

Visually Guided Haptic Search

Myrthe A. Plaisier, Astrid M.L. Kappers, Wouter M. Bergmann Tiest, and Marc O. Ernst, *Member, IEEE*

Abstract—In this study, we investigate the influence of visual feedback on haptic exploration. A haptic search task was designed in which subjects had to haptically explore a virtual display using a force-feedback device and to determine whether a target was present among distractor items. Although the target was recognizable only haptically, visual feedback of finger position or possible target positions could be given. Our results show that subjects could use visual feedback on possible target positions even in the absence of feedback on finger position. When there was no feedback on possible target locations, subjects scanned the whole display systematically. When feedback on finger position was present, subjects could make well-directed movements back to areas of interest. This was not the case without feedback on finger position, indicating that showing finger position helps to form a spatial representation of the display. In addition, we show that response time models of visual serial search do not generally apply for haptic serial search. Consequently, in teleoperation systems, for instance, it is helpful to show the position of the probe even if visual information on the scene is poor.

Index Terms—Haptic search, psychophysics, perception.

1 INTRODUCTION

TELEOPERATION systems and minimally invasive surgery techniques often involve a combination of haptic (force-feedback) and visual information (e.g., camera images and ultrasound). There are several factors that influence image-guided operations. For instance, to facilitate integration of information from the image and the workspace, the image is often superimposed on the workspace. Integration is even facilitated further if the image is projected in-depth [1]. It has also been shown that performance in image-guided surgery is influenced by the shape of the image aperture and that performance is better if the surgeon controls the camera position manually [2]. It is clear that research into how haptic and visual information are combined is important for optimizing performance through such systems.

In the present study, we aim to provide more insight into how several types of visual feedback influence haptic exploration. One example is visual feedback of finger position. Finger position is important for keeping track of which parts of a scene have already been explored. Of course, finger position can also be perceived through proprioception. Although virtual haptic environments created with force-feedback devices like the PHANTOM (SensAble Technologies) often allow haptic exploration through only a single contact point with the virtual environment, this does not necessarily prevent the user from forming a spatial representation of the virtual environment. It has been shown that a spatial representation

can be established even through short kinesthetic contact [3]. Although spatial representations can be formed through proprioception, spatial representation through vision is usually better [4], [5]. It has been shown that humans integrate information from the visual and haptic modality in a statistically optimal fashion [6]. This means that the modality with the highest accuracy is weighed most heavily in the combined percept. Therefore, we expect that vision will play a dominant role in combined haptic and visual spatial representation. It has also been shown that there is transfer of spatial context from visual to haptic search [7]. In that study, subjects first performed a block of visual search trials, and later, a block of haptic search trials. Some displays in the haptic condition were the same as in the visual condition, while others were new. Subjects were significantly faster when the display had already been shown in the visual condition. Because of these interactions between visual and haptic perception and the fact that a visual spatial representation is usually better than a haptic one, we expect that providing visual feedback of finger position will make haptic spatial exploration more efficient. In this study, we investigate the influence of visual information on haptic exploration and compare haptic search to visual search.

An important difference between visual and haptic exploration of a scene is that, in vision, a spatial representation of the scene is readily available, which can be used to plan, for instance, saccades directly to areas of interest. In haptic exploration, this spatial representation is not readily available. Adding this type of spatial information through visual feedback could therefore facilitate haptic exploration. To study the influence of visual feedback on haptic exploration, a haptic search task was designed to which different types of visual feedback could be added. The haptic search paradigm has been extrapolated from the visual search paradigm. In visual search, subjects typically search for a certain target item (e.g., a red dot) among varying numbers of distractor items (e.g., green dots) presented on a screen. Usually, response times are measured as a function of the number of items on the display, but eye movements can also be recorded. In daily

• M.A. Plaisier, A.M.L. Kappers, and W.M. Bergmann Tiest are with the Helmholtz Institute, Universiteit Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands.

E-mail: {M.A.Plaisier, A.M.L.Kappers, W.M.BergmannTiest}@uu.nl.

• M.O. Ernst is with the Max Planck Institute for Biological Cybernetics, Spemannstrasse 41, 72076 Tuebingen, Germany.

E-mail: marc.ernst@tuebingen.mpg.de.

Manuscript received 27 Apr. 2009; revised 10 July 2009; accepted 25 Aug. 2009; published online 2 Sept. 2009.

Recommended for acceptance by L. Jones.

For information on obtaining reprints of this article, please send e-mail to: toh@computer.org, and reference IEEECS Log Number TH-2009-04-0028. Digital Object Identifier no. 10.1109/ToH.2009.43.

life, the visual modality is not the only modality that is used to perform search tasks. When we try to take our keys out from our pocket or a pen out of our bag, we search using touch. Contrary to visual search, only a few studies have addressed haptic search in the past. Recently, however, the haptic search paradigm has been gaining attention [8], [9], [10], [11], [12], [13].

Although items are normally presented on a screen in visual search studies, there are several different ways to present items for haptic search. One way is by pressing the items onto separate fingers. Items can consist of different types of materials or raised lines [8], [14], [9]. Items can also be 3D shapes fixed in a grid and subjects have to explore the different shapes sequentially [10]. Another way of presenting 3D shapes that does not force subjects to explore the items sequentially is to let subjects grasp a number of shapes simultaneously in the hand [12], [13]. Finally, the way of item presentation most similar to the way this is done in vision is to present items on a surface [11]. Items can consist of, for instance, rough patches on a smooth surface. Such a “tactile display” can be actively explored.

The advantage of using stimuli that are actively explored is that subjects can adjust their exploration strategy in order to optimize their performance. It has been shown that there are typical exploratory procedures (EPs) for extracting object properties [15] and that object recognition can be impaired by constraining the exploratory movements [16]. Analysis of exploratory movements has shown that haptic object recognition is viewpoint-dependent [17], [18]. Thus, characterization of the exploratory movements that subjects make in combination with response times provides insight into the search strategy used. In two previous studies, we have shown the importance of analyzing exploratory strategy for interpreting response times in haptic search tasks [11], [12].

In visual search studies, usually only response times are analyzed to determine which search strategy was used. When the response times do not increase with the number of items in the display, the search strategy is referred to as “parallel” meaning that all items were processed simultaneously. When items are processed one by one, response times increase with the number of items in the display, and the search strategy is referred to as “serial” [19], [20], [21], [22], [23]. We have shown in a previous study that the response time slopes can be very shallow in a haptic search task, while analysis of the exploratory movements that were made clearly indicate that the search strategy was serial [11]. This suggests that visual search models cannot readily be used to distinguish haptic parallel and serial search based on response times alone. As mentioned before, in vision, a spatial representation of the scene is readily available whereas this is not the case in haptics. Adding this type of spatial information could make haptic serial search performance more similar to visual serial search performance.

