Guidance-based Two-Dimensional Haptic Contour Rendering for Accessible Photography

Jongho Lim, Yongjae Yoo, and Seungmoon Choi

***Abstract*— Touching the faces of loved ones on a photograph can be a strong means to invoke associated pleasant memories to visually-impaired users, as seeing is to sighted users. Inspired by this long-term goal of multimodal accessible photography, we present in this paper a two-dimensional contour rendering algorithm based on haptic guidance using force feedback for tactile graphics. We begin with a basic algorithm for one contour and find its optimal guidance parameter values by a user study. This basic algorithm is then applied to rendering geometric primitives, and its perceptual performance is evaluated by a shape identification experiment. We extend the basic algorithm to cover multiple contours and apply it to rendering portrait photographs. Major facial features are extracted from photographs, and their contours are rendered by our algorithms. It is demonstrated that users can accurately recognize emotional expressions on the faces, suggesting the high potential of our approach for accessible photography.**

I. Introduction

This work stemmed from our on-going project on accessible photography named TouchPhoto. We envision an integrated system that provides visually-impaired users with multimodal assistance in order to support their independent undertaking of photo-related activities. Our initial prototype, which is described in a companion paper [1] with full detail, consists of two main modules. One module runs on a regular smartphone and enables a visually-impaired user to take portrait photographs independently by providing speech guidance for camera aiming. The user is also allowed to add several types of audio recordings to facilitate later recall of the memory and experiences associated with the photographs. The other module is a large electrostatic variable friction display for tactile graphics. The user can explore a photograph in detail by listening to audio tags and also feeling main facial features in the photograph expressed by electrovibration.

The addition of the tactile display was to emphasize the social and emotional function of photography, as a way to remember people and related events. For example, imagine a mother with visual impairment touching and feeling the face of her baby in an old photograph, who is now 20 years old and about to leave the home. Reviving such personal, emotional, often intimate memories conforms to the general nature of tactile sense. However, rendering facial landmarks in portrait photographs using an electrovibration display turned out to be not very effective in delivering pictorial information (see [1] for details). The goal of this work was to find an effective and simple method that enables users to perceive tactile graphics composed of contours, without any visual information, as a way to improve the haptic rendering function of TouchPhoto.

One of the most important requirements is related to the form factor of a haptic device. We wish to use a small and light haptic device for easy integration into a smartphone or tablet, which visually-impaired users carry everyday. This convenience of installation was the main reason we chose an electrovibration display in our first attempt [1]. Our next choice is a two-dimensional (2D), small, light force-feedback device, e.g., those driven with cables [2], mounted on a touchscreen device. This type of device can exert both assistive and resistive lateral force, while only resistive lateral force is available with electrovibration displays.

In broader context, our work is part of research efforts for tactile graphics. Visual objects, such as pictures, maps, and graphs, are converted to tactile representations so that visually-impaired users perceive them using hands. Tactile graphics generally goes through two steps. First, a tactile representation is made by reducing or modifying the visual content, while considering the limited information transmission capacity of tactile sense [3]. Second, the tactile content is printed on a physical medium, e.g., by embossing papers using a special machine, or rendered using a haptic interface, e.g., Optacon [4] and other mechanical ones [5], [6], [7], [8], [9].

Some prior studies used a force feedback interface for tactile graphics. For example, Jason et al. implemented a haptic rendering technique to deliver 2D and 3D graphic data using a PHANToM [5]. Jansson et al. used a PHANToM to render several 3D shapes (cylinders, cones, cube, and spheres) of different sizes and measured users' recognition accuracies and task completion times [10]. Yu et al. tried to transmit information about graphs to visually-impaired people by rendering graphs with different textures or grid lines using a PHANToM [11].

The rendering algorithm we present in this paper is for rendering 2D contours, and it is based on *haptic guidance* to help visually-impaired users easily follow the contours. For haptic guidance, a force-feedback device exerts active force to the user to assist task execution [12]. Our guidance-based contour rendering algorithm strictly assumes that no visual cues are available. Also, our method is not to facilitate motor learning; it uses haptic guidance to aid cognition, somewhat similar to a previous approach utilizing the haptic sensory channel to improve the memorization performance [13].

In Section II, we introduce an elementary 2D contour

This research was supported by the X-Project program (NRF-2016R1E1A2914792) from the National Research Foundation of Korea.

