https://doi.org/10.2352/issn.2169-2629.2021.29.66

Visual perception of surface properties through manipulation

James A. Ferwerda, Snehal A. Padhye; Center for Imaging Science, Rochester Institute of Technology, Rochester, NY/USA

Abstract

Vision is a component of a perceptual system whose function is to support purposeful behavior. In this project we studied the perceptual system that supports the visual perception of surface properties through manipulation. Observers were tasked with finding dents in simulated flat glossy surfaces. The surfaces were presented on a tangible display system implemented on an Apple iPad, that rendered the surfaces in real time and allowed observers to directly interact with them by tilting and rotating the device. On each trial we recorded the angular deviations indicated by the device's accelerometer and the images seen by the observer. The data reveal purposeful patterns of manipulation that serve the task by producing images that highlight the dent features. These investigations suggest the presence of an active visuo-motor perceptual system involved in the perception of surface properties, and provide a novel method for its study using tangible display systems

Introduction

In the laboratory, it's customary to study vision under highly controlled conditions, both to eliminate the influence of extraneous factors, and to allow visual functions (color, motion, shape, etc.) to be studied in isolation. Yet as Gibson [1] observed, under normal conditions, vision is a component of a complex sensori-motor perceptual system whose function is to provide information in support of purposeful behavior. For example, as shown in Figure 1, to perceive the shape of the guitar back, the luthier holds it up to the light, sights down the surface, and tilts it back and forth to reveal its shape. While much has been learned about vision from controlled laboratory studies, there is undoubtably still more that could be learned by studying vision under more natural conditions.

Over the past several years, Ferwerda et al. [2] have been developing tangible display technologies that leverage the form factors, graphics, and sensors in modern tablet devices to allow observers to view and manipulate virtual objects in ways similar to the ways they do with real ones. Figure 2 shows a sequence of images from a video of an observer interacting with a tiled surface rendered on an iPad-base tangible display. Note that as the observer tilts the iPad back and forth, the highlights and textures in the images of the tiled surface change as if the observer were holding the surface in their hands.

In this paper we first describe a set of tools we've developed that allow us to run Web-based psychophysical experiments on the role of manipulation in surface perception using tangible display systems, and then present the findings of studies that investigate how observers view and manipulate surfaces to understand their properties. The goal of this paper is to introduce a new paradigm for the study of surface perception that enables the quantitative investigation of visuo-motor behavior as a perceptual system.

Related work

While Gibson was the first to describe the complex, information-seeking, visuo-motor behavior perceivers engage in as a perceptual system, Yarbus's studies of task-dependent eye movements in the 1950's [3] are often cited to illustrate the properties of active vision. Using a primitive eye-tracker, Yarbus showed that observers' gaze patterns when viewing a painting of a domestic scene were dramatically affected by the high-level conceptual task (e.g. estimate the wealth, ages, positions, etc. of people in the scene) given by the experimenter. While more recent studies [4, 5] have failed to replicate Yarbus's impressive findings, there is widespread agreement that visuo-motor behavior is shaped by the observer's task and goals.

Figure 1. Perceiving surface properties through manipulation.

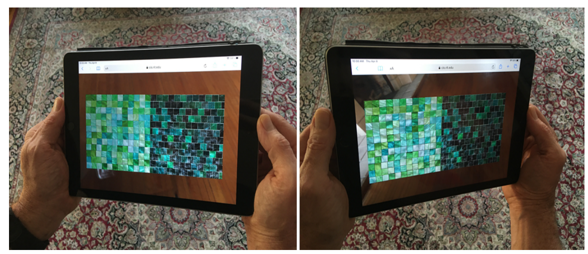

Figure 2. Interacting with a virtual surface using a tangible display system.

Pelz and colleagues have published a series of papers (mostly summarized in [6, 7]) that look at the coordination of eye, head, and hand movements in natural tasks such as hand washing and sandwich making, and they have discovered sophisticated interactions between visual, motor, and haptic systems including "look-ahead" and "just-in-time" information gathering that supports fluid and adaptive actions. Hayhoe [8] has recently written a review of the literature in this area.

