of 3-D objects. Experiment 4 examined 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 (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 (Figure 6). I used two different object pairs: one consisted of shapes where inter-element separation and trench depth were *correlated* within each object (as in the case of real-life situations), and the other consisted of shapes where inter-element 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 (Figure 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 that were 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 inter-element separation is used to judge the depth of untouched parts of 3-D objects in an automatic manner, the crossmodally congruent tactile patterns will reduce reaction times (RTs) to the visual correlation pair.

Insert Figure 6 about here

Method

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

Participants. A total of 24 sighted volunteers (aged between 19 and 45; 15 female, nine male) participated in the experiment. All were naive as to the purpose of the experiment. Eight 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-inch LCD screen (Apple Cinema Display, 1,680 × 1,050 pixels, frame rate of 60 Hz). Stimulus presentation and data collection were controlled by a personal computer (Apple Macbook) using Matlab with Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli, 2007).

A visual display consisted of two simulated cube-like 3-D objects placed side by side on a simulated rectangle viewed from above (Figures 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 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

Mitsudo Fig. 6

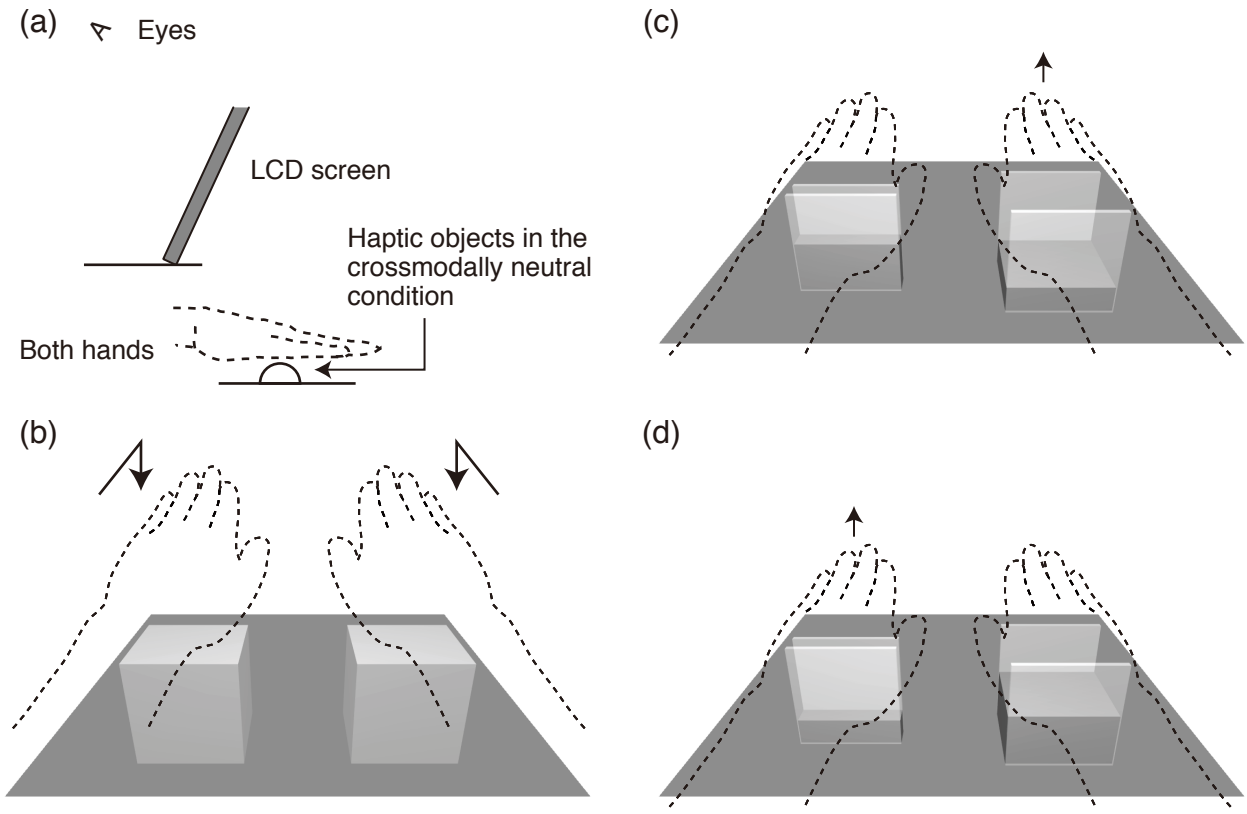

Figure 6. Setup and visual stimuli for Experiment 4. Dashed lines schematically represent both hands that were masked from the participant' s view. (a) Schematic illustration of the side view of the apparatus in the crossmodally neutral condition (i.e., haptic objects were hemisphere objects). (b)-(d) The 3-D objects were the visual stimuli presented. Arrows represent required hand movements during a trial. (b) The starting position of the hands and the visual display before touch. (c) The visual correlation stimulus. (d) The visual anticorrelation stimulus. In (c) and (d), arrows represent 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 text for details.