Authors are with the Department of Computer Science and Engineering at Pohang University of Science and Technology (POSTECH), Korea.
E-mail: {lcdplayer, dreamseed, choism}@postech.ac.kr

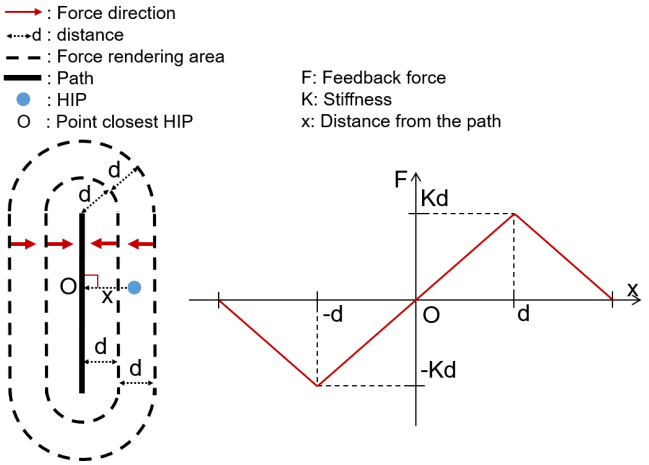

Fig. 1. Force rendering rule for guidance to a line segment.

rendering method based on haptic guidance. This method has two parameters that control the strength and spatial range of guidance, respectively. We investigated their effects on contour tracking performance by a user study (Experiment I) and determined the best values. Next, we carried out another user study (Experiment II) to test the applicability of our method to the identification of geometric primitives. Its results are described in Section III. Lastly, in Section IV, we generalize the basic contour rendering algorithm to handle general cases of multiple contours and apply the generalized algorithm to the rendering of facial features. A user study (Experiment III) was conducted to measure the performance of our algorithm in transmitting the emotional state of facial expressions in portrait photographs.[1]

This work has contributions in 1) the design of 2D contour haptic rendering algorithms that allow for blind exploration and cognition on the basis of active haptic guidance and 2) the optimization and verification of its effectiveness by means of the three user experiments.

II. Generic Guidance-Based Rendering

We began with designing an effective 2D force rendering method that enables a user to track planar contours without visual information. Our method is based on active haptic guidance. We also carried out a user study in order to find the parameter values optimal to contour following accuracy and completion time.

A. Force Rendering Rule

To aid contour following without vision, we render 2D guidance force as follows. For a line segment, we use a force profile illustrated in Fig. 1. We exert attractive force to the line segment, but its magnitude becomes the largest when HIP (haptic interface point; the end point of a force-feedback device representing the finger position) has a distance d from the line. The attractive force is linearly decreased when the HIP moves closer to or farther from the line. This rule provides attractive guidance to the line segment while allowing for easy and comfortable tracking of the line (no guidance force when HIP is exactly on the line). Mathematically, the rule can be written as

$$F = \begin{cases} 0 & \text{if } x < -2d \\ -K(x+2d) & \text{if } -2d \le x < -d \\ Kx & \text{if } 0 \le x < d \\ -K(x-2d) & \text{if } d \le x < 2d \\ 0 & \text{if } 2d \le x \end{cases}, \qquad (1)$$

which has two parameters, stiffness K and guidance range d.

[1]In TouchPhoto [1], our multimodal accessible photography system, the identify of a person is delivered by speech embedded in the photograph. As said earlier, touching is mostly for the recollection of emotional memories.

Fig. 2. Paths used in Experiment I.

B. Methods

The performance of our basic guidance-based rendering method is determined by the stiffness K and the guidance range d. Therefore, we designed a user study, Experiment I, to find the most adequate values for K and d. We used a PHANToM Premium 1.5 (Geomagic, Inc.) in this experiment (also in the next two) along with the CHAI3D 3.0 library. All the user experiments reported in this paper were approved by the IRB of POSTECH (PIRB-2016-E037).

1) Experimental Conditions: With the PHANToM device, we tested five stiffness values (0.1, 0.2, 0.3, 0.4, and 0.5 N/mm) and five guidance ranges (2, 4, 6, 8, and 10 mm).

As 2D contours, we used a circle, an equilateral triangle, and a curved path (Fig. 2). The circle had a radius of 4 cm. The height of the triangle was 5 cm. The curved path had 20 cm of distance between the left and right ends. These paths were rendered on a horizontal virtual wall. We also applied a soft unilateral virtual spring above the virtual wall to gently constrain the stylus movement to the horizontal plane.