While the focus of the previous research is on eye, head, and hand movements per se, with objects and tasks serving as foils to reveal patterns and interactions, Johansson et al. [9] brought the focus to objects, and looked at eye-hand coordination in object manipulation. They found that grasping, wielding, and pressing actions are supported by oculo-motor behaviors that suggest the presence of a perceptual system that is predicting the future locations of objects to accurately guide hand movements.

Beyond the question of how the visual, motor, and haptic systems interact to allow objects to be manipulated, is the question of what role manipulation plays in the perception of object and surface properties. While there is a vast literature on the visual perception of surface shape and a rapidly growing literature on the perception of surface materials (see [10, 11] for reviews), to our knowledge, little study has been done to understand how people manipulate objects and surfaces to perceive their properties. To this end we have developed the systems and methods described in the following sections.

Tools developed

To investigate the role that manipulation plays in the perception of surface properties we need tools that allow us to 1) create surfaces that vary systematically in their properties and 2) record how observers manipulate the surfaces when judging their properties. Tangible display systems support both of these needs. In previous publications [2,12], Ferwerda et al. have described the technical issues involved in implementing tangible displays in detail so we will just summarize their features here.

A tangible display system starts with a modern mobile tablet. These devices typically include a high-resolution color display, hardware accelerated 3D graphics, high-speed wireless networking, and a variety of sensors (cameras, gyroscopes, accelerometers). Custom software senses how the tablet is being manipulated and uses that information to drive a 3D rendering engine that takes a model of a surface (topography and reflectance properties) and a model of illumination (simple or environment mapped) that renders photorealistic images of the surface to the tablet's screen in real-time. The software is designed so that tilting and rotating the tablet causes the lighting on the rendered surface to change appropriately with respect to the lighting environment. Thus, the tangible display allows observers to interact with virtual surfaces as naturally as they would with real ones, while providing precise control of visual stimulus properties and the ability to record patterns of interaction. While the original tangible displays were implemented in C++ under iOS, advances in Web-based 3D graphics have enabled us to implement the current system in Javascript using the three.js library [13], which is supported by most modern browsers across a wide variety of platforms.

The basic capabilities of the tangible display support the first of our needs, since by changing the properties of the surface model we can systematically vary its appearance. To support the second need, we have linked the system to a server-based back-end we have designed that through standard network protocols running in the observer's browser, 1) initializes an experimental session, 2) provides model data to the tangible display code, 3) records sensor data from the observer's device detailing how the display is being manipulated, as well as the observer's responses to experimental tasks, and 4) closes the experimental session.

Preliminary studies

Since tangible display systems provide a new tool for investigating the relationships between manipulation and the perception of surface appearance, the studies presented here are preliminary - seeking to answer basic questions. We have focused on three:

- How do observers manipulate surfaces when judging their properties?
- What kinds of visual features do observers use to judge surface properties?
- What are the relationships between the visual features and the patterns of manipulation?

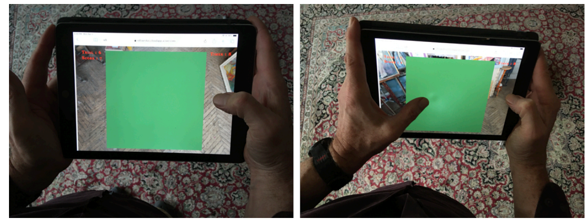

Figure 3. Observer's view of the dent-finding experiment on the tangible display.

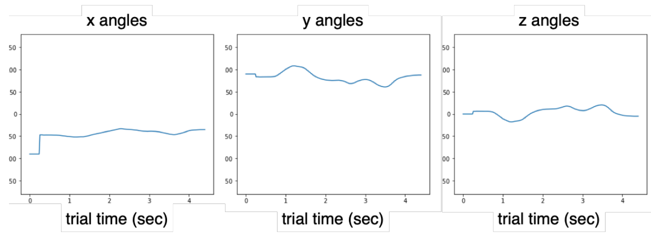

Figure 4. Accelerometer records showing angular deviations of the display/surface with respect to the observer's line-of-sight.