To investigate how visual information can be used to guide haptic search and which types of visual information are most important for enhancing haptic search efficiency, a haptic display was generated using a force-feedback device. On this display, items were defined by regions with a higher friction coefficient than the background of the display.

Frictional forces were chosen to define the virtual display, because friction is a property present in the real world that is perceived through lateral motion when you move your finger over a certain material [15]. Subjects haptically explored the display with one finger only, ensuring that the task could only be performed in a serial manner. In the different conditions, varying amounts of visual information could be provided. The effects of the different types of visual feedback were compared to simulations of two extreme types of search strategies. For the first strategy, it was assumed that subjects moved from item to item along the shortest pathway; in this case, exploration was efficient and completely guided by item positions. In the second strategy, it was assumed that subjects scanned the whole display with their finger, so exploration was inefficient and completely independent of item positions.

2 METHODS

2.1 Participants

Ten paid subjects (mean age 25 ± 5 years, four male) participated in the experiment. One participant was left-handed, while the others were all right-handed according to Coren’s test [24]. They had normal or corrected-to-normal vision. All subjects were naive as to the purpose of the experiment and gave their informed consent. None of the subjects reported any known hand deficits.

2.2 Apparatus

The setup consisted of a custom-built visuohaptic workbench. The haptic stimulus was presented using a PHANTOM 1.5 A force-feedback device. Subjects placed the index finger of their dominant hand in a thimble-like holder that was connected to the PHANTOM. The visual stimulus was presented on a computer screen. The subjects looked via a mirror onto the screen such that the visual and haptic stimuli were spatially aligned, as illustrated in Fig. 1a. The finger position was recorded at 50 Hz by sampling the position of the thimble-like holder as a single point in space.

2.3 Stimuli

The haptic working range was restricted to the size of the haptic display ($15 \text{ cm} \times 15 \text{ cm}$) in the horizontal plane. Subjects could not move outside of the haptic display, and the edges of the display felt like a continuous wall. The working range was restricted in height such that subjects could raise their finger 4 cm upward from the display plane, but they were instructed not to lift their finger at all. On the square surface of the haptic display, items consisting of circular areas (1.6 cm diameter) with an increased friction coefficient were placed at random positions (the edges of the items were at least 1.6 cm apart and 1 cm from the boundaries of the display). Both the static and dynamic friction coefficients of the display background were set to 0.2, while distractor items had friction coefficients of 0.5 and the target had both friction coefficients set to 0.8. There could be three, five, or seven items on the display.

There were four different visual conditions, but the haptic display was always defined in the same way. The visual display was represented with a blue square while

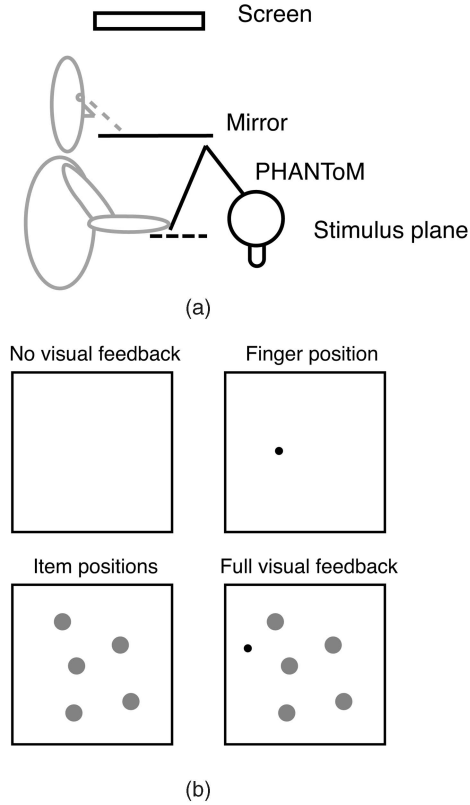


Fig. 1. (a) Illustration of the setup. Subjects placed their index finger in the holder connected to the force-feedback device that was used to create the haptic display. They viewed the visual stimulus via a mirror projecting the image onto the same plane as the haptic stimulus. (b) Examples of the visual display in each of the four conditions.

items were indicated with light-colored disks and finger position with a small sphere. In the first condition, only the square representing the display was shown on the display ("No visual feedback" condition); in the second condition, the square was shown together with the finger position ("Finger position" condition); in the third condition, only the square and the item positions were shown ("Item positions" condition); in the last condition, the square, the item positions, and finger positions were shown ("Full visual feedback" condition). The different conditions are shown in Fig. 1b. Note that there was never visual information present on which item was the target item.

2.4 Experimental Design

Subjects were instructed to indicate as fast as possible and accurately whether or not a target item was present. They were informed that the friction coefficient of the target and distractor items would be constant throughout the experiment and that there could at most be one target on the display. They were also told that, like in reality, frictional forces depended on the amount of downward pressure. Responses were made through key presses using keys that were situated next to each other on the keyboard ("f" and "g" keys). The response key on the left side corresponded to "yes" and the key on the right side corresponded to "no." To help subjects remember which button corresponded to which answer, the words "yes" and "no" were shown to the left and right of the visual display, respectively. After pressing a response key, feedback on whether the answer

was correct was shown on the screen. Subjects explored the display with their dominant hand, and answered with the other hand. Before the next trial started, subjects moved their finger to the starting position in the upper left corner of the display. During this period, the finger position was shown regardless of the experimental condition.

All conditions were performed in separate blocks of trials. Prior to the experiment, subjects performed a single block of training trials in the full visual feedback mode until they were comfortable with the task and it was clear that they had understood the task. Then, prior to each block of trials, they performed at least 20 training trials in the experimental condition of that block. Trials were continued until nine out of 10 were answered correctly. On average, subjects performed 25 ± 9 training trials and the maximum number of training trials that was needed was 52. Each subject performed all four conditions in a roughly counter-balanced order. Each block consisted of 60 trials (20 trials per number of items) in random order. In half of the trials, a target item was present. After 30 trials, there was a 5-minute break. The blocks of trials were performed on separate days. Trials that were answered incorrectly were repeated at the end of the block until all trials were answered correctly. If a repeated trial was answered incorrectly, then this trial would be repeated again (but this only happened in 25 of the total of 2,400 trials). This ensured that there were 10 correctly answered trials for each number of items in each experimental condition. Only the trials that were answered correctly were included in the analysis. Error rates were calculated as the percentage of correctly answered trials of the total number of performed trials.

