A Psychophysical Study of Surface Reconstructions Based on Binocular Disparity and Motion Parallax

Aya Shiraiwa, Takefumi Hayashi
Kansai University, Osaka, Japan

A visible-surface can be reconstructed either from binocular disparity or motion parallax, and several previous studies have revealed the similarity of the surface properties from either type of reconstruction. In the present research, we have studied the ability for interpolation of reconstructed surfaces, that is, a gap-spanning task was conducted using both static and dynamic random dot patterns which give the same surface perception viewed binocularly and monocularly. We found that: (1) The ability for surface interpolation based on binocular disparity is more effective than that based on motion parallax; and (2) The experimental data correspond to the simulation result of a computational model is based on standard regularization theory. Based on these results, the brain mechanisms of surface reconstruction were discussed.

Keywords: surface reconstruction, binocular disparity, motion parallax, computational model, standard regularization theory

Introduction

A surface can be perceived from a sparse RDS (random dot stereogram) viewed binocularly by means of surface reconstruction (Julesz, 1971; White, 1962). A surface can also be perceived from a dynamic random dot pattern, in which the velocity gradient of the dots is defined and viewed monocularly (Rogers & Graham, 1979). Unlike the differences obtained with different input modalities (binocular disparity and motion parallax), the characteristics of surface reconstruction in previous research have found to be similar. Modulation transfer functions show the band-pass properties, and a maximum sensitivity was found at a spatial frequency of 0.2-0.5 cycle/deg in both cases (Rogers & Graham, 1982). When a blank region without dots is created in random dot patterns, the surface is interpolated there both in the case of a binocularly disparity-defined pattern (Grimson, 1981; Wilcox & Duke, 2005) and in the case of a monocularly parallax-defined pattern (Saidpour, Braunstein, & Hoffman, 1992; 1994).

When we consider the mechanism of the surface reconstruction process from the 2-D retinal image to the 3-D object surface, a priori knowledge of the 3-D world is indispensable. Poggio, Torre, and Koch (1985) showed that the reconstruction (from 2-D to 3-D) problems in early vision are inherently ill-posed, and proposed a computational model based on the framework of standard regularization theory (Tikhonov & Arsenin, 1977). In their model, a priori knowledge of the 3-D world is defined as a constraint to solve the
equations to recover the 3-D information. In the case of surface reconstruction, a general differential operator of the smoothness constraint was suggested (Poggio et al., 1985). This differential operator consists of the first and second derivatives of the disparity, corresponding respectively to the “membrane” potential energy (Terzopoulos, 1986) and the “thin plate” potential energy (Grimson, 1981). Norman, Lappin, and Zucker (1991) reported that the human visual system is sensitive to the second derivatives of a continuous surface, whereas Bülthoff (1991) and Yuille, Geiger, and Bülthoff (1991) reported that the membrane potential energy accurately fits data from psychophysical experiments.

In the present study, we examined the difference between surface reconstruction for static and dynamic random dot patterns viewed binocularly and monocularly, respectively. In the following psychophysical experiments on surface reconstruction, we evaluated the results quantitatively by comparing them with the results of numerical calculations for a membrane potential energy and a thin plate potential energy. Furthermore, on the basis of a report that the perception of 3-D surface orientation is processed in the parietal lobe (Tsutsui, Taira, & Sakata, 2005), we discussed the brain localization and the pathways over which surface reconstruction takes place. Note that the use of the term “motion parallax” in this article actually refers to a kinetic depth effect, but we use these terms interchangeably.

**Experiment 1**

In Experiment 1, we examined the characteristics of surface interpolation of disparity-defined patterns using Julesz’s (1960) RDS, and considered the applicability of two types of potential energy for standard regularization theory. In the RDSs used as stimuli, we formed a region without dots (a gap) around the peak of the sinusoidal wave, and examined the characteristics of the surface reconstruction in that region by depth matching (that is, a gap-spanning task).