The circle and triangle appeared at random locations in a 20-cm $\times$ 20-cm horizontal workspace. Any point on the circle and one of the three triangle vertices could be a starting point for contour following. The location of the curved path was fixed in the horizontal workspace for its large size. Contour following began at either the left or right end point.

Experiment I had 75 conditions by combining 5 stiffness values, 5 guidance ranges, and 3 paths. They were repeated twice. The order of the experimental conditions was randomized for each participant.

The average tracking errors and task completion times are presented for all stiffness values and guidance ranges in Fig. 3. Both of the tracking error and the completion time were generally improved when the stiffness or guidance range was increased. Some saturation behavior was also observed for the largest stiffness values or guidance ranges.

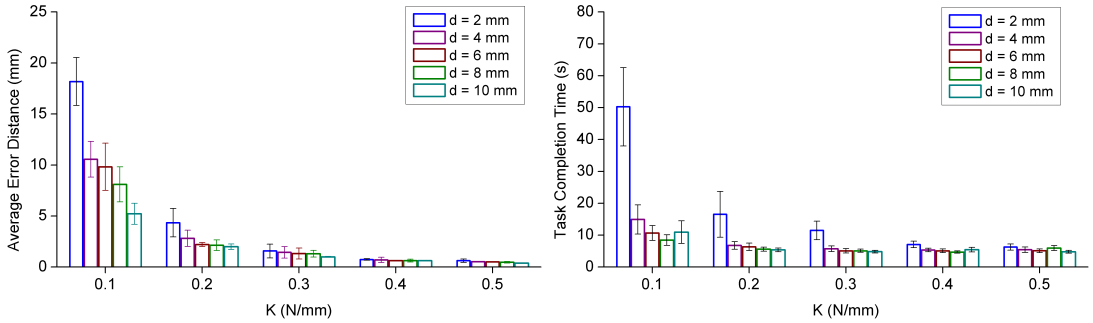

Fig. 3. Average tracking errors (left) and completion times (right).

2) Task and Procedure: Participants first had a training session in which a simple straight line was presented. They learned how to use the PHANToM while experiencing guidance force.

On each main trial, participants' task was as follows. They grasped the PHANToM's stylus with their dominant hands. If they pressed the Enter key on a keyboard, the stylus was moved to a starting point automatically. When ready, they pressed the Enter key again. Then they moved the stylus to follow the *invisible* contour relying on only haptic guidance force. When the stylus reached the end point, the trial was terminated. We also specified a number of knot points that the stylus had to pass. If participants moved the stylus to the end point somehow without passing the knot points, i.e., not following the path, they were asked to repeat that trial.

For the circle and triangle, the start and end points were the same. For the curve path, the end point for tracking was the other end point of the path. Participants were instructed to track the paths as fast and accurately as possible and to close their eyes during tracking. They finished Experiment I in approximately one hour.

3) Data Analysis: We used tracking error and task completion time as performance measures. For the former, we calculated the error distance between the HIP and the path for each sample and then computed their average.

4) Participants: Ten university students (six male and four female; 18–26 years old, average 22.2; one left-handed and nine right-handed) participated in this study. Nobody reported sensorimotor disorders. The participants were paid 10,000 KRW($\simeq$ 9 USD) after the experiment.

C. Results and Discussion

We conducted a three-way repeated-measure ANOVA with the three independent variables of stiffness K, guidance range d, and path P for tracking error. K and d had significant effects ($F(4, 36) = 27.65$, $p < 0.0001$ and $F(4, 36) = 41.97$, $p < 0.0001$), but path did not ($F(2, 18) = 2.59$, $p = 0.102$). The interaction term between K and d was also significant ($F(16, 144) = 6.668$, $p < 0.001$). The same procedure was applied to task completion time. All K, d, and path were significant ($F(4, 36) = 30.14$, $p < 0.0001$, $F(4, 36) = 30.00$, $p < 0.0001$, and $F(2, 18) = 43.01$, $p < 0.0001$), as well as all interactions ($F(16, 144) = 16.4$, $p < 0.0001$ for $K \times d$, $F(8, 72) = 8.698$, $p < 0.0001$ for $K \times Path$, $F(8, 72) = 11.06$, $p < 0.0001$ for $d \times Path$, and $F(32, 288) = 4.293$, $p < 0.0001$ for $K \times d \times Path$). We also ran post-hoc Tukey's HSD tests on the main effects, and results are shown in Fig. 4.