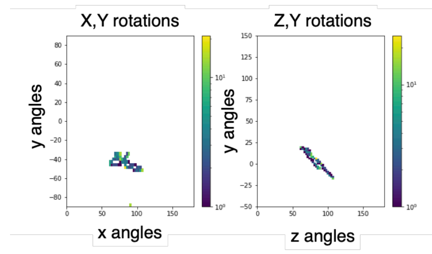

Figure 5. 2D plots showing the observer's patterns of manipulation. Positions indicate display/surface orientations, and colors indicate time spent at a given orientation.

To answer these questions, we designed an experiment where observers were tasked with finding dents in an otherwise flat surface. The observer's view in the experiment is illustrated in Figure 3.

The experiment consisted of a series of 20 trials. On each trial, the observer viewed a green, glossy, textured surface in natural lighting rendered on the tangible display. The surface was flat except for a randomly positioned gaussian dent. Over the course of the experiment the depth of the dent was controlled by a staircase procedure to keep detectability near threshold. The dents were thumb-sized and were deep enough to be visible, but were shallow enough that they could not be seen by diffuse shading alone. Therefore, observers were forced to manipulate the display/surface to find the dent, at which point they were instructed to indicate its position by touching it on the screen.

To encourage the observers to be both fast and accurate in their searches, we gamified the experiment. A 20 second countdown timer started at the beginning of each trial and stopped when the observer touched the screen. If they correctly indicated the location of the dent, the value of the countdown timer was added to their point total. If they were incorrect it was subtracted.

Over the course of the trials, we recorded haptic data related to the observer's manipulations of the display/surface, and visual data related to the rendered images produced by these manipulations.
Haptic data included the output of the tangible display's accelerometer (Figure 4), from which we derived measures of the angular deviations of the surface with respect to the observer's line of sight (Figure 5). Visual data included luminance images of the

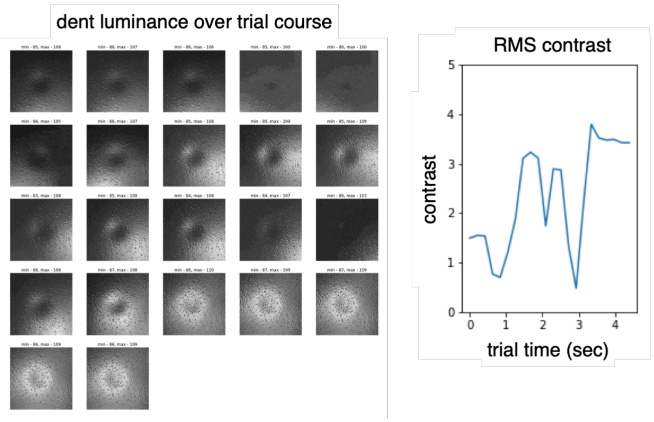

Figure 6. Luminance images of a surface dent over the course of a trial as the display/surface is being manipulated and RMS contrasts of the images.

dent over the course of a trial (Figure 6), for which we calculated measures of RMS image contrast.

In total, 20 observers participated in the studies (ages 16-65, normal or corrected-to-normal vision). All showed similar patterns of behavior/response/performance. Since at this point in our research, our goal is to broadly understand how observers manipulate surfaces to understand their properties, representative results from several observers will be presented in the following section.

Results

Haptic data: Figure 7 shows the accelerometer data from trials 13-16 of a typical run of the experiment. Each graph indicates the angular deviation of the display/surface with respect to the observer's modeled line-of-sight. From trial-to-trial, the staircase procedure is decreasing the depth of the dent, making the search/detection task harder and harder.

There are several patterns that can be observed in the haptic data from this trial series. First, the trial durations increase (2.5s, 7.0s, 7.5s, and 25s respectively) suggesting the task is becoming more difficult. Second, the amplitudes of the angular deviations increase, suggesting that the observers are making more extensive manipulations and searching a greater range of viewing angles. Finally, it appears that the frequencies at which the manipulations occur also increase, suggesting more vigorous manipulations, however it's possible that this is a confound produced by the different durations of the trial records. Further analysis would be required to answer this question.