**Methods**

**Subjects.** Four university undergraduates (two males), one graduate student (female) and one of the authors (AS) participated in each of the experiments reported here. The subjects were in their 20s, with normal or corrected-to-normal visual acuity, and normal stereopsis. Except for AS, the subjects received payment for their participation and were unaware of the purpose of the study.

**Apparatus.** A personal computer (Dell Precision PWS470) was used to generate the stimuli, which were displayed on a 20-inch CRT monitor (Iiyama HM204DA; 1600 × 1200 pixels) with a refresh rate of 75 Hz. The subjects viewed the display in a darkened room from a distance of 100 cm through a haploscope with both eyes. The viewing area of the RDS used for each stimulus was the 12.5 × 12.5 cm, equivalent to a visual angle of 7.13 × 7.13 deg. The mean luminance of the display was 2.93 cd/m² when viewed through the haploscope.

**Stimulus.** Figure 1 shows examples of RDSs used as stimuli in Experiment 1. In the right eye image of the RDS, the horizontal positions of the dots $x$ were modulated sinusoidally by their vertical positions $y$ ($x = \sin y$). A 3-D convex surface of a half-period sinusoid is perceived when the images from the both eyes are fused (see Figure 1 (a)). The distribution of disparity was constant along the horizontal axis, and the peak of the sinusoidal wave was assigned to the center of the screen. A gap (a region without dots) was formed near the center of the peak of the sinusoidal wave (see Figure 1(b)). Figure 2(a) and Figure 2(b) show the images of the perceived surface by subjects corresponding to the stimuli shown in Figure 1(a) and Figure 1(b), respectively.
A PSYCHOPHYSICAL STUDY OF SURFACE RECONSTRUCTIONS

Figure 1. Examples of random dot patterns used in the experiments. It can be seen that, for ease of understanding, the colours of the background and the dot patterns have been reversed. In the actual experiment, white dots on black background stimuli were used. The density of dots and the disparity amplitude of these stimuli are 2.46 dots/deg$^2$ and 0.40 deg, respectively: (a) shows the pattern with dots on the peak of the sinusoidal wave; and (b) shows the pattern with a region at the centre of the area without dots. The width of this region is 1.72 deg.

Figure 2. Images of the perceived surface. Images: (a) with dots on the peak of the sinusoidal wave; and (b) without dots at the centre region of the pattern (with gap).

The dot density of the RDSs was 2.46 dots/deg$^2$ (the number of dots was 125), and the spatial frequency was 0.07 cycle/deg. In preliminary experiments, we examined the cases where dot density is high. When the dot density increased beyond 4.92 dots/deg$^2$ (the number of dots was 250), subjects detected an edge at the border of the gap and were not able to perceive any surface at all in the gap. Thus, we lowered the dot density of the RDSs, so that every subject could perceive a surface in the gap region.

In the present experiments, there were three kinds of stimulus pattern that changed the amplitude of the sinusoidal wave (maximum disparity at the peak of the sinusoidal wave at 0.20, 0.40 and 0.60 deg), and five kinds that changed the width of the gap (0.00, 0.86, 1.72, 2.58 and 3.43 deg). The total number of stimuli was 15.

Task. The stimulus pattern of the RDS and the marker used to assess the depth of the perceived surface were presented at the same time. The vertical position of the marker was at the center (in the gap), and its horizontal position was 0.40 deg from the right-side end of the stimulus. Its size was 5.16 × 5.16 min. The subjects altered the binocular disparity of the marker from the keyboard to get the same magnitude of depth as the amplitude of the interpolated surface, and the value of the binocular disparity was recorded. The marker was put on the peak of interpolated surface. Ten trials were repeated for the pattern of each amplitude. The subjects proceeded to the next stimulus once they were satisfied with their depth judgment for the interpolated
surface. We presented five kinds of stimulus pattern in each trial at random by changing the width of the gap. A total of \(10 \times 3 \times 5 = 150\) trials was given.