3 RESULTS

The results consist of response times, error rates, and recorded movement tracks. The error rates were well below chance level for all subjects in each of the conditions; statistical analysis (repeated measures ANOVA) of the error rates did not show an effect of condition. There were false negatives (9 percent of all trials) as well as false positives (6 percent of all trials). Only correct trials were included in the analyses that follow. Fig. 2 shows a representative selection of tracks over the display of one subject in each of the four conditions. For each condition, two target-present and two target-absent trials are shown. It can be seen that there is a clear strategy difference between the conditions in which there was no visual feedback of the item locations compared to the two conditions in which this information was present. In the first case, the subjects systematically scanned the whole display, whereas in the second case, exploratory movements concentrated around the item positions. It is clear from the tracks in the "Item positions" condition that the subject could use visual information about the item positions without visual feedback of the finger position.

3.1 Time Spent Touching the Edges

Fig. 2 suggests that in both conditions without visual feedback on item positions ("No visual feedback" versus. "Finger position"), subjects touched the edges of the display more often when feedback on finger position was absent.

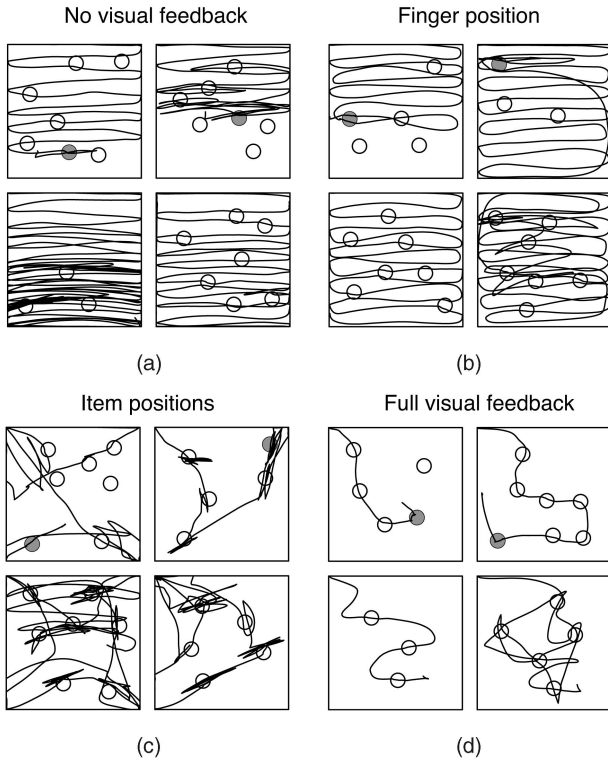


Fig. 2. A selection of four tracked movements over the display from the same subject in (a) no visual feedback condition, (b) finger position condition, (c) item position condition, and (d) full visual feedback condition. In each panel, the top two tracks were target present trials, while the bottom two were target absent trials.

The same holds comparing both conditions with visual feedback on items positions (“Item positions” versus. “Full visual feedback”). Fig. 3 shows the percentage of time that the subject spent touching the edges of the display (i.e., finger positions at 2 mm or less from the edges) for each condition. Statistical analysis of the percentages of the duration of a trial that subjects were touching the edges showed that there was an effect of condition (repeated measures ANOVA, $F(3,27) = 15.7$, $p < 0.001$). To determine whether there was an effect of the presence of visual feedback on how much time subjects were touching the edges, posthoc t -tests were performed to compare the “No visual feedback” to the “Finger position” condition and to compare the “Item positions” condition to the “Full visual feedback” condition. This analysis showed that the proportion of time subjects touched the edges in the “No visual feedback” condition was significantly larger than in the “Finger position” condition ($t = 5.3$, $p < 0.001$) and also significantly larger in the “Item positions” condition than in the “Full visual feedback” condition ($t = 2.4$, $p = 0.04$).

3.2 Response Time Slopes

Fig. 4 shows the response times as a function of the number of items for target-present and target-absent trials in each of the conditions. For the two conditions without visual information about the item positions, the slopes are not significantly different from zero ($p > 0.05$). There is a difference in offset as the target-absent trials yield larger response times than the target-present trials. For the conditions with visual feedback of target positions, the target-absent slope was significantly different from zero. The value is indicated in the figure. Both

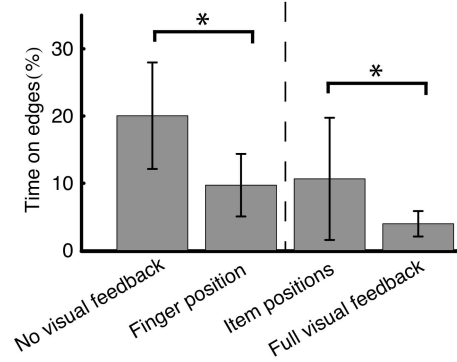


Fig. 3. Percentage of the duration of a trial subjects spent at distances smaller than 2 mm to the edges of the display, averaged over all subjects for each condition. The error bars indicate the standard deviation of the single-subject means. An asterisk indicates that the difference was significant.

the target-present and absent slopes were significantly different from zero for the “Full visual feedback” condition; the slope values are indicated in the figure and the ratio between the target-absent and target-present slopes in this last condition was 1.5.

3.3 Strategy Analysis

In Fig. 5, the distribution of the distances from the sampled finger position to the center of the nearest item combined is shown for all subjects combined. These distributions can be interpreted as probability density functions of the chance that a finger position was sampled at a certain distance from an item. The bars at distances smaller than the item radius (to the left of the dashed line) represent the time that subjects spent on items. The remainder of the distribution represents the parts of the trials where subjects were moving on the background of the display. It can be seen that this part of the distribution centers on smaller distances for the conditions with visual feedback of item position than for the conditions without this feedback. This means that subjects spent a larger portion of time moving relatively far away from items when visual feedback of item locations was absent than when this feedback was present.

