Recognizing faces defined by texture gradients

CHANG HONG LIU
University of Hull, Hull, England

CHARLES A. COLLIN
University of Ottawa, Ottawa, Ontario, Canada

and

REZA FARIVAR and AVI CHAUDHURI
McGill University, Montréal, Québec, Canada

Texture gradients can reveal surface orientation in a manner similar to shape from shading, and therefore provide an important cue for object recognition. In this study, we tested whether a complex 3-D object, such as a face, can be identified from texture gradients alone. The stimuli were laser-scanned faces for which the texture element was a fractal-noise pattern mapped onto the 3-D surface. An eight-alternative forced choice task was used in which participants matched a face defined by texture gradients to one of eight faces defined by shape from shading (Experiment 1) or by texture gradients (Experiment 2). On average, participants scored 24% and 18%, respectively, above chance in these experiments. Although this performance was much poorer than the performance based entirely on shape-from-shading stimuli (Experiment 3), the results suggest that texture gradient information may be used to recover surface geometry of complex objects.

Artists have long recognized that the orientation of a surface relative to a vantage point can be described by a projection of the surface's texture onto a picture plane. For centuries, artists have applied this technique to create illusions of a 3-D world on a 2-D plane. It was not until recently, though, that this observation was introduced into the science of perception. Using photographs of natural scenes and drawings of patterned planar surfaces, Gibson (1950) demonstrated that an impression of depth can be created from a gradation of texture density when the uniformly textured surface is deviated from the perpendicular plane of a vantage point. He coined the term "texture gradient" to refer to the gradation of projected texture density. Gibson noted that the rate of change in the projected texture density provides useful information about surface orientation. He also showed that texture gradients can provide information about angular changes and surface discontinuities. Based on this analysis, Gibson argued that texture gradients contain important information for object recognition.

It is clear from Gibson's (1950) formulation that the role of texture gradients is very similar to that of shape

from shading. As with shape from shading, texture gradients can also reveal surface geometry of 3-D scenes. Nonetheless, the role of texture gradients in object recognition is relatively unknown compared with that of shape from shading. Studies on texture perception have largely been concerned with low-level processing, such as texture segregation or orientation of geometrically regular and irregular surfaces (e.g., Braunstein & Payne, 1969; Cavanagh, 1987; Christou, Koenderink, & van Doorn, 1996; Knill, 1998; Li & Zaidi, 2003; Rosenholtz & Malik, 1997; Todd & Oomes, 2002; Zaidi & Li, 2002). The contribution of texture processing to higher level object recognition/identification, however, is rarely studied. This lack of research may be due to the belief that object recognition is by and large edge-based, so that surface properties of objects do not play an important role (see a review by Bruce & Humphreys, 1994; also see Tarr, Kersten, & Bülthoff, 1998, for an alternative account).

Although recognition of common objects may be predominantly edge-based, there is now ample evidence that surface information is important for face recognition. Studies have shown that faces are better recognized in photographs that provide shading information than in line drawings that lack such information (Bruce, Hanna, Dench, Healey, & Burton, 1992; Davies, Ellis, & Shepherd, 1978; Leder, 1996). Furthermore, the use of unt textured 3-D laser-scanned faces devoid of pigmentation has shown that faces can be identified from 3-D facial surface information defined mainly by shape from shading (Bruce, Burton, & Doyle, 1992). Finally, it is well-known that face recognition can be disrupted by changing contrast polarities (Bruce & Langton, 1994; Galper, 1970;

This study was supported by research grants to A.C. from the Medical Research Council (MRC) of Canada and the Natural Sciences and Engineering Research Council (NSERC) of Canada. We are grateful to Vicki Bruce and Mike Burton for kindly providing us with the face database used in Experiment 3. We also thank our reviewers for their constructive comments on an earlier version of this article. Correspondence concerning this article should be addressed to C. H. Liu, Department of Psychology, University of Hull, Cottingham Road, Hull HU6 7RX, England (e-mail: c.h.liu@hull.ac.uk).

Galper & Hochberg, 1971; Hayes, 1988; Johnston, Hill, & Carman, 1992; Kemp, Pike, White, & Musselman, 1996; Liu & Chaudhuri, 1998; Phillips, 1972) or lighting direction (Hill & Bruce, 1996; Johnston et al., 1992; Liu, Collin, Burton, & Chaudhuri, 1999), even when edge information under these circumstances is identical. Together, the evidence supports the view that recognition of homogeneous 3-D objects such as faces at the level of individual identity depends on the encoding of surface geometry. This information may be acquired from a number of depth cues, such as shape from shading, stereopsis, motion parallax, and texture gradients.