Increasing K or d provides stronger haptic guidance, and it generally improved the participants' performance. However, using large K also makes it difficult to move from one contour to another contour when there exist multiple contours in the scene. Using large d degrades the spatial resolution of presenting multiple contours. These trade-off relationships need to be carefully considered when K and d are determined, and the data in Fig. 4, which shows statistically not-different groups, is pertinent to that purpose. In later experiments, we chose to use $K = 0.2$ N/mm and $d = 4$ mm.

III. Identification of Geometric Primitives

The next step was to apply the guidance-based contour rendering method probed in Experiment I to 2D geometric primitives. We also evaluated users' cognitive ability of identifying geometric primitives in Experiment II.

A. Methods

In what follows, the methods commonly used in Experiment I are not repeated for brevity (also for Experiment III).

1) Stimuli: We used five simple geometric primitives—circle, triangle, square, pentagon, and hexagon. Each primitive had three sizes so that its circumscribed circle had a radius of 2, 5, or 8 cm. They were rendered using the guidance-based rendering algorithm described in Section II with the parameters $K = 0.2$ N/mm and $d = 4$ mm.

2) Task and Procedure: Prior to the experiment, participants were trained shortly (about 1 min) using three simpler forms (straight, broken, and curved lines) to become familiar with haptic guidance. Geometric primitives always appeared at the workspace center. In each trial, participants found a geometric primitive by scanning on the plane using the PHANToM stylus and followed the contour of the primitive to identify its shape, all through only the tactile sense without vision. Then they told the perceived shape, and the experimenter recorded it. Each geometric primitive was tested three times. So each participant completed 45 trials (5 geometric primitives $\times$ 3 sizes $\times$ 3 repetitions), and their presentation order was random. Participants were blindfolded to block any visual cues. Task completion time was also