To analyze the differences between the conditions, the distributions were split into distances smaller and larger than the item radius. From the distances smaller than the item radius, the percentage of time that subjects touched an item was calculated (see Fig. 6a). A large percentage indicates that subjects spent relatively little time on the display background, indicating well-directed movements toward the items. The largest percentage of time was found for the “Full visual feedback” condition. The distributions of the distances larger than the item radius were analyzed in terms of the mean and the kurtosis, which are shown in Figs. 6b and 6c. A smaller mean indicates that subjects moved, on average, closer to an item. The kurtosis is a measure for how heavy the peak in the distribution is; a large value means that a large portion of data was located near the peak and less in the flanks (for comparison: the normal distribution has a kurtosis of 3). Percentage of time on an item, mean distance to an item, and kurtosis were calculated from the distributions

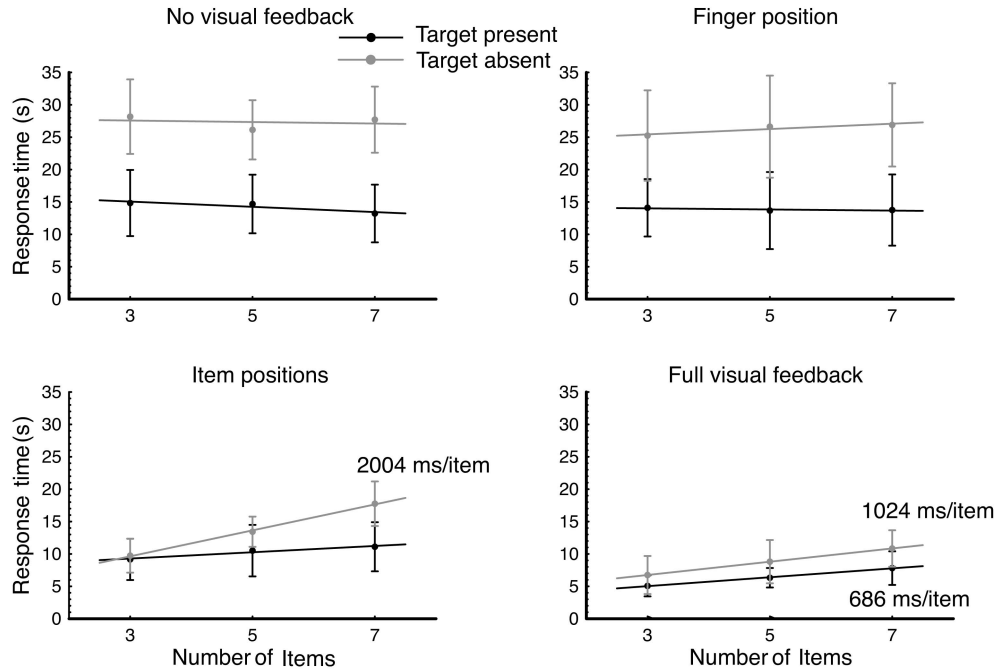


Fig. 4. Response times averaged over subjects as a function of the number of items for target-present and target-absent trials. Error bars indicate the standard deviations of the subject means. Solid lines represent linear regression to the mean response times. Slope values are indicated for significant slopes only ($R^2 > 0.9$, $p \leq 0.03$).

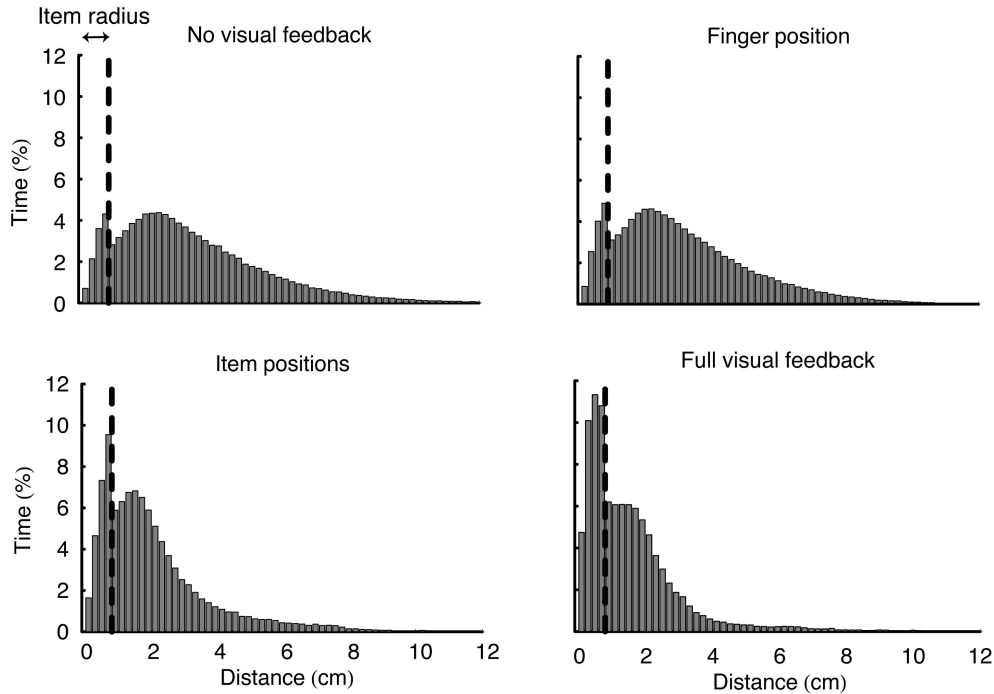


Fig. 5. Distributions of the distances from the sampled finger position to the center of the nearest item from all subjects in each of the four conditions. The dashed line indicates the item radius (8 mm).

from each subject. Repeated measures MANOVA was performed on these three measures (Pillai's trace, $F(9, 81) = 13$, $p \leq 0.001$). Follow-up analysis using univariate tests (ANOVAs) showed that there was a main effect for each measure ($F(1.6, 18) \geq 20$, $p \leq 0.001$, Greenhouse-Geisser correction was used when appropriate). Posthoc Bonferroni-corrected t -tests showed that most differences between the conditions were significant ($p \leq 0.02$). The

nonsignificant differences between conditions are indicated in Fig. 6. These results show that when visual feedback of item positions was provided, subjects spent a larger portion of time touching items and less time moving in between items. Furthermore, subjects moved at smaller distances to items and spent less time at distances far away from items when visual feedback of item location was provided.