### Results and Discussion

The results of Experiment 1 are shown in Table 1 and Figure 3. In Table 1, the estimated amplitude of the perceived surface for each subject is presented (matched disparity of the marker, in the unit of degrees). As the width of the gap increased, the estimated amplitude decreased for each subject. According to the introspective report of subjects, we confirmed that the gap was perceived as a continuous surface which is extrapolated from the region with dots. However, there were large individual differences concerning the width of the gap that was perceived as a surface. In particular, three of the six subjects (AS, NF, MM) could perceive the interpolated surface in the gap in the case of 2.58 and 3.43 deg; it was difficult for two other subjects (TK, CN), and it was completely impossible for one subject (CF).

Table 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>0.00 deg</th>
<th>0.86 deg</th>
<th>1.72 deg</th>
<th>2.58 deg</th>
<th>3.43 deg</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>0.220</td>
<td>0.217</td>
<td>0.215</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>TK</td>
<td>0.225</td>
<td>0.227</td>
<td>0.231</td>
<td>0.230</td>
<td>0.230</td>
</tr>
<tr>
<td>CN</td>
<td>0.235</td>
<td>0.226</td>
<td>0.207</td>
<td>0.207</td>
<td>0.180</td>
</tr>
<tr>
<td>AS</td>
<td>0.214</td>
<td>0.206</td>
<td>0.200</td>
<td>0.193</td>
<td>0.169</td>
</tr>
<tr>
<td>NF</td>
<td>0.205</td>
<td>0.209</td>
<td>0.203</td>
<td>0.215</td>
<td>0.203</td>
</tr>
<tr>
<td>MM</td>
<td>0.252</td>
<td>0.242</td>
<td>0.211</td>
<td>0.196</td>
<td>0.169</td>
</tr>
</tbody>
</table>

In addition, a two-way repeated measures ANOVA (analysis of variance) was performed with: (1) disparity (0.2, 0.4, and 0.6 deg); and (2) width of the gap (0.00, 0.86, 1.72, 2.58, and 3.43 deg) as factors. The interaction was statistically significant \(F(8, 40) = 5.41, p < 0.01\). The main effects of disparity and width of the gap were significant (disparity: \(F(2, 10) = 272.06, p < 0.01\); gap width: \(F(4, 20) = 3.68, p < 0.05\)).

To compare the experimental results with the output of the computational model, experimental data for each subject are normalized by the estimated value of “the gap width = 0”, and averaged for all subjects. The
results are shown in Figure 3 with the numerical simulation results based on the standard regularization theory (Poggio et al., 1985). The horizontal axis shows the width of the gap, and the vertical axis shows the amplitude of the surface that subjects estimated. Figures 3 (a), (b) and (c) corresponds to stimulus disparities of 0.20, 0.40 and 0.60 deg, respectively. The values of “membrane” and “thin plate” in the figures are the numerical simulation results based on standard regularization theory using the membrane potential energy function (Terzopoulos, 1986) and the thin plate potential energy function (Grimson, 1981), respectively.

The experimental results showed a tendency similar to the characteristics of the surface reconstruction using the thin plate potential energy function (which uses the information of the curvature of the surface). This result supports the earlier experimental studies of Norman et al. (1991) and Saidopour et al. (1994), which reported that the human visual system is sensitive to the second derivatives of a continuous surface.

**Experiment 2**

In Experiment 2, we examined the characteristics of surface interpolation by parallax-defined patterns with
sinusoidal velocity gradients. Here, we formed a gap, similar to Experiment 1, and we asked subjects to match the depth of the interpolated surface to that of a marker whose perceived depth increased or decreased by disparity (a gap-spanning task). In Experiment 2, we converted the velocity gradient of random dot patterns into equivalent disparity (Rogers & Graham, 1979, 1982; M. Ono, Rivest, & H. Ono, 1986).

Methods

Subjects. As subjects, one male and five female university undergraduates participated in each of the experiments reported here. The subjects were in their 20s, with corrected-to-normal vision and good stereopsis. The subjects received payment for their participation and were unaware of the purpose of the study.