The purpose of this study was to determine whether statue-like face stimuli devoid of color information could be identified from texture gradient information alone. To date, there has been little research that directly addresses the role of texture gradients in identifying objects with complex surfaces. However, there have been studies on the role of dynamic information in face perception that may be relevant to this issue. Using a method developed by Johansson (1973), a number of studies (Bassili, 1978, 1979; Bruce & Valentine, 1988; Hill, Jinno, & Johnston, 2003) have examined face perception in static and dynamic point-light displays. Their displays consisted of small dots attached to faces and then filmed in such a way that the facial surfaces were concealed, so only the dots could be seen by the observers. Most of Bassili's (1978, 1979) participants were able to realize that the stimuli represented faces and were able to discriminate certain facial expressions when the stimuli were presented in motion. Bruce and Valentine's participants were also able to recognize facial expressions in such dynamic displays. In addition, their participants could recognize the sex and identity of familiar faces above chance level. Participants in Hill et al.'s (2003) study were able to judge the sex of faces above chance level as well using dynamic point-light displays.

Because the point-light displays consisted of stochastically regular texture elements (dots), it may be argued that the overall 3-D shape of a face may have been revealed by the texture gradient information, so that participants in these studies were assisted by this information in their tasks. However, a notable result of all of these studies was that recognition performance was extremely poor when the point-light displays were static. Although participants in static conditions could still discriminate sex and some facial expressions, they failed to judge the identity of faces better than chance, even with familiar faces (Bruce & Valentine, 1988). From these results, it may be concluded that texture gradient information alone does not play an important role in recognizing facial identity, but that it may do so when combined with motion.

An alternative account is that texture gradients may have been more useful if the point-light display had been designed to produce a stronger and more detailed impression of depth. According to this argument, the reason that these studies did not detect the role of texture gradients was that the images did not have sufficient resolution. Resolution in this case can be simply defined as

the number of texture elements (dots) per object (face). The four studies reviewed above used 100 dots per face at most, a resolution level that may have been too low for recognizing faces in the static displays. Image resolution has been shown to be very important in other kinds of displays, such as those featuring shape from shading. Studies using blocking or low-pass filtering have shown that face identification using shape from shading can be disrupted or may even be impossible when the resolution is below a certain limit (Costen, Parker, & Craw, 1994, 1996; Harmon, 1973; Peli, Lee, Trempe, & Buzney, 1994).

To study the potential role of texture gradient in recognition, we applied fractal noise to the surface of 3-D face models¹ (see examples in Figure 1). The resolution was greatly increased compared to previous studies through fine gradients at different luminance levels of the fractal-noise patterns. The improved resolution permitted a more effective examination of the role of static texture gradient information and whether it alone can be used in facial identification.

EXPERIMENT 1

The primary goal of this experiment was to find out whether statue-like faces defined by static texture gradients could be identified better than chance. We know that recognition of unfamiliar faces is more difficult than recognition of familiar ones due to the susceptibility of faces to variations of lighting, pose, level of spatial frequency overlap, and representational media (e.g., Bruce, Valentine, & Baddeley, 1987; Hill & Bruce, 1996; Liu et al., 1999; Liu, Collin, Rainville, & Chaudhuri, 2000; Rhodes, Brennan, & Carey, 1987; Troje & Bühlhoff, 1996). To ensure that face shape recognition from texture gradients was genuinely related to face processing, participants were not only required to match faces in different representational media (i.e., faces defined by texture gradients or shape from shading), but also to match faces in different poses. Participants were thus prevented from using strategies, such as matching the silhouettes of the face stimuli, that are not directly related to face processing. Based on the fact that even familiar faces in static point-light displays were not recognized beyond chance level (Bruce & Valentine, 1988), we expected the task to be very difficult. Thus, to maximize the detectability of texture gradient effects, participants were allowed to take as much time as needed for their judgments.

We were also interested in determining the viewing angles of faces that were more easily identified, in the event that participants performed better than chance. We thus displayed faces at a wide range of angles along both the horizontal and vertical meridians.

We chose a paradigm in which participants were required to match faces defined by texture gradients to those defined by shape from shading. Matching between different depth cues (i.e., texture gradient vs. shape from shading) was employed to examine whether facial shape could be generalized from one cue to another.

Figure 1. Example faces with shape from texture and shape from shading. (A) Horizontal rotation at -30° . (B) Full-face view at 0° . (C) Vertical rotation at -20° . (D) Vertical rotation at 30° . (E) Full-face view at 0° . (F) Horizontal rotation at 20° . In Experiment 1, faces with shape from texture were targets, whereas faces with shape from shading were used as matching stimuli. In Experiment 2, both target and matching faces were defined by shape from texture. In Experiment 3, both target and matching faces were defined by shape from shading.

Method

Participants. Twenty-three undergraduate students from the University of Hull participated in this study. All had normal or corrected-to-normal vision. None was familiar with the 3-D face models.

Materials. We used a 3-D face database obtained from the Max Planck Institute for Biological Cybernetics in Tübingen, Germany. Each face in this database contained over 262,000 x -, y -, and z -coordinates. A full description of this face database is found in Troje and Bühlhoff (1996). In this study, eight male and eight female faces from the set were used.