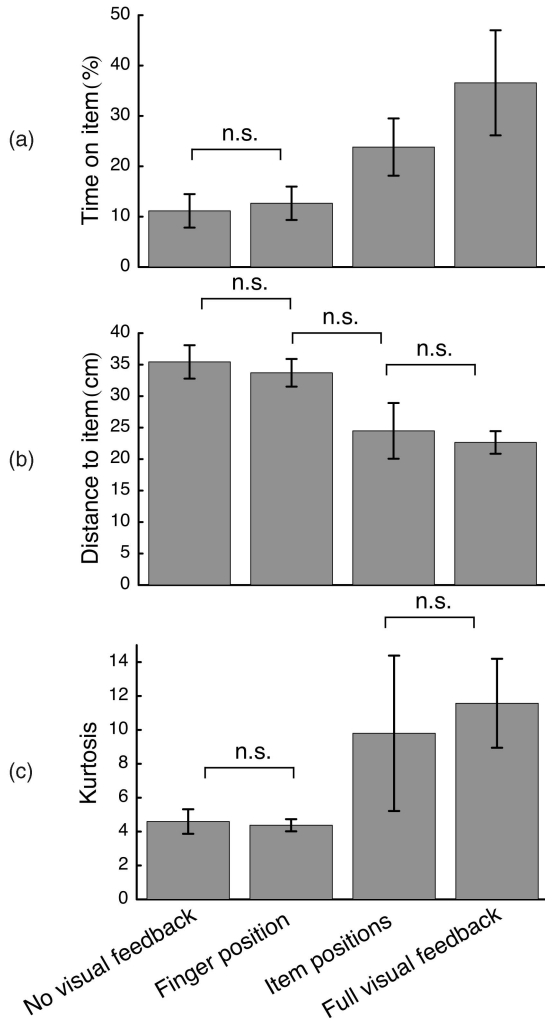


Fig. 6. (a) Percentage of time on an item, (b) distance to the nearest item, and (c) the kurtosis of the distribution of sampled distances to the nearest item for each of the conditions averaged over subjects. The error bars indicate the standard deviation of the single subject means and nonsignificant differences are indicated.

3.4 Global versus Local Exploration

When there was no visual information about item positions, subjects scanned the display systematically with their finger. It is possible that they returned to previously visited items after scanning the whole display. To investigate whether subjects did this and whether they were able to use a spatial representation of the items on display, the tracks from the conditions in which there was no visual feedback of item positions were divided into two parts. To this end, the display was divided into an 8×8 grid. Consequently, the grid elements had a height and width of 1.9 cm. This size was of the order of the diameter of an item (1.6 cm), because it can be expected that subjects made scan paths approximately an item diameter apart. Decreasing grid size increases the chance that subjects did not visit a certain element during a trial, while they did search the whole display, which is not desirable. The second part of the track was defined from the moment that all elements in the grid were visited at least once, because, from that moment, subjects started exploring previously explored parts of the display again. The remaining part of track had to be at least

2 seconds long to be considered as a second part of the track. Scanning direction differed between subjects, but also between trials and even within a trial. This way of defining the track parts works regardless of the subjects' scanning direction. Not all trials had a second part, as subjects could answer when they had found a target or immediately after scanning the whole display. Trials without a second part were not included in the analysis. There was a second part in 20 percent of the trials in the "No visual feedback" condition and in 31 percent of the trials from the "Finger position" condition. Fig. 7a shows examples of a track with two parts for the "No visual feedback" condition and for the "Finger position" condition. It can be seen that particularly in the "Finger position" condition, well-directed movements toward previously touched items were made during the second part of the trial. In the "No visual feedback" condition, this was not as clearly the case, although in the bottom-left panel, it can be seen that the subject had a rough idea of where in the display the items were located. The distributions in time of distances from the sampled finger position to the nearest item for the first (light bars) and the second part (dark bars) of trials for all subjects combined are shown in Fig. 7b. It can be seen that the distributions from the two parts differ mainly in the "Finger position" condition. The peak from the distribution of the second part is shifted toward smaller distances from items with respect to the peak of the first part. Also, the distribution from the second part of the trials decreases faster for distances far away from items than the distribution from the first part of the trials. This suggests that there was a difference in exploratory strategy between the first and the second part of the trial.

For the distributions from each subject, the percentage of time on an item, mean distance to an item, and kurtosis were calculated for the two parts of the trials. Fig. 7c shows these measures averaged over all subjects. Significant differences between the first and the second part are indicated with an asterisk (paired samples t -tests, $t \geq 2.7, p \leq 0.02$). There were only significant differences between the first and second part in the "Finger position" condition. In this condition, subjects spent a larger proportion of exploration time on items than in the first part. Furthermore, on average, they moved at a smaller distance to items, and in combination with the larger kurtosis, this indicates that they spent more time near items than further away from items than in the first part of the trial. In Fig. 7a, it can be seen that, sometimes, subjects were still systematically scanning after all grid elements were visited. Note that this does make the distributions of the two parts more similar rather than dissimilar. So, the significant differences between the different distributions cannot be due to the criterion we used for splitting up the movement tracks.

4 SIMULATIONS

Simulations of two extreme search strategies were performed, representing the most efficient and most inefficient strategy. "Guided search" assumed that the subjects moved with constant speed (corresponding to a movement speed of 10 cm/second or position being sampled every 2 mm at 50 Hz sampling rate) to the nearest untouched item along the shortest pathway. Search was terminated when a target was