Apparatus. The apparatus was the same as in Experiment 1. In addition, although the refresh rate of the CRT monitor was again 75 Hz, the frame rate during stimulus presentation in Experiment 2 (the moving random dot pattern) was 60 Hz due to software effects. Because the refresh rate was higher than the frame rate, however, the desynchronization did not in fact prevent the smooth motion of the moving random dot pattern.

Stimulus. We used moving random dot patterns with sinusoidal velocity gradients of the half period as stimuli. The distribution of dot velocity \( v_x \) (along the horizontal axis) was constant, but was modulated sinusoidally by the vertical position of the dot \( y \) \((v_x = \sin y)\). The dot density, the spatial frequency and the nature of the gap (region without dots) were the same as those of the patterns used in Experiment 1. The moving random dots were displayed for 2.5 seconds, accompanied with acceleration and deceleration. The direction of movement was horizontal (from the subject’s left to right). When the patterns were observed monocularly, surfaces, such as those in Experiment 1 (as shown in Figure 2(a) and Figure 2(b)) were perceived.

In Experiment 2, there were three kinds of stimulus patterns that changed the amplitude of the sinusoidal wave (equivalent disparity; 0.20, 0.40 and 0.60 deg), and three kinds that changed the width of the gap (0.00, 0.86 and 1.72 deg). The total number of stimuli was 9. Note that, in the results of a preliminary experiment, it is difficult for most subjects to interpolate the surface to the gap when the width of the gap exceeded 2.58 deg, therefore, gap widths of 2.58 and 3.43 deg were not used in Experiment 2.

Task. Because there was no velocity gradient in the initial stimulus pattern of random dots, subjects could not perceive depth. Following a key press, the dots moved for 2.5 seconds and the stimulus surface became visible. The subjects then altered the binocular disparity of the marker from the keyboard to get the same magnitude of depth as the amplitude of the interpolated surface. During the experiment, the subjects observed the random dot pattern monocularly (with the dominant eye) and the marker binocularly through the haploscope. The subjects could observe the movement of the random dot pattern repeatedly, until they were satisfied with their depth judgment for the interpolated surface. We presented three kinds of stimulus pattern in each trial at random by changing the width of the gap. A total of 10 (repeat) \( \times \) 3 (disparity) \( \times \) 3 (width) = 90 trials was given.

Results and Discussion

The results of Experiment 2 are shown in Table 2 and Figure 4. In Table 2, the estimated amplitude of the perceived surface for each subject is presented. There were large individual differences in the results, but all subjects underestimated the amplitude of the surface. Despite this, each subject reported that a 3-D surface was perceived when the width of the gap was 0 deg. When the width of the gap was 0.86 deg, four of six subjects (CU, NS, KS, KN) perceived a continuous smooth surface, and the others perceived a flat surface. Moreover, all subjects perceived the gap region as flat when the width of the gap was 1.72 deg.
Table 2

Results of Experiment 2: Estimated Amplitude of Perceived Surfaces for Each Subject