Two types of stimuli were created from the 3-D face database. One set were in texture gradient and were used as target faces, the other set were shape from shading and were used as matching faces. Face images defined by texture gradients were generated on 3D Studio Max R3 using a fractal-noise texture map. Default parameters were used for the fractal-noise texture (Threshold High = 1.0, Low = 0.0, and Levels = 3.0). The size of each texture element

was 23 distance units in the 3D Studio environment. The head models averaged 200 distance units in height. The calculation of texture size was based on surface area rather than projected distances to ensure that a texture was evenly spread across the surface. Texture was applied procedurally, meaning that each texture element was generated individually on the face surface and was created to tile seamlessly with adjacent texture elements. No shading was used. The face stimuli were set to maximum self-illumination. The resulting face stimuli resembled marble statues. Figure 1 shows some examples of the texture gradient faces at four angles. Each fractal-noise image has a mean gray value of 128 ($SD = 24$), and the spectral energy profiles (i.e., the RMS contrasts at different spatial frequencies) closely approximate the function of $1/f^2$, which is typical of the energy distribution of natural scenes.

Each face was rendered in 17 views. Among these were eight spanning the horizontal and eight along the vertical meridian of the face. The origin (0°) was shared by both horizontal and vertical

meridians. The angles employed for both meridians were -40° , -30° , -20° , -10° , $+10^\circ$, $+20^\circ$, $+30^\circ$, and $+40^\circ$. Negative horizontal values represent views facing left, whereas negative vertical values represent views facing down.

The size of the face images on the monitor was about 90×65 mm for the full-face views. The height in this measure was taken from the top of the forehead to the chin, and the width was measured between the outer contours of cheekbones. There was a minor variation among the size of face stimuli, which reflected the veridical differences among the real faces from which the 3-D laser scans were obtained. Because of foreshortening, the height and width of each image assumed different values depending on the angular rotation.

Each face defined by shape from shading was captured in three views—one full-face view (0°) and two three-quarter views (-22.5° and $+22.5^\circ$)—resulting in a total of 48 images (16 faces $\times$ 3 views) to form the set of matching stimuli. The views were produced by setting the angle between the central axis of the face and the y -axis of object space. All faces were illuminated at 45° above the horizontal meridian of the faces.

The stimuli were presented on a 21-in. Apple monitor. The screen resolution was set to $1,024 \times 768$ pixels. Three windows were created to hold the stimuli and to record the responses: One window was designated for the target face, one for the eight matching faces with shape from shading, and one for entering the response. The sizes of these windows were 19×17 , 26×17 , and 7×17 cm, respectively. The response window was always displayed beneath the target window. The width of each window covered half of the screen width. The target face was presented in the center of the target window, and the eight matching faces were arranged in 2 columns $\times$ 4 rows evenly distributed in the matching window. The display was freely viewed binocularly at a distance of 60 cm, from which point the target and matching faces subtended approximately $6.5^\circ \times 9.0^\circ$ and $5.7^\circ \times 4.3^\circ$ arc, respectively.

Design and Procedure. An eight-alternative forced choice task was used in which participants were instructed to match the target face with shape from texture to one of eight matching faces with shape from shading. The gender of target and matching faces was always the same. Each of the 16 matching faces was assigned a number between 1 and 16 that appeared at the top left corner. On each trial, the positions of the faces in the matching window were randomized. Participants were required to type the number of the face in the response window that best matched the face in the target window. Participants were told to respond as accurately as possible.

For half of the participants, the target and response windows were displayed on the left and the matching window on the right side of the screen. For the remaining half, this configuration was reversed. To prevent participants from using the silhouettes of faces as a cue for their judgments, the target and matching faces always appeared in different views. Specifically, when the target face was in any view other than the full-face view, the matching faces were presented only in full-face view. When the target face was in a full-face view, the matching faces were presented in three-quarter (22.5°) view, either to the left or to the right, as determined randomly. Because the left and the right views were more or less symmetric, only half of the eight angles that rotated away from 0° were tested for the horizontal conditions. This was done by randomly selecting a direction (either facing left or facing right) for each of the four view angles (i.e., 10° , 20° , 30° , and 40°). Thus, a total of 13 views were tested, in which 4 varied in horizontal direction, 4 in upward direction, and 4 in downward direction, in addition to the view at 0° . The 13 views were tested in a block of 13 trials. The order of the 13 angles was completely randomized within each block. The 13 angles were randomly assigned to 13 of the 16 available faces for each block. A constraint in this random assignment ensured that roughly equal numbers (6 or 7) of the male and female faces were selected. Each participant was tested on four blocks. It took on average about 30 min to complete the four blocks of trials.

Results and Discussion

The mean accuracy score for each of 13 angles across four blocks was computed within each participant. Figure 2 shows the results based on these mean scores. The grand mean across these conditions was 36.9% ($SD = 10.5$), which was 24.4% higher than the chance level (12.5%).