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 inter-element separation and depth. In a visual correlation stimulus (Figure 6c), one object had an inter-element 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 (Figure 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 object were used. In a crossmodally congruent condition, the top parts of touched objects were the square-shaped gratings used in Experiment 1 and placed at the corresponding position for the two visual objects. The inter-element 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 of 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 position, approximately 10 cm above the tabletop. On the visual screen, two identical, simulated cubes (edge length, 80 mm) were presented side by side (Figure 6b), and a fixation cross (approximately $1 \times 1^{\circ}$ of visual angle) 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 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 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, a "×" 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 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

Median correct RTs and error rates were calculated for each condition and for 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 (10 trials in total, 0.35% all trials). Data were collapsed across the response position (right, left). Because the main concern of Experiment 4 was to see whether crossmodally congruent tactile patterns influenced 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-participant factors, and SOA was a within-participant factor. Figure 7 shows the mean values of the median RTs and error rates, averaged over 12 participants in each condition, and collapsed across the response types.

Insert Figure 7 about here

Reaction Time. For the visual correlation stimulus, the main effect of SOA was significant [$F(1, 20) = 21.1$, $p = .0002$, $\eta^2 = .10$]. The two-way interaction between crossmodal congruency and SOA was 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 SOAs 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 200 ms 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 [$F(1, 20) = 7.3$, $p = .014$; $F(1, 20) = 33.3$, $p < 0.0001$, respectively]. The two-way interaction between crossmodal congruency and response type was significant [$F(1, 20) = 4.7$, $p = .043$, $\eta^2 = .16$].

Mitsudo Fig. 7

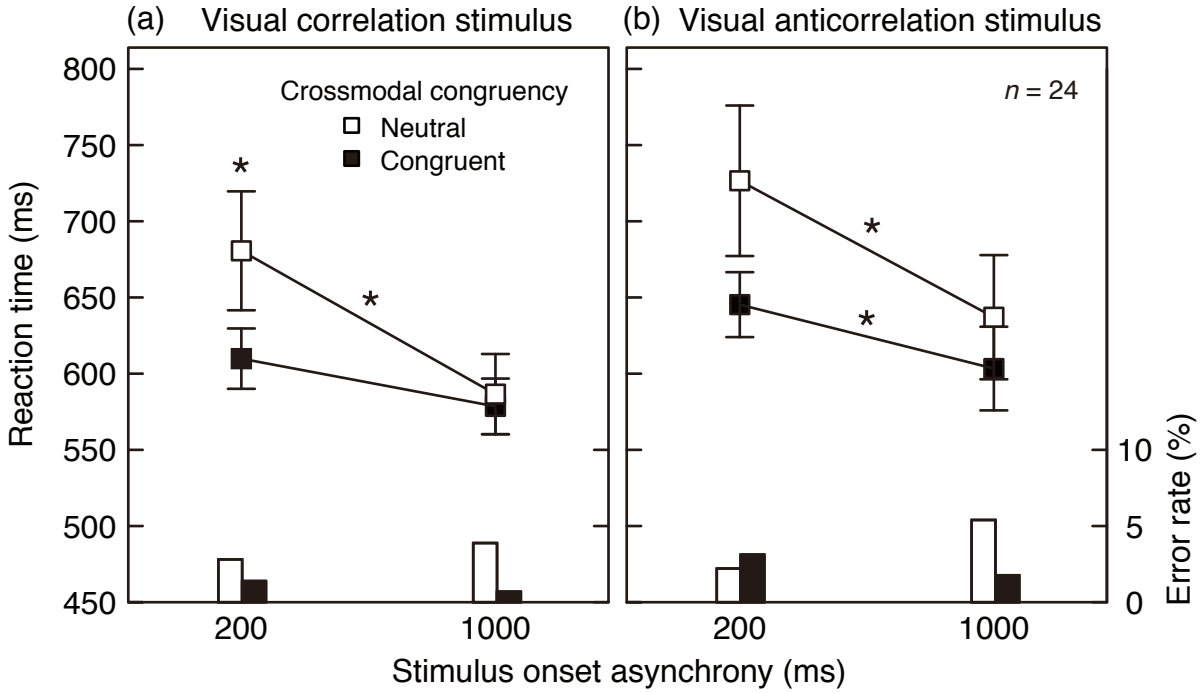

Figure 7. Mean values of median RTs and error rates as a function of SOA in Experiment 4. (a) The visual correlation stimulus; (b) the visual anticorrelation stimulus. Asterisks indicate *p*s < .05. Error bars represent standard error.