<table>
<thead>
<tr>
<th>Subject</th>
<th>0.00 deg</th>
<th>0.86 deg</th>
<th>1.72 deg</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20 deg amplitude subject:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CU</td>
<td>0.068</td>
<td>0.074</td>
<td>0.080</td>
</tr>
<tr>
<td>NS</td>
<td>0.104</td>
<td>0.143</td>
<td>0.107</td>
</tr>
<tr>
<td>KS</td>
<td>0.028</td>
<td>0.030</td>
<td>0.016</td>
</tr>
<tr>
<td>YS</td>
<td>0.057</td>
<td>0.037</td>
<td>0.036</td>
</tr>
<tr>
<td>KN</td>
<td>0.074</td>
<td>0.076</td>
<td>0.075</td>
</tr>
<tr>
<td>HM</td>
<td>0.061</td>
<td>0.050</td>
<td>0.033</td>
</tr>
<tr>
<td>0.40 deg amplitude subject:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CU</td>
<td>0.112</td>
<td>0.108</td>
<td>0.120</td>
</tr>
<tr>
<td>NS</td>
<td>0.131</td>
<td>0.127</td>
<td>0.124</td>
</tr>
<tr>
<td>KS</td>
<td>0.054</td>
<td>0.048</td>
<td>0.028</td>
</tr>
<tr>
<td>YS</td>
<td>0.101</td>
<td>0.071</td>
<td>0.071</td>
</tr>
<tr>
<td>KN</td>
<td>0.138</td>
<td>0.130</td>
<td>0.117</td>
</tr>
<tr>
<td>HM</td>
<td>0.145</td>
<td>0.108</td>
<td>0.088</td>
</tr>
<tr>
<td>0.60 deg amplitude subject:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CU</td>
<td>0.173</td>
<td>0.199</td>
<td>0.165</td>
</tr>
<tr>
<td>NS</td>
<td>0.139</td>
<td>0.163</td>
<td>0.145</td>
</tr>
<tr>
<td>KS</td>
<td>0.038</td>
<td>0.037</td>
<td>0.033</td>
</tr>
<tr>
<td>YS</td>
<td>0.114</td>
<td>0.100</td>
<td>0.070</td>
</tr>
<tr>
<td>KN</td>
<td>0.161</td>
<td>0.169</td>
<td>0.147</td>
</tr>
<tr>
<td>HM</td>
<td>0.142</td>
<td>0.158</td>
<td>0.123</td>
</tr>
</tbody>
</table>

In addition, a two-way repeated measures ANOVA was performed with: (1) equivalent disparity (0.20, 0.40 and 0.60 deg); and (2) width of the gap (0.00, 0.86 and 1.72 deg) as factors. The interaction was statistically significant ($F_{(4, 20)} = 3.20, p < 0.05$) and the main effects of disparity and width of the gap were also significant (disparity: $F_{(2, 10)} = 7.72, p < 0.01$; gap width: $F_{(2, 10)} = 4.43, p < 0.05$).

Figure 4 shows a comparison of the experimental results to those of the output of the computational model of surface reconstruction. The experimental data were normalized and averaged for all subjects in the same manner as Experiment 1 (as shown in Figure 3). Using the mean values of all subjects, we could not conclude whether the experimental results indicate a tendency towards the characteristics of either simulated surface reconstruction. In fact, the data were similar to the numerical results of the thin plate potential energy for three subjects (CU, NS, KN) and similar to that of the membrane potential energy for three subjects (KS, YS, HM).

Despite the fact that subjects could reliably perceive surfaces in Experiment 1, they perceived a flat surface when the width of the gap was 1.72 deg in Experiment 2. In Experiment 1, if the width of the gap was expanded so that subjects could not perceive an interpolated surface, they did not perceive a flat surface. Therefore, it is suggested that the surface reconstruction was not achieved for two subjects (YS, HM) when the width of gap exceeded 0.86 deg, and for all subjects when it exceeded 1.72 deg.

**General Discussion**

We examined the properties of surface reconstruction between binocular disparity and motion parallax, and found a difference with regard to the interpolation of the gap. Here, we consider this difference from the viewpoint of the processing pathways and receptive fields in the visual cortex of the brain.
A PSYCHOPHYSICAL STUDY OF SURFACE RECONSTRUCTIONS

The diameter of the receptive field with central vision is known to be three deg in V4 and six deg in TEO (Desimone & Duncan, 1995). The IT (inferior temporal) is divided into substructures TEO (posterior) and TE (anterior). Moreover, it is one to three deg in MT (middle temporal), and 40 deg in MST (medial superior temporal) (Zeki, 1974). Mishkin, Ungerleider, and Macko (1983) argued that the ventral pathway and the dorsal pathway are parallel systems. The ventral pathway processes “what” information, such as shape and color, and the dorsal pathway processes “where” information, such as depth and motion. Most cortico-cortical connections are bidirectional with interactions between feedforward and feedback inputs (Mishkin et al., 1983).