To examine whether there were effects of tested views, a two-way repeated measures analysis of variance (ANOVA) was performed on the data set in Figure 2, using three levels of head rotation (horizontal, upward, and downward) and four levels of face angle (10° , 20° , 30° , and 40°). The data for the 0° view were excluded from this analysis. The main effect of head rotation was not significant [$F(2,44) = 0.02$, $p = .98$], but the main effect of face angle was [$F(3,66) = 14.66$, $p < .001$]. The interaction between these two factors was not significant [$F(6,132) = 0.62$, $p = .72$].

Further analyses using Tukey post hoc tests ($\alpha = .05$) revealed that faces at a 10° angle (50.4%) were identified better than faces at 20° (36.6%), 30° (35.1%), or 40° (26.4%). Furthermore, faces angled at 20° were identified better than faces angled at 40° . The rest of the pairwise comparisons were not significant.

The upward and downward conditions can be considered a continuum along the vertical rotation axis around 0° , and the 0° view can likewise be considered part of the horizontal rotation axis. We therefore performed two separate one-way ANOVAs, one for the vertical and the other for the horizontal rotations, taking the data at the 0° view into account in both cases. The main effects were significant both for vertical rotation [$F(4,88) = 2.89$, $p < .03$] and horizontal rotation [$F(8,176) = 3.81$, $p < .001$]. Tukey tests showed that the identification rate was significantly higher for 10° than for 40° faces for horizontal rotation. Faces rotated 10° upward were identified significantly better than were those rotated 30° or 40° upward or 40° downward. These results are not surprising if performance is predicted according to the amount of angular rotation separating the matching views: When a 10° target face was used, the angular difference between it and matching faces (at 0°) was smallest among all pairings.

To examine whether there was a practice effect over the four blocks, a one-way repeated measures ANOVA was applied to the overall accuracy score for each participant in each block. No significant differences were found between the blocks [$F(3,66) = 0.44$, $p = .72$].

The data show that participants were able to identify unfamiliar faces defined by static texture gradients at better than a chance level. This result was quite surprising considering the previous finding that even familiar faces were not recognized better than chance in static point-light displays (Bruce & Valentine, 1988). It would appear that facial identity can be ascertained when texture images contain finer gradient information. Our analysis showed that this result was not due to a practice effect over the four blocks of trials but rather was present in all blocks.

Figure 2. Accuracy as a function of viewing angle in three directions in Experiment 1. The task was to match faces defined by texture to faces defined by shading. The dotted line represents the chance performance level. Error bars represent standard errors.

We found that identification performance depended on the direction and angle of head rotations. The 10° rotation from matching faces generated best performance in this experiment. The results thus show that performance depends on angular difference between a target and matching faces, as expected from previous literature (Troje & Bühlhoff, 1996). The direction of rotation did not affect performance in this experiment.

EXPERIMENT 2

It is clear that both shape-from-texture and shape-from-shading displays contributed to the matching results in Experiment 1. To ascertain whether texture gradients alone could be used to identify statuelike faces and to assess the effects of matching faces based on different types of depth cues (i.e., shape from texture and shape from shading), we employed the task from Experiment 1 again using only shape-from-texture displays.

Methods

Participants. Twenty-one undergraduate students from the University of Hull participated in this study. All had normal or corrected-to-normal vision.

Materials. The matching faces defined by shape from shading in Experiment 1 were replaced by faces defined by texture gradient, created using the same methods described in Experiment 1. The fractal-noise texture map was generated individually for each matching face, so that the matching stimuli had unique and different textures from those appearing on the target faces. The target faces were identical to those in Experiment 1.

Design and Procedure. These were the same as in Experiment 1.

Results and Discussion

The mean accuracy score for each of 13 angles across four blocks was computed within each participant. Figure 3 shows the results based on these mean scores. As the figure indicates, performance in general was well above chance. The grand mean score across the conditions was 30.5% ($SD = 7.0$), which was 18.0% above the chance level.

As in Experiment 1, we examined whether there were effects of tested views by a two-way repeated measures ANOVA using three levels of head rotation (up, down, and horizontal) and four levels of face angle (10°, 20°, 30°, and 40°). Both main effects of head rotation and face angle were significant [$F(2,40) = 4.15, p < .05$, and $F(3,60) = 2.75, p < .05$, respectively]. The interaction between these two factors was not significant [$F(6,120) = 0.29, p = .94$].

Further analyses using Tukey HSD post hoc tests revealed that recognition for the downward (36.3%) and horizontal (29.2%) rotations were better than for the upward rotation (26.2%). Recognition of faces angled at 10° (36.9%) was better than for those at 40° (25.8%). The rest of the pairwise comparisons were not significant.

Two separate one-way ANOVAs, one for horizontal and the other for vertical rotations, were also applied to the data, taking the data at the 0° view into account in each case. The main effect was not significant for either vertical rotations [$F(8,20) = 1.83, p = .08$] or horizontal rotations [$F(4,20) = 0.71, p = .59$].