Plotting the angular deviations in pairs reveals more information about the observer's patterns of manipulation. Figure 8 shows 2D plots of the X,Y rotation data, where position indicates angular rotation about the horizontal and vertical axes of the display/surface, and color represents the amount of time the device was at that orientation. There are several insights that can be taken from these plots. First, when the task is relatively easy (trials 13,14), the manipulation paths are simple lines or curves. A path might be retraced (trial 14), but detection is relatively quick, and time spent at any particular orientation is low. On the other hand, when the task is more difficult (trial 15, and especially 16), the manipulation path becomes more two dimensional (trial 15) and is ultimately exhaustive (trial 16), searching all locations in angle-space. Second, as the task gets harder, more time is spent at some orientations than others, presumably because these orientations produce images that either reveal the dents, or suggest dent-like features.

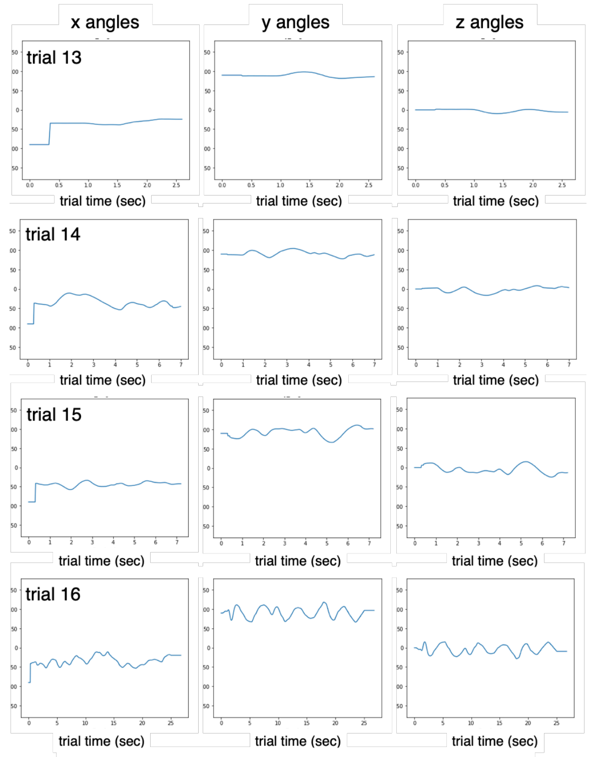

Figure 7. Representative accelerometer data from trials 13-16 of the experiment. Note how the trial durations and angular deviations of the manipulations increase as the search/detection task becomes harder.

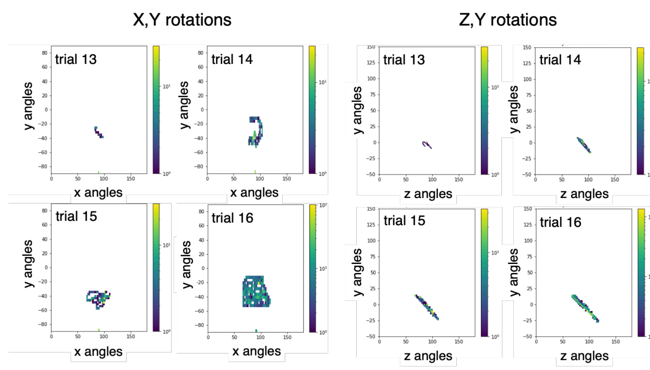

Figure 8. 2D plots of angular deviations with respect to the X,Y and Z,Y axes. Note that as the task gets harder, the manipulations become more extensive and exhaustive in X,Y but remain highly correlated in Z,Y.

However, in contrast to the changes found in the X,Y rotation plots with increasing task difficulty, the Z,Y axis plots all show similar patterns of manipulation. The amplitudes of the paths increase, but in all there are strong correlations between Y axis and Z axis rotations (hence the linear paths). Since the Z data represents rotations around the line-of-sight, and this action produces little change in the displayed images, we believe these paths represent unintentional "steering wheel" rotations that occur while the observer is attempting to "slant" and "tilt" the display/surface with respect to their line-of-sight.

Image data: The dynamic relationships between the haptic manipulations described above and the image features produced by the manipulations are illustrated in Figure 9. Each frame in each sequence shows the luminance image in the vicinity of the dent taken every 200ms as the observer manipulates the display/surface.