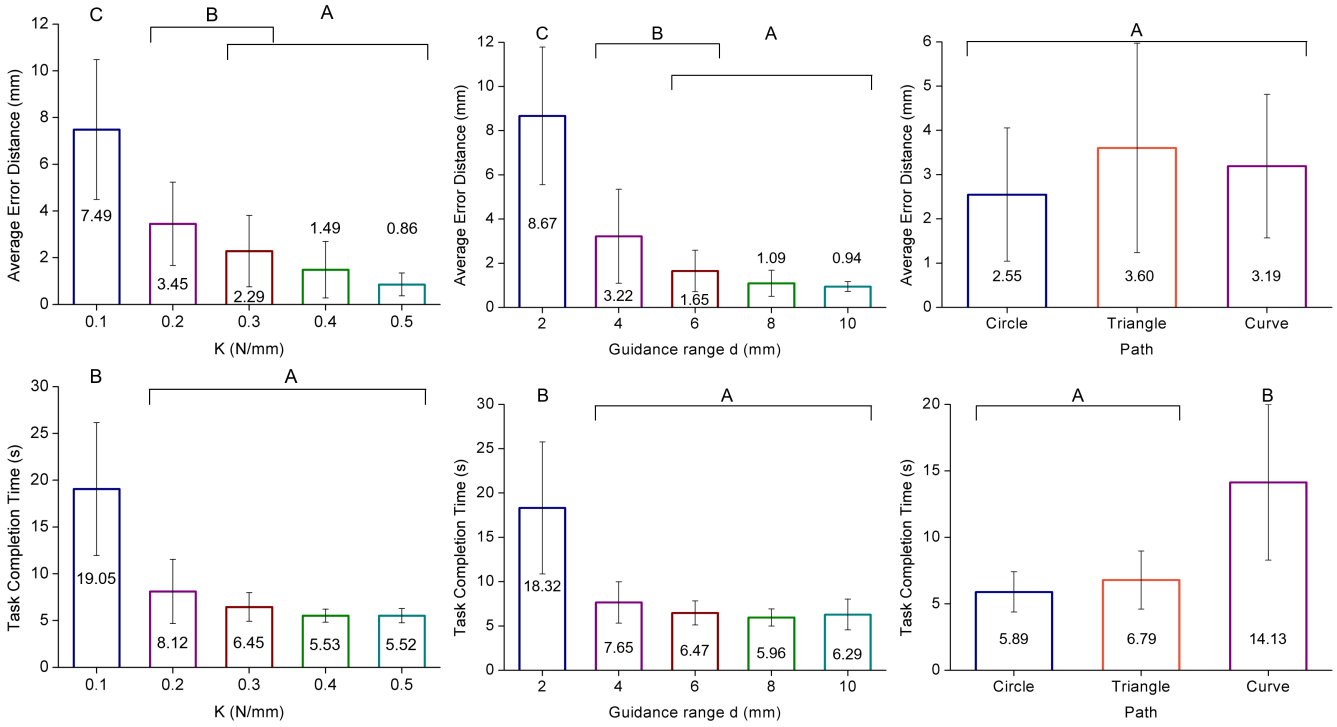

Fig. 4. Mean tracking errors (top) and task completion time (bottom) for stiffness values (left), guidance ranges (middle), and path (right). Error bars represent standard errors. In each plot, factor means grouped by the same letter were not statistically different by Tukey's HSD test.

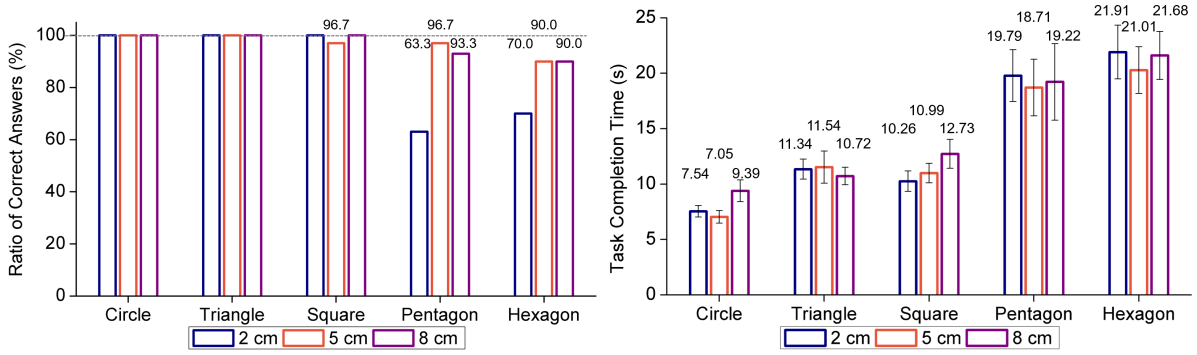

Fig. 5. Results of Experiment II: (left) correct recognition ratio and (right) task completion time. Error bars show standard errors.

recorded in each trial. The entire experiment took 30–40 min to complete per participant.

3) Participants: Ten volunteers (six male and four female; 18–26 years old, average 22.2) without sensorimotor disorders participated in this experiment.

B. Results and Discussion

Experimental results are summarized in Fig. 5 in terms of average correct recognition ratio and task completion time. Geometric primitive identification was nearly perfect for the circles, triangles, and squares of all sizes. It was also easy for pentagons and hexagons when their sizes were 5 and 8 cm, but the correct recognition ratios dropped for the smallest size (2 cm) to 63% and 70%, respectively. Two-way repeated-measure ANOVA with geometric primitive and size as independent factors showed that both factors and their interaction were significant for identification ($F(4, 36) = 3.74$, $p = 0.012$; $F(2, 18) = 27.0$, $p < 0.0001$; $F(8, 72) = 8.59$, $p < 0.0001$). The average task completion time increased from 7.9 s for circle to 21.2 s for hexagon, but it did not depend on size critically. These were confirmed by two-way repeated-measure ANOVA ($F(4, 36) = 16.35$, $p < 0.0001$; $F(2, 18) = 0.578$, $p = 0.571$ for size). Their interaction was not significant ($F(8, 72) = 0.535$, $p = 0.827$).

According to the above results, the participants could identify the 2D geometric primitives with very high accuracy (minimum 90%) if they were sufficiently large (5 cm or more), even the complex ones (pentagon and hexagon). This is a promising finding considering that only very limited sensory cues (mostly kinesthetic) were available, without spatially-distributed tactile cues or visual information. Our results are consistent with those presented by Jansson [10], where

Fig. 6. Faces showing different emotions. Facial features extracted for haptic rendering are shown with blue curves.

participants perceived 3D shapes, such as cone, cylinder, cube, and sphere, rendered by a force-feedback device without vision. Correct recognition was more difficult for small, complex objects (cylinder and cone).