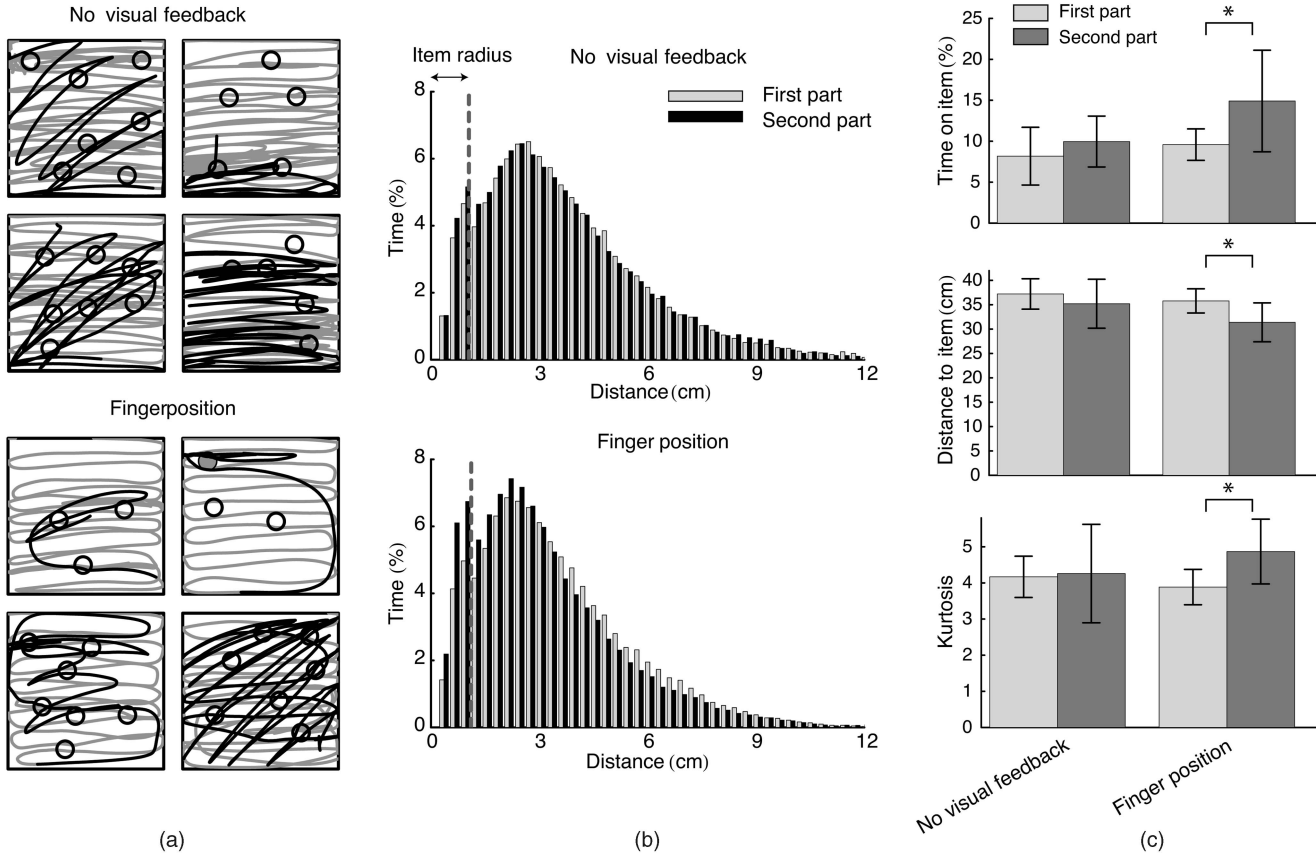


Fig. 7. (a) Four examples of trials with a second part for the “No visual feedback” and the “Finger position” conditions. The first part of the track is shown in gray and the second part is shown in black. A target item is indicated with a filled disk. All examples were trials from the same subject. (b) Distributions of the distances from the sampled finger position to the nearest item from all subjects for the two stages of the conditions without visual feedback of item positions. (c) Time on an item, average distance to an item, and kurtosis of these distributions. Error bars represent the standard deviation of the single-subject means, and an asterisk indicates a significant difference.

found or when all items were visited. The resulting distribution of distances to the nearest item is shown in Fig. 8a.

The most inefficient search strategy would be when the whole display is searched, regardless of the item positions. Note that, in this case, the distribution of the distances is completely driven by the distribution of the items on the displays. The chance that a random point on the display is located at a certain distance from an item is not equal for all distances. For instance, the chance that a random point is very far from an item is quite small. Therefore, in the simulation labeled “Unguided search,” 2,500 positions were homogeneously distributed over each display in the set (at 50 Hz sampling rate, this would correspond to a response time of 50 seconds) and the distance from each position to the nearest item was calculated. The resulting distribution is shown in Fig. 8b. As the item positions were carefully randomized, the distributions of the distances did not differ significantly for the sets of displays from the different conditions.

Comparison of the distributions of the sampled finger positions (Fig. 5) to that of the simulations (Fig. 8) shows that the distribution from the “No visual feedback” and the “Finger position” condition resemble the “Unguided search” simulation, while the “Full visual feedback” and “Item positions” conditions are most similar to the “Guided search” simulation. So, when visual feedback of item

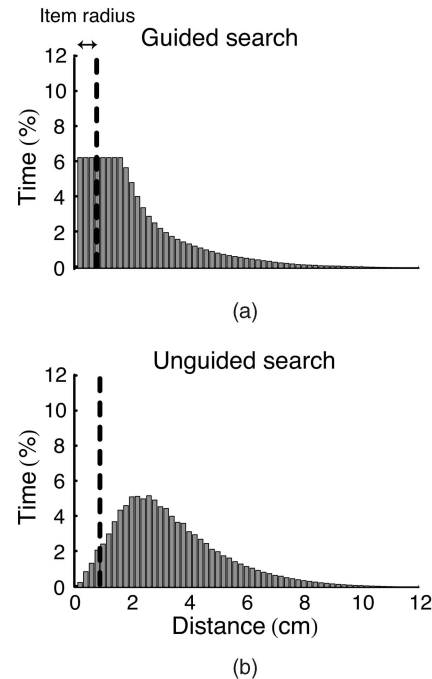


Fig. 8. Distributions of the distances from the simulated finger position to the nearest item from all displays in (a) “Guided search” strategy and (b) “Unguided search” strategy.

locations was present, subjects used a search strategy most similar to the “Guided search” strategy. Thus, when this feedback was not present, subjects used a strategy similar to the “Unguided search” strategy. For the “Finger position” condition, it was found that movements during the second part of trials were, on average, at distances closer to items and had a larger kurtosis than movements during the first part. This shows that the second stage of exploration was shifted toward the “Guided search” strategy. This indicates that exploration during the second part of trials in this condition was more similar to the conditions with visual feedback on the item positions than the first part. This was, however, not the case for the “No visual feedback” condition.

In all experimental conditions in Fig. 5, there is a peak for distances smaller than the item radius, which indicates that subjects spend relatively more time on an item. This peak is absent in the simulations, because a constant movement speed was assumed without distinction between movement on an item or on the background. In the “Unguided search” strategy, the chance that a simulated point was close to or on the center of an item is very small. However, in the “Guided search” strategy, it was assumed that movements were made to the center of an item and then to the next. This explains why the flat part of the distribution ranges beyond the item radius.

5 DISCUSSION AND CONCLUSIONS

In the present study, we show that adding visual information strongly influences haptic exploratory strategy. In the absence of visual information about item positions, subjects systematically scanned the whole display; when information about item positions was added, exploratory movements concentrated around the item positions. Studies into spatial representation have shown that spatial locations and layout can be learned through proprioception. It has been shown that subjects can quite accurately return to a certain target position that has only been briefly touched before [3]. This indicates that the representation of spatial location through proprioception is fairly good. This is in agreement with our finding that subjects can use visual information about item positions in the absence of visual feedback of finger position.