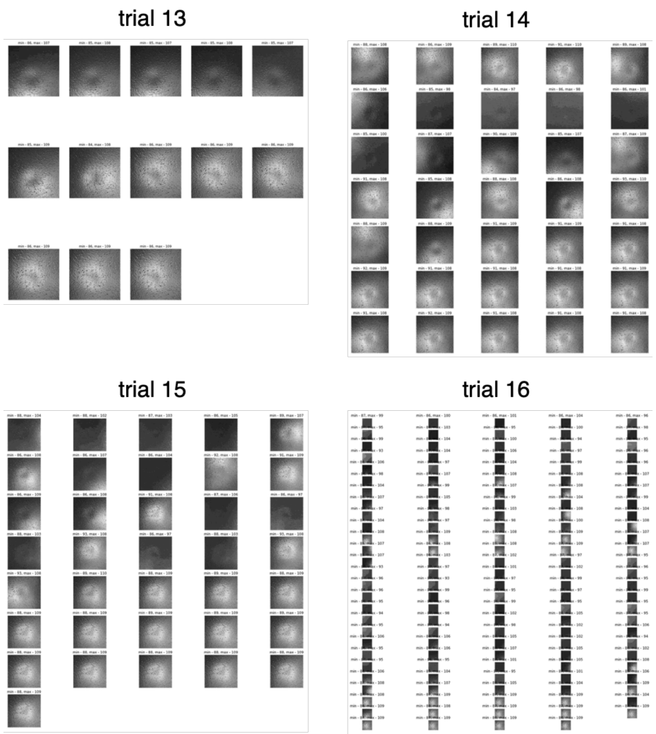

Figure 9. Luminance images of the dents over the course of a trial. Images were capture every 200ms. Note the repeated high- lighting, un-lighting cycles prior to detection/response.

Since the dent in trial 13 is relatively deep, the detection task is relatively easy, and the image sequence shows that 1) the dent is always visible, and 2) the observer's manipulations serve to illuminate the dent uniformly, producing a clear light/dark ring pattern. The corresponding trial 13 graph in Figure 10 shows that these manipulations also cause the RMS contrast of the images to first decrease and then increase to a high-level prior to detection/response.

As the dent depth decreases and the search/detection task gets harder (trials 14, 15, 16), longer yet similar image patterns are seen. For example, the luminance image record for trial 15 (Figure 9) shows a number of cycles where the observer is manipulating the display surface to alternately high-light and un-light the dent area. It's as if the observer is using light sources in the scene as flashlights, and is orienting and re-orienting the surface to create a dynamically changing image stimulus. This behavior is also reflected in the corresponding RMS contrast graph for trial 15 (Figure 10) where the over the course of the trial, the contrast of the dent image increases and decreases several times.

Although the long duration of trial 16 and the large number of luminance images (Figure 9) make it difficult to see, this dynamic high-lighting behavior was especially prominent when the shallow geometry of the dent made it hard to detect. The back-and-forth manipulations and correlated image contrast changes (Figure 10) provide a visual stimulus that appears to facilitate detection.

Discussion

At the outset of our studies, we posed three related questions: How do observers manipulate surfaces when judging their properties? What kinds of visual features do observers use to judge surface properties? What are the relationships between the visual

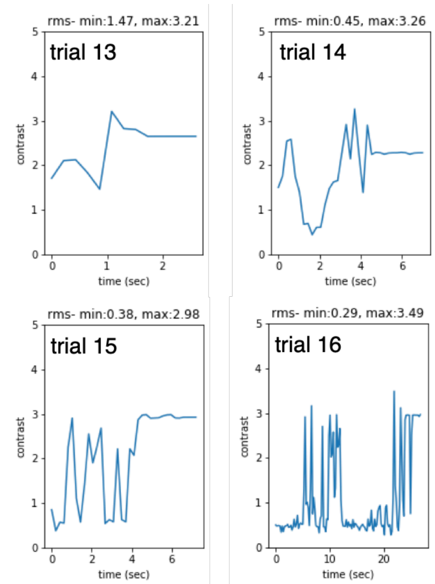

Figure 10. RMS contrasts of the dent images over the course of a trial. Note how the observer's manipulations cause the image contrasts to increase and decrease several time prior to detection/response.

features and the patterns of manipulation? While the results of our studies are preliminary, we can begin to address these questions.