V4 is normally included in the ventral pathway, and it is known that there is a small number of neurons responding selectively to the direction of movement (Zeki, 1977), but there are many more neurons responding selectively to form and spatial frequency (Van Essen & Gallant, 1994). It has also been reported that there are many disparity selective neurons in IT, which takes projections from V4 (Uka, Tanaka, Yoshiyama, Kato, & Fujita, 2000), and they indicate form selectivity in the case of RDS (Tanaka, Uka, Yoshiyama, Kato, & Fujita, 2001). Moreover, MT is included in the dorsal pathway, and it is known that there are many neurons with moving direction selectivity (Zeki, 1974).

In the present experiment, the width of the region at which all subjects could perceive a continuous surface was 1.72 deg in Experiment 1 and 0.86 deg in Experiment 2. However, it became difficult to perceive...
the surface when the width of the region exceeded 2.58 deg in Experiment 1. In addition, subjects perceived a flat surface in the gap when the width was 1.72 deg in Experiment 2. When the subjects perceived a flat surface, they detected edges of the borders between the region with dots and the gap were perceived as a discontinuity. Because detection of a discontinuity was different from the process of interpolation of a smooth surface, we surmised that subjects could not interpolate the surface. The crucial widths in Experiments 1 and 2 were smaller than the receptive field of V4 in the ventral pathway, and were approximately equal to that of MT in the dorsal pathway.

Tsutsui, Taira, and Sakata (2005) reported that the perception of 3-D surfaces was processed in the parietal lobe. Therefore, the final stages of the process of surface reconstruction might similar in both binocular disparity and motion parallax. In addition, it is thought that the width of the gap is affected by the size of the receptive fields in the process of surface interpolation. Therefore, our results suggest that surface reconstruction with binocular disparity passes through the ventral pathway, whereas that with motion parallax passes through the dorsal pathway. Janssen, Vogels, and Orban (2000a; 2000b) reported that surface interpolation might be accomplished at a subsequent stage of extraction of disparity gradients, such as TE. In the ventral pathway, TE is the last stage, but Felleman and Van Essen (1991) reported that the signal is output to 7a of the dorsal pathway. This finding supports the effects of receptive fields and pathways deduced from our results and those reported by Tsutsui et al. (2005).

We can conclude that surface reconstructions by binocular disparity and motion parallax are different subsequent to processing at the lateral geniculate nucleus, but all such processes ultimately converge in visual cortex and rely on the same mechanism regardless of the monocular or binocular origin of the disparity. Previous studies indicated that stereo and motion information are not independently processed (Domini, Caudek, & Tassinari, 2006), but our results suggest that, with regard to surface reconstruction, two different brain pathways are employed.

Conclusions

In the present study, we examined differences in surface reconstruction between binocular disparity and motion parallax using a gap-spanning task. The results were compared with the output of a computational model based on standard regularization theory. It was shown that the maximum gap width at which the reconstructed surface can be interpolated was notably larger in the case of the binocular disparity compared to that of motion parallax. Assuming that the 3-D surface perception is processed in the parietal lobe, this gap width difference can be explained by considering the magnitude of the receptive fields in the pathways of information in which binocular disparity and motion parallax are processed.

For both experiments, the results tended to favor the computational results based on standard regularization theory. Especially for binocular disparity, the thin plate potential energy function (which uses the second derivatives of the surface) is a good fit for the data. This result supports earlier experimental studies indicating that human visual system is sensitive to the second derivatives of a continuous surface. However, for motion parallax, because of individual differences between subjects, we can not conclude which energy function is a better fit to the experimental data. Further research is required.

We can conclude that standard regularization theory can handle various processes of early vision in a unified manner, and supports the neurophysiological validity of such models. Henceforth, we will compare psychophysical data with numerical simulations that apply standard regularization theory in surface reconstruction by motion parallax.
References