Again, as was shown in Experiment 1, participants were able to identify faces defined by static texture gra-

Figure 3. Accuracy as a function of viewing angle in three directions in Experiment 2. The task was to match faces defined by texture. The dotted line represents the chance performance level. Error bars represent standard errors.

dients well beyond the chance level, although performance was slightly poorer than in Experiment 1. To verify the difference, we performed two post hoc analyses on the results for horizontal and vertical rotations, treating experiments as a between-participants variable. For the vertical rotations, the accuracy in Experiment 1 (37.0%) was significantly higher than in Experiment 2 (31.1%) [$F(1,42) = 4.15, p < .05$], and there was a significant main effect of vertical rotation [$F(8,336) = 4.05, p < .001$]. The interaction of these two factors was not significant [$F(8,336) = 1.49, p = .16$]. For the horizontal rotations, there was no significant difference between accuracy in Experiment 1 (36.1%) and in Experiment 2 (29.3%) [$F(1,42) = 2.63, p = .11$]. The main effect of rotation was significant [$F(4,168) = 3.11, p < .05$]. There was no significant interaction [$F(4,168) = 0.39, p = .81$].

The post hoc ANOVA results suggest that matching faces using different types of displays (shape-from-texture and shape-from-shading) was not more difficult than matching faces using shape-from-texture displays only. On the contrary, matching faces having texture with faces having shading produced a slightly higher accuracy.

EXPERIMENT 3

To assess the degree to which the poor performance in our previous experiments can be attributed to texture gradient limitations, we report this experiment in which participants performed the same matching task as Experiments 1 and 2, except that both target and matching faces were presented in shape-from-shading format. Be-

cause this experiment was carried out in a separate study, the laser-scanned faces were not from the same face database. However, because the test procedure and conditions in this experiment were identical to those in Experiments 1 and 2, this experiment should still provide a fairly good comparison between the effects of the two depth cues.

A few prior studies have used faces devoid of pigmentation, similar to the shape-from-shading displays in this study, to assess how shape-from-shading information alone is used in face recognition under different views (Hill, Schyns, & Akamatsu, 1997; Troje & Bülthoff, 1996). However, a main difference between the present and previous tasks was that the prior studies used a sequential matching paradigm in which a memory load was required. Furthermore, we also evaluated the effects of vertical rotation of views on identification. To our knowledge, this manipulation has not been examined in untextured faces using this paradigm.

Methods

Participants. Twelve undergraduate students from McGill University participated in this study. All had normal or corrected-to-normal vision. None was familiar with the 3-D face models.

Materials. The laser-scanned faces were originally developed at University College London by recording 3-D surface information with a laser beam. Over 20,000 *x*-, *y*-, and *z*-coordinates were recorded for each face, giving a very detailed description of the surface. The coordinates were connected by polygons, each having four vertices. This face database has been used in a number of prior studies, where a full description of the stimuli may be found (Bruce et al., 1991; Bruce & Langton, 1994; Hill & Bruce, 1996; Moss, Linney, Grindrod, Arridge, & Clifton, 1987). The face images were created on a Silicon Graphics workstation using 3-D rendering software called Geomview 2.0 (www.geom.umn.edu/projects/visualization).

The projection from the 3-D faces to the camera view assumed perspective. The camera distance was three times the face width. A Gouraud shading function was used to apply gradually interpolated shading to each facet and thus to smooth the appearance of the polygonal surface. The standard OpenGL illumination model (see Woo, Neider, & Davis, 1997) was used with a single light source at infinity and a small amount of ambient light. The relative intensities of these two components were 0.8 and 0.2, respectively. These values relate to the OpenGL lighting models and represent coefficients to the formulae for ambient and diffuse light. They were chosen to mimic common daylight conditions, in which a strong directional light from the sun is combined with a weaker ambient light from the sky. The target faces were created by the same method that was used to produce the matching faces. The 17 views of each face used as target stimuli were identical to those used in the texture-gradient format specified in Experiment 1. The set of target stimuli consisted of 272 images ($16 \text{ faces} \times 17 \text{ views}$). The target stimuli were enlarged to match the size of target stimuli used in Experiment 1. The faces were rendered with a material having a diffuse reflection factor of 1.0, a specular reflection factor of 0.3, an ambient reflection factor of 0.3, and a specular reflection exponent of 13.7. No space variant albedo was applied.

The captured images were cropped using Adobe Photoshop 5.0. An elliptical shape was used to clip off parts of the face above the hairline and below the chin. All images were then reduced from 16-bit RGB format to 8-bit grayscale using Graphic Converter 2.2 for Macintosh (Lemke Software, www.lemkesoft.de). Matching stimuli were scaled to a size of about $57 \times 43 \text{ mm}$ for a full-face view. Because matching faces were scaled to the same size, it should have been difficult for the participants to use the size variation of the tar-

get faces for their judgments. Figures 1E and 1F show two examples of the shape-from-shading matching stimuli.