For the dent search/detection task we created, at first observers manipulate the surface causing light from scene illuminants to play across the surface, with the goal of revealing surface features through image contrast patterns. Once a potential feature has been localized, the observer then manipulates the surface causing the contrast pattern of the feature to increase and decrease. At this point in our studies, we cannot say definitively why the observer does this, but one thought is that it might provide multiple opportunities to detect the dent feature or confirm its absence.

While these findings are exploratory, they offer novel insights into visual behavior and provide evidence for a sophisticated visuo-motor perceptual system that supports the active perception of surface properties.

Conclusion

In this paper we introduced a tool, tangible display systems, that allow us to run Web-based psychophysical experiments on the role of manipulation in surface perception. We then described the design and findings of some preliminary studies that investigate how observers view and manipulate surfaces to understand their properties. The findings are still exploratory, but suggest a complex interplay between the haptic manipulation of an observed surface and the visual information generated that supports perception. Our intent in this work is to introduce a new paradigm for the study of surface perception that enables the quantitative investigation of visuo-motor behavior as a perceptual system.

Since this is the first work of its kind, there are wide opportunities for future work, including studies of surface property detection, discrimination, and scaling, with respect to both geometric and material properties. We plan to investigate these topics in subsequent studies.

References

[1] Gibson, James Jerome, and Leonard Carmichael. *The senses considered as perceptual systems*. Vol. 2. No. 1. Boston: Houghton Mifflin, 1966.

[2] Ferwerda, James A. "Tangible imaging systems." *Imaging and Printing in a Web 2.0 World IV*. Vol. 8664. International Society for Optics and Photonics, 2013.

[3] Yarbus, Alfred L. "Eye movements during perception of complex objects." *Eye movements and vision*. Springer, Boston, MA, 1967. 171-211.

[4] Lipps, Marianne, and Jeff B. Pelz. "Yarbus revisited: task-dependent oculomotor behavior." *Journal of Vision* 4.8 (2004): 115-115.

[5] Greene, Michelle R., Tommy Liu, and Jeremy M. Wolfe. "Reconsidering Yarbus: A failure to predict observers' task from eye movement patterns." *Vision research* 62 (2012): 1-8.

[6] Pelz, Jeff B., and Roxanne Canosa. "Oculomotor behavior and perceptual strategies in complex tasks." *Vision research* 41.25-26 (2001): 3587-3596.

[7] Hayhoe, Mary M., et al. "Visual memory and motor planning in a natural task." *Journal of vision* 3.1 (2003): 6-6.

[8] Hayhoe, Mary M. "Vision and action." *Annual review of vision science* 3 (2017): 389-413.

[9] Johansson, Roland S., et al. "Eye–hand coordination in object manipulation." *Journal of Neuroscience* 21.17 (2001): 6917-6932.

[10] Todd, James T. "The visual perception of 3D shape." *Trends in cognitive sciences* 8.3 (2004): 115-121.

[11] Fleming, Roland W. "Material perception." *Annual review of vision science* 3 (2017): 365-388.

[12] Ferwerda, James A. "ImpastoR: A realistic surface display system." *Vision Research* 109 (2015): 166-177.

[13] https://threejs.org/

Author Biographies

James A. Ferwerda is an Associate Professor in the Chester F. Carlson Center for Imaging Science at RIT. He received a B.A. in Psychology, M.S. in Computer Graphics, and a Ph.D. in Experimental Psychology, all from Cornell University. The focus of his research is on building computational models of human vision from psychophysical experiments and developing advanced imaging systems based on these models.

Snehal A. Padhye is a fourth year PhD. student in the Chester F. Carlson Center for Imaging Science at RIT. She received a BS in Electronics and MS in Signal Processing from India. Her dissertation research work focuses on designing hardware and software systems for capturing and visualizing realistic digital models of cultural heritage objects.