In both conditions without visual feedback of finger position (“No visual feedback” and “Item location”), subjects tended to touch the edges more often than when feedback of finger position was present. The tracks over the display also show that they sometimes followed the edges of the display in this condition. Although it is possible that subjects touched the edges in conditions without visual feedback of finger position because they simply overshot their movement, the fact that subjects often moved along the edges before moving to the next item suggests that subjects used the edges as a reference to re-calibrate their finger position. It has been shown that here there is indeed an advantage for creating a spatial representation if an external reference frame (like a bounding square) is provided [25].

Our data from the conditions without visual feedback of item positions show that subjects, sometimes, used a two-stage exploratory strategy. First, the whole display was

scanned and then subjects explored parts of the display again. Lederman and Klatzky have shown that such a two-stage strategy of global exploration followed by local exploration is often present in haptic exploration [26]. An object’s shape, for instance, can be explored globally by enclosure, followed by a local exploration procedure like contour following. Interestingly, when there was visual feedback of finger position, exploration in the second stage was clearly different from the first stage. Subjects spent a larger proportion of exploration time on items and moved at distances closer to items. This indicates that subjects had built a spatial representation of the item positions in the display during the first stage and could use this representation to move efficiently back to areas of interest during the second stage. This made the exploratory strategy during the second stage more similar to the strategy used in the conditions with visual feedback on item positions. When there was no feedback of finger position, however, exploratory movements were not correlated more closely to item positions in the second part than in the first part of the track. This shows that forming and using a spatial representation of the display was facilitated by providing visual feedback of finger position.

Because spatial representations can be formed through proprioception alone, the question arises why visual feedback of finger position was required. It has been suggested that visual spatial learning is easier, because in this modality, cues like walls of a room that provide a reference frame are readily available [27]. In another study that was mentioned earlier, it was shown that spatial learning can be aided by providing an external reference frame [25]. In that case, subjects explored a map with one hand while touching the external reference frame with the other. In this way, subjects could easily keep track of the position of the exploring finger with respect to the reference frame. In the “Finger position” condition of the present study, the boundaries of the display and the finger position could be viewed simultaneously; therefore, the position of the finger relative to the display boundaries could also be easily extracted. When the finger position was not shown, extracting this information was much more difficult. This could explain why subjects were able to use a spatial representation of the display in the second stage of exploration in the “Finger position” condition, but not in the “No visual feedback” condition. Note, however, that in the present study, subjects were not instructed to learn the spatial layout of the display. Therefore, our results do not mean that the layout of the displays could not be learned through proprioception alone. If the subjects were instructed to, they might possibly have been able to do so. Rather, our results show that during a search task, subjects returned to locations where they had previously felt something quite accurately when visual feedback of finger position was provided. It is likely that also in the “No visual feedback” condition, a spatial representation formed during scanning, but probably a much less accurate one than when visual feedback of finger position was available.

Analysis of the response times as a function of the number of items showed that response times were relatively constant for the conditions without visual feedback of item positions. Search strategy was essentially serial in each of the experimental conditions, but in visual search,

a flat response time slope is usually interpreted as parallel search. This shows that visual search models cannot readily be used to interpret haptic response time slopes. Search strategy analysis showed that there was serial self-terminating search comparable to visual search only when a spatial representation of the display was available. Therefore, in haptic search tasks, it is usually important to also analyze the exploratory strategy that was used when interpreting response time slopes [11].

Summarizing, visual feedback of item locations could be used to efficiently move from item to item. When this feedback was absent, subjects systematically scanned the whole display. When visual feedback of finger position was provided, they could use the scanning stage to build a spatial representation of the display and move efficiently to items after scanning the whole display. Furthermore, when visual feedback of finger position was absent, subjects used the edges to calibrate their finger position. Finally, response time models from visual search are only applicable to haptic search when a spatial representation of the display is readily available.

Concluding, in teleoperation systems, it is clearly most desirable to have full visual feedback, but this may not always be possible as the camera image might be blurred due to fog, for instance. Our results show that providing either visual feedback of finger position only or feedback of item positions can guide haptic exploration. Consequently, in teleoperation systems, visual information on the scene can be used to guide exploration even when the probe is not visible. On the other hand, there is also an advantage of showing the position of the probe even if visual information on the scene is poor because the camera image is blurred.

ACKNOWLEDGMENTS

This research was supported by a grant from The Netherlands Organisation for Scientific Research (NWO) for Astrid M.L. Kappers, a grant from Immersence for Marc O. Ernst, and a grant from IEEE's Technical Committee on Haptics (Student Exchange Program for Cross-Disciplinary Fertilization) for Myrthe A. Plaisier.