Design and Procedure. These were the same as in Experiment 1. Each participant required 30 min on average to complete the trials.

Results and Discussion

As in Experiment 1, the mean accuracy score for each of 13 angles across four blocks was first computed within each participant. The results of this calculation are presented in Figure 4. Overall, the accuracy was very high. The only exception was the data point at 40° of downward rotation (37.5%). The remaining means ranged from 70.8% to 89.6% .

As in Experiment 1, a $3 \text{ (head rotation)} \times 4 \text{ (angle)}$ repeated measures ANOVA was applied to the results in Figure 4, excluding the data at the 0° angle. Both the effects of head rotation and face angle were significant [$F(2,22) = 3.61, p = .04$, and $F(3, 33) = 12.94, p < .0001$, respectively]. There was also a significant interaction between these two factors [$F(6,66) = 3.45, p < .005$].

Tukey HSD post hoc tests revealed that faces at 40° of downward rotation were recognized more poorly than any of the other 12 views tested. All other pairwise comparisons were not significant. Clearly, this uneven effect of angular change on identification for different directions of rotation was the source of the significant interaction.

Figure 4. Accuracy as a function of viewing angle in three directions in Experiment 3. The task was to match faces defined by shading. The dotted line represents the chance performance level. Error bars represent standard errors.

The data at the 0° angle were included in two separate one-way ANOVAs performed on data obtained from horizontal and vertical rotations. The main effect for horizontal rotation approached the level of significance [$F(4,44) = 2.39, p = .07$], reflecting a drop of accuracy at 30° and 40°. The main effect for vertical rotation was significant [$F(8,88) = 8.43, p < .0001$]. Tukey HSD tests found that accuracy at 40° of downward rotation was worse than accuracy at any of the other angles. All other pairwise comparisons were not significant.

Compared with the results in Experiments 1 and 2, recognition in shape-from-shading displays was much superior. Even the lowest recognition performance (37.5% at 40° of downward rotation) was better than, or equivalent to, most of the mean scores from the 13 angles in Experiments 1 and 2. The superior performance in this experiment suggests that the difficulty involved in Experiments 1 and 2 was largely due to the limited information contained in the shape-from-texture images.

GENERAL DISCUSSION

The goal of this study was to determine whether recognition of complex 3-D objects such as faces could be accomplished by texture gradients alone. Our main findings were as follows: First, unfamiliar faces were identified at reliably better than chance level in shape-from-texture displays (Experiments 1 and 2). Second, matching faces with shape from texture to faces with shape from shading (Experiment 1) was easier than matching two faces both defined by shape from texture (Experiment 2). Finally, faces in shape-from-texture displays were much more difficult to recognize than faces in shape-from-shading displays (Experiment 3).

Compared with point-light displays used in the prior studies, the fractal-noise displays in our study provide finer gradient information, which may explain why our participants were able to perform much better than those in previous studies. The role of texture gradients may be largely attributable to the occluding edges that reveal key facial features. For example, in Figure 1A, the occluding edges from horizontal rotation reveal vertically oriented facial features, such as the nose line on the left and the boundary of the lips. In addition to occluding edges, fractal-noise displays also seem to produce fairly good descriptions of surface orientation.

However, in comparison with the results based on shape from shading, the role of texture gradients in recognition shows clear limitation. The difficulty of identifying faces with shape from texture is fairly obvious. At certain angles or for certain facial areas, the shape-from-texture display provides very little information about facial features. In contrast, facial features are defined quite well in the shape-from-shading display. This difference may be due to an inherent deficiency of shape-from-texture in comparison with shape-from-shading information for this kind of 3-D object. Although texture gradients are useful for detecting the ori-

entation of relatively simple surfaces, such as planar shapes, cubes, or simple waveforms, highly complex 3-D objects such as faces may be harder to describe by texture gradients alone. Todd and Akerstrom (1987) showed that a complex, undulating surface covered with texture elements produced only a very weak impression of its underlying 3-D shape. They pointed out that the perception of 3-D form from texture gradients tended to break down if the depicted surface was too complex.

We conjecture that the diminished role of texture gradients in complex object recognition is attributable to gradient information that is insufficient to fully describe the surfaces. That is, the number of texture elements per unit of surface area may be insufficient in our texture gradient faces to capture the actual changes in surface orientation within relatively small regions. Yet, such changes often contain crucial information; for example, the subtle surface structure of the eyes consists of the curved surface of eyeballs and the thin eyelids lying on top. It is very difficult to describe the abrupt change of surface orientation at the border of eyeballs and eyelids because the narrow surface of the eyelids is difficult to describe by texture gradient alone. To display more texture elements within this narrow surface would require enlarging its area, or the texture elements would be too small to be seen. However, enlarging the area would also require scaling the entire face image to a much larger size. Shape-from-shading information, on the other hand, can reveal this 3-D structure more easily through an accelerated change of luminance values within the narrow surface of the eyelids.