However, it seems that we have some room for further improvement in our guidance-based rendering algorithm in terms of the task completion time. Exploration of the geometric primitives took considerable time, e.g., more than 20 s for hexagons. Following the contour of a real hexagonal object, which provides natural tactile cues at edges, can be done much faster. It is also interesting that the task execution time did not depend on the primitive size. This indicates that the participants somehow modulated the contour scanning velocity depending on the primitive size.

IV. Rendering Algorithm for Facial Features

The results of Experiment II were quite positive in that participants could recognize geometric primitives with high accuracy by force guidance-enabled contouring following only. We proceeded one step further by applying the contour rendering method to the rendering of facial features on photographs. This application is based on our use scenario for multimodal accessible photography described in Section I. Then we carried out another user study (Experiment III) to verify whether our facial feature rendering algorithm allows users to recognize the emotions expressed in the photographs.

A. Algorithm

Input to our algorithm is a photograph showing a person's face. Using an external API (Face++ [14]), we extract the coordinates of eyes, mouth, and facial contour. Although we can obtain position information on other facial features, we observed in pilot experiments that rendering only the contours of eyes and mouth was most effective. Including other facial features tended to make emotion recognition more difficult.

Algorithm 1 Guidance-based force rendering

procedure ForceRendering(HIP, VertexSet Q, K, d)
  $\mathbf{v_1} \leftarrow$ the nearest vertex in Q from HIP
  $\mathbf{v_2} \leftarrow$ the second nearest vertex in Q from HIP
  ▷ Returns the projected point of HIP to the line extending $\mathbf{v_1}$ and $\mathbf{v_2}$
  $\mathbf{v} \leftarrow$ ProjectToLine($\mathbf{v_1}$, $\mathbf{v_2}$, HIP)
  if $\mathbf{v}$ is on the line segment ($\mathbf{v_1}$, $\mathbf{v_2}$) **then**
    if $|\mathbf{v} - \text{HIP}| \leq d$ **then**
      $F \leftarrow |\mathbf{v} - \text{HIP}|$
    else if $|\mathbf{v} - \text{HIP}| \leq 2d$ **then**
      $F \leftarrow 2d - |\mathbf{v} - \text{HIP}|$
    else
      $F \leftarrow 0$
    end if
    $\mathbf{n} = \dfrac{\mathbf{v} - \text{HIP}}{|\mathbf{v} - \text{HIP}|}$
  else
    if $|\mathbf{v_1} - \text{HIP}| \leq d$ **then**
      $F \leftarrow |\mathbf{v_1} - \text{HIP}|$
    else if $|\mathbf{v_1} - \text{HIP}| \leq 2d$ **then**
      $F \leftarrow 2d - |\mathbf{v_1} - \text{HIP}|$
    else
      $F \leftarrow 0$
    end if
    $\mathbf{n} = \dfrac{\mathbf{v_1} - \text{HIP}}{|\mathbf{v_1} - \text{HIP}|}$
  end if
  $\mathbf{F} \leftarrow K \cdot F \cdot \mathbf{n}$
end procedure

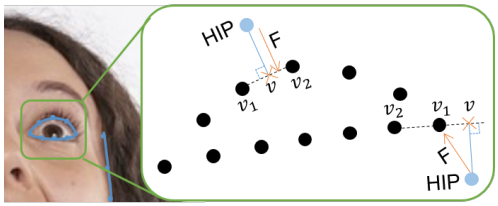

Fig. 7. Two cases considered for haptic guidance.

We also include the facial contour to help users find the face easily on a 2D plane. The contour points founded by Face++ are further smoothed by interpolations using lines and quadratic Bézier curves. See Fig. 6 for example.

Using the facial features, output force is computed by Algorithm 1. We find the point $\mathbf{v_1}$ closest to HIP and the second closest point $\mathbf{v_2}$ among all the vertices representing the contours of facial features. Then we project HIP to the line extending $\mathbf{v_1}$ and $\mathbf{v_2}$ and call it $\mathbf{v}$ (Fig. 7). If $\mathbf{v}$ is between $\mathbf{v_1}$ and $\mathbf{v_2}$, we use the distance between $\mathbf{v}