REFERENCES

- [1] B. Wu, R.L. Klatzky, D. Shelton, and G.D. Stetten, "Psychophysical Evaluation of In-Situ Ultrasound Visualization," *IEEE Trans. Visualization and Computer Graphics*, vol. 11, no. 6, pp. 684-693, Nov./Dec. 2005.
- [2] P.R. DeLucia, R.D. Mather, J.A. Griswold, and S. Mitra, "Toward the Improvement of Image-Guided Interventions for Minimally Invasive Surgery: Three Factors that Affect Performance," *Human Factors*, vol. 48, pp. 23-38, 2006.
- [3] R.L. Klatzky and S.J. Lederman, "Representing Spatial Location and Layout from Sparse Kinesthetic Contacts," *J. Experimental Psychology: Human Perception and Performance*, vol. 29, pp. 310-325, 2003.
- [4] S. Cashdan, "Visual and Haptic Form Discrimination Under Conditions of Successive Stimulation," *J. Experimental Psychology*, vol. 76, pp. 215-218, 1968.
- [5] P. Worchel, "Space Perception and Orientation in the Blind," *Psychological Monographs*, vol. 65, no. 15, pp. 1-28, 1951.
- [6] M.O. Ernst and M.S. Banks, "Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion," *Nature*, vol. 415, pp. 429-433, 2002.
- [7] T. Nabeta, F. Ono, and J.I. Kawahara, "Transfer of Spatial Context from Visual to Haptic Search," *Perception*, vol. 32, pp. 1351-1358, 2003.
- [8] S.J. Lederman and R.L. Klatzky, "Relative Availability of Surface and Object Properties during Early Haptic Processing," *J. Experimental Psychology: Human Perception and Performance*, vol. 23, pp. 1680-1707, 1997.
- [9] K.E. Overvliet, K.M. Mayer, J.B.J. Smeets, and E. Brenner, "Haptic Search Is More Efficient When the Stimulus Can Be Interpreted as Consisting of Fewer Items," *Acta Psychologica*, vol. 127, pp. 51-56, 2008.
- [10] K.E. Overvliet, J.B.J. Smeets, and E. Brenner, "The Use of Proprioception and Tactile Information in Haptic Search," *Acta Psychologica*, vol. 129, pp. 83-90, 2008.
- [11] M.A. Plaisier, W.M. Bergmann Tiest, and A.M.L. Kappers, "Haptic Pop-Out in a Hand Sweep," *Acta Psychologica*, vol. 128, pp. 368-377, 2008.
- [12] M.A. Plaisier, W.M. Bergmann Tiest, and A.M.L. Kappers, "Salient Features in Three-Dimensional Haptic Shape Perception," *Attention, Perception and Psychophysics*, vol. 71, no. 2, pp. 421-430, 2009.
- [13] M.A. Plaisier, I.A. Kuling, W.M. Bergmann Tiest, and A.M.L. Kappers, "The Role of Item Fixation in Haptic Search," *Proc. Third Joint EuroHaptics Conf. and Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 417-421, 2009.
- [14] S.J. Lederman, R.A. Browse, and R.L. Klatzky, "Haptic Processing of Spatially Distributed Information," *Perception and Psychophysics*, vol. 44, pp. 222-232, 1988.
- [15] S.J. Lederman and R.L. Klatzky, "Hand Movements: A Window into Haptic Object Recognition," *Cognitive Psychology*, vol. 19, pp. 342-368, 1987.
- [16] S.J. Lederman and R.L. Klatzky, "Haptic Identification of Common Objects: Effects of Constraining the Manual Exploration Process," *Perception and Psychophysics*, vol. 66, pp. 618-628, 2004.
- [17] F.N. Newell, M.O. Ernst, B.S. Tjan, and H.H. Bühlhoff, "Viewpoint Dependence in Visual and Haptic Object Recognition," *Psychological Science*, vol. 12, pp. 37-42, 2001.
- [18] M.O. Ernst, C. Lange, and F.N. Newell, "Multisensory Recognition of Actively Explored Objects," *Canadian J. Experimental Psychology*, vol. 61, pp. 242-253, 2007.
- [19] A. Treisman and G. Gelade, "A Feature-Integration Theory of Attention," *Cognitive Psychology*, vol. 12, pp. 97-136, 1980.
- [20] K.R. Cave and J.M. Wolfe, "Modeling the Role of Parallel Processing in Visual Search," *Cognitive Psychology*, vol. 22, pp. 225-271, 1990.
- [21] J.M. Wolfe, "Guided Search 2.0: The Upgrade," *Proc. Human Factors and Ergonomics Soc.*, vol. 2, pp. 1295-1299, 1993.
- [22] J. Duncan and G.W. Humphreys, "Visual Search and Stimulus Similarity," *Psychological Rev.*, vol. 96, pp. 433-458, 1989.
- [23] J. Theeuwes, "Visual Selective Attention: A Theoretical Analysis," *Acta Psychologica*, vol. 83, pp. 93-154, 1993.
- [24] S. Coren, *The Left-Hander Syndrome: The Causes and Consequences of Left-Handedness*. Vintage Books, 1993.
- [25] S. Millar and Z. Al Attar, "External and Body-Centered Frames of Reference in Spatial Memory: Evidence from Touch," *Perception and Psychophysics*, vol. 66, pp. 51-59, 2004.
- [26] S.J. Lederman and R.L. Klatzky, "Haptic Classification of Common Objects: Knowledge-Driven Exploration," *Cognitive Psychology*, vol. 22, pp. 421-459, 1990.
- [27] N. Yamamoto and A.L. Shelton, "Path Information Effects in Visual and Proprioceptive Spatial Learning," *Acta Psychologica*, vol. 125, pp. 346-360, 2007.



Myrthe A. Plaisier received the MSc degree in experimental physics from Utrecht University in 2006. Currently, she is working toward the PhD degree in the Department of Physics and Astronomy at Utrecht University and is with the Human Perception Group at the Helmholtz Institute. Her project includes haptic and visual search as well as numerosity judgement.



received the prestigious VICI grant. She is a member of the editorial boards of *Acta Psychologica* and *Current Psychology Letters* and an associate editor of the *IEEE Transactions on Haptics*.



haptic counting, and haptic perception of volume, mass, and material properties such as roughness, thermal conductance, friction, compliance, and viscosity.

Astrid M.L. Kappers received the PhD degree from Eindhoven University of Technology. She studied experimental physics at Utrecht University, The Netherlands. Since 1989, she has been with the Department of Physics and Astronomy of Utrecht University, where she is the head of the Human Perception Group. She was promoted to full professor in 2005. Her research is conducted at the Helmholtz Institute. Her research interests include haptic and visual perception. In 2003, she

Wouter M. Bergmann Tiest received the MSc degree in experimental physics in 1999 and the PhD degree from Utrecht University, The Netherlands. Until 2004, he was employed by The Netherlands Institute for Space Research. He is currently a postdoctoral researcher at the Department of Physics and Astronomy of Utrecht University, where he is with the Human Perception Group of the Helmholtz Institute. His research interests include haptic searching,



California, Berkeley, where he was engaged in research on psychophysical experiments and computational models investigating the integration of visual-haptic information under professor Martin Banks. At the end of 2001, he returned to the Max Planck Institute. He is currently the principal investigator of several international grants including the two European Projects, namely, ImmerSence and CyberWalk. He is the leader of the independent Research Group Multisensory Perception and Action at the Max Planck Institute for Biological Cybernetics in Tübingen, Germany. The group is interested in human multimodal perception and sensorimotor integration. Methodologically, the group mainly uses quantitative psychophysical and neuropsychological methods together with Virtual Reality techniques and Bayesian models of sensory perception. He is a member of the IEEE.

Marc O. Ernst received the PhD degree from the Max Planck Institute for Biological Cybernetics for investigations on the human visuomotor behavior. For this, he was awarded the Attempto Prize (2000) from the University of Tübingen and the Otto-Hahn-Medaille (2001) from the Max Planck Society. He studied physics in Heidelberg and Frankfurt/Main. Starting in 2000, he spent almost two years as a postdoctoral researcher at the University of

► **For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.**