A comparison of the results in Experiments 1 and 2 shows that the difficulty of the tasks involving shape-from-texture displays cannot be attributed to face matching in different representational media (i.e., between faces defined by texture gradients and shape from shading). In fact, our results show that matching in different representational media produced a more reliable advantage than face matching defined by texture gradients. This result appears to be incompatible with the principle of encoding specificity, which predicts better performance for similar matching stimuli (Tulving & Thomson, 1973), or with the idea that recognition of unfamiliar faces relies on low-level image similarity. However, the result is analogous to a previous finding that face matching with images of incongruent resolution (i.e., between low and high resolution) produces an advantage over images of the same but low resolution (Liu, Seetzen, Burton, & Chaudhuri, 2003). Although it is likely that the greater facial information contained in the shape-from-shading displays is not present in the shape-from-texture displays, such information may allow the option of interpolating some of the missing information in the shape-from-texture displays.

The difference between Experiments 1 and 2 also suggests that any remaining similarity between silhouettes of the targets and the matching stimuli also cannot be the determinant of the performance. Had the participants re-

lied on silhouettes alone, their performance in Experiments 1 and 2 would have been identical. We do not completely rule out the possibility that silhouettes played a role in our results, but the magnitude of their influence must have been quite negligible, given both our analysis of Experiments 1 and 2 and the general difficulty of matching silhouettes in different views.

Although shape from texture is less effective than shape from shading for defining complex objects such as faces, it may be easier to extract invariant information from texture gradients. An object at any single angle can be presented in an infinite number of lighting conditions, resulting in many possible intensity maps that vary both locally and globally. In contrast, the global pattern of the texture gradients on an object viewed from one angle is more or less similar regardless of lighting conditions, despite the fact that the local positions of the particular individual texture elements may vary. Therefore, it would be easier to derive invariant properties from texture gradient than from shading displays.

In conclusion, although evidence from prior studies has shown that face identification was impossible in static point-light displays, we have shown here that recognition performance is considerably better if such displays have sufficient resolution. Texture gradient information is especially effective when it forms occluding edges that reveal salient facial features.