Error rate. Three-way ANOVAs were performed on the arcsine-transformed error rates, with the same three factors as those 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 an SOA of 1,000 ms [$F(1, 20) = 4.0$, $p = .060$].

The main results are summarized as: (a) the crossmodally congruent tactile patterns reduced 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 depth 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 (Figure 7b). One might think that this contradicts the above idea because tactile inter-element separation did not "predict" the depth of objects in the visual anticorrelation pair. Note that, to perform the task, the participants would be required to accurately recognize overall 3-D shapes defined by both inter-element separation and depth. Therefore, although uninformative, the crossmodally congruent tactile patterns may act 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 judged shapes are likely to appear in real-life situations, as in the case of the visual correlation pair.

In almost all 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 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., tactile patterns in this experiment) when the onset of a target is temporally separated from that of 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 inter-element separations), but not in the neutral condition (i.e., two hemispheres of the same size). Participants might therefore form 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 visual stimuli with a correlation between inter-element separation and depth.

General Discussion

The results of the four experiments are consistent with the idea that inter-element separation biases participants' interpretation of an untouched concave part of a 3-D object.

The dependency of the estimated depth magnitude on inter-element separation, found in Experiment 1, was somewhat similar to that reported in studies on perceived roughness, whereas the spatial scale was much greater in the present study (inter-element separations of 20-80 mm) than in the previous studies (e.g., less than 1.25 mm, Lederman & Taylor, 1972). Although inter-element 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 depth of an untouched part of a 3-D object.

Several psychophysical studies have investigated 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 showing that humans can pick up information about the spatial layout of nearby objects (that are not directly touched) 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 used 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 (a) to touch the outer edges or elements of the object first, and then (b) to move the hand or fingers into the concave part. In (a), tactile spatial patterns are formed on the skin. I believe that a depth-estimation process like that considered here occurs in (a) and is useful for efficiently preparing a subsequent hand movement in (b).

In the present experiments, although the participants were instructed to keep the hand as flat as possible when touching the gratings, the hand might slightly be flexed toward the untouched bottom plate especially for the gratings with an inter-element separation of 80 mm. Therefore, one could think that hand flexure may 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 inter-element 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 inter-element separations of 20 and 80 mm were evidently greater than 5 mm (> 40 mm on average, Figure 3). Second, in Experiment 2 or 3, no main effect of distractor type (i.e., inter-element separation) was found. Nevertheless, in an attempt to keep the hand flat when touching the different gratings, the participants might make reactive kinematic responses that differed across the configurations. This study does not rule out the possibility that such reactive responses contribute to the estimated depth of the tactile gratings. To examine this issue, it will be helpful to use dynamic touch or more natural hand movements in further studies.

References

Bancroft, T., & Servos, P. (2011). Distractor frequency influences performance in vibrotactile working memory. *Experimental Brain Research, 208*, 529-532.

Barac-Cikoja, D., & Turvey, M. T. (1993). Haptically perceiving size at a distance. *Journal of Experimental Psychology: General*, *122*, 347-370.

Baud-Bovy, G., & Viviani, P. (1998). Pointing to kinesthetic targets in space. *Journal of Neuroscience*, *18*, 1528-1545.

Békésy, G. von (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.

Brainard, D. H. (1997). The Psychophysics Toolbox, *Spatial Vision, 10,* 443-446.

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.

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.

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.

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, Netherlands: John 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.

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.

Hollins, M., & Risner, S. R. (2000). Evidence for the duplex theory of tactile texture perception. *Perception & Psychophysics*, *62*, 695-705.

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., Gabias, 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.

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.

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, Vol. 2. Cognitive processes and performance* (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.

Miyazaki, M., Hirashima, M., & Nozaki, D. (2010). The "cutaneous rabbit" hopping out of the body. *Journal of Neuroscience*, *30*, 1856-1860.

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: Lawrence Erlbaum Associates.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies, *Spatial Vision, 10,* 437-442.

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.