REFERENCES

- BASSILI, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception & Performance*, **4**, 373-379.
- BASSILI, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality & Social Psychology*, **37**, 2049-2058.
- BRAUNSTEIN, M. L., & PAYNE, J. W. (1969). Perspective and form ratio as determinants of relative slant judgments. *Journal of Experimental Psychology*, **81**, 584-590.
- BRUCE, V., BURTON, M., & DOYLE, T. (1992). Faces as surfaces. In V. Bruce & M. Burton (Eds.), *Processing images of faces* (pp. 232-252). Norwood, NJ: Ablex.
- BRUCE, V., HANNA, E., DENCH, N., HEALEY, P., & BURTON, M. (1992). The importance of "mass" in line drawings of faces. *Applied Cognitive Psychology*, **6**, 619-628.
- BRUCE, V., HEALEY, P., BURTON, M., DOYLE, T., COOMBS, A., & LINNEY, A. (1991). Recognizing facial surfaces. *Perception*, **20**, 755-769.
- BRUCE, V., & HUMPHREYS, G. W. (1994). Recognizing objects and faces. In V. Bruce & G. W. Humphreys (Eds.), *Object and face recognition* (pp. 141-180). Hove, U.K.: Erlbaum.
- BRUCE, V., & LANGTON, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception*, **23**, 803-822.
- BRUCE, V., & VALENTINE, T. (1988). When a nod's as good as a wink: The role of dynamic information in facial recognition. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 169-174). London: Wiley.
- BRUCE, V., VALENTINE, T., & BADDELEY, A. D. (1987). The basis of the 3/4 view advantage in face recognition. *Applied Cognitive Psychology*, **1**, 109-120.
- CAVANAGH, P. (1987). Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity, and shape. *Computer Vision, Graphics & Image Processing*, **37**, 171-195.
- CHRISTOU, C., KOENDERINK, J. J., & VAN DOORN, A. J. (1996). Surface gradients, contours, and the perception of surface attitude in images of complex scenes. *Perception*, **25**, 701-713.
- COSTEN, N. P., PARKER, D. M., & CRAW, I. (1994). Spatial content and spatial quantisation effects in face recognition. *Perception*, **23**, 129-146.
- COSTEN, N. P., PARKER, D. M., & CRAW, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics*, **58**, 602-612.
- DAVIES, G., ELLIS, H. D., & SHEPHERD, J. (1978). Face recognition accuracy as a function of mode of representation. *Journal of Applied Psychology*, **63**, 180-187.
- GALPER, R. E. (1970). Recognition of faces in photographic negative. *Psychonomic Science*, **19**, 207-208.
- GALPER, R. E., & HOCHBERG, J. (1971). Recognition memory for photographs of faces. *American Journal of Psychology*, **84**, 351-354.
- GIBSON, J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- HARMON, L. D. (1973, November). The recognition of faces. *Scientific American*, **229**, 71-82.
- HAYES, A. (1988). Identification of two-tone images; some implications for high- and low-spatial-frequency processes in human vision. *Perception*, **17**, 429-436.
- HILL, H., & BRUCE, V. (1996). The effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 986-1004.
- HILL, H., JINNO, Y., & JOHNSTON, A. (2003). Comparing solid-body with point-light animations. *Perception*, **32**, 561-566.
- HILL, H., SCHYNS, P. G., & AKAMATSU, S. (1997). Information and viewpoint dependence in face recognition. *Cognition*, **62**, 201-222.
- JOHANSSON, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, **14**, 201-211.
- JOHNSTON, A., HILL, H., & CARMAN, N. (1992). Recognizing faces: Effects of lighting direction, inversion, and brightness reversal. *Perception*, **21**, 365-375.
- KEMP, R., PIKE, G., WHITE, P., & MUSSELMAN, A. (1996). Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. *Perception*, **25**, 37-52.
- KNILL, D. C. (1998). Surface orientation from texture: Ideal observers, generic observers, and the information content of texture cues. *Vision Research*, **38**, 1655-1682.
- LEDER, H. (1996). Line drawings of faces reduce configural processing. *Perception*, **25**, 355-366.
- LI, A., & ZAIDI, Q. (2003). Observer strategies in perception of 3-D shape from isotropic textures: Developable surfaces. *Vision Research*, **43**, 2741-2758.
- LIU, C. H., & CHAUDHURI, A. (1998). Are there qualitative differences in face recognition between positive and negative? *Perception*, **27**, 1107-1122.
- LIU, C. H., COLLIN, C. A., BURTON, A. M., & CHAUDHURI, A. (1999). Lighting direction affects recognition of untextured faces in photographic positive and negative. *Vision Research*, **39**, 4003-4009.
- LIU, C. H., COLLIN, C. A., RAINVILLE, S. J. M., & CHAUDHURI, A. (2000). The effects of spatial frequency overlap on face recognition. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 956-979.
- LIU, C. H., SEETZEN, H., BURTON, A. M., & CHAUDHURI, A. (2003). Recognition is robust with incongruent image resolution: Relationship to security video images. *Journal of Experimental Psychology: Applied*, **9**, 33-41.
- MOSS, J. P., LINNEY, A. D., GRINDROD, S. R., ARRIDGE, S. R., & CLIFTON, J. S. (1987). Three-dimensional visualization of the face and skull using computerized tomography and laser scanning techniques. *European Journal of Orthodontics*, **9**, 247-253.
- PELI, E., LEE, E., TREMPER, C. L., & BUZNEY, S. (1994). Image enhancement for the visually impaired: The effects of enhancement on face recognition. *Journal of the Optical Society of America A*, **11**, 1929-1939.
- PHILLIPS, R. J. (1972). Why are faces hard to recognize in photographic negative? *Perception & Psychophysics*, **12**, 425-426.
- RHODES, G., BRENNAN, S., & CAREY, S. (1987). Identification and ratings of caricatures: Implications for mental representation of faces. *Cognitive Psychology*, **19**, 473-497.
- ROSENHOLTZ, R., & MALIK, J. (1997). Surface orientation from texture: Isotropy or homogeneity (or both)? *Vision Research*, **37**, 2283-2293.
- TARR, M. J., KERSTEN, D., & BÜLTHOFF, H. H. (1998). Why the visual

- recognition system might encode the effects of illumination. *Vision Research*, **38**, 2259-2275.
- TODD, J. T., & AKERSTROM, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 242-255.
- TODD, J. T., & OOMES, A. H. J. (2002). Generic and non-generic conditions for the perception of surface shape from texture. *Vision Research*, **42**, 837-850.
- TROJE, N. F., & BÜLTHOFF, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research*, **36**, 1761-1771.
- TULVING, E., & THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352-373.
- WOO, M., NEIDER, J., & DAVIS, T. (1997). *OpenGL programming guide*. New York: Addison-Wesley.
- ZAIDI, Q., & LI, A. (2002). Limitations on shape information provided by texture cues. *Vision Research*, **42**, 815-835.

NOTE

1. We used facial surfaces merely as examples to illustrate the kind of role texture gradients may play in recognition of complex objects, particularly when they are to be identified at subordinate or exemplar level. Our selection of textures in this sense was more or less arbitrary, although there is a close resemblance between our fractal-noise stimuli and textured marble statues. We do not imply that the fractal-noise patterns used in our study are anywhere close to real facial textures or that genuine facial textures produce the same kind of effects found in our study. Undoubtedly, the role of texture gradients also depends on the type and salience of the textured surface.

(Manuscript received June 12, 2003;
revision accepted for publication April 12, 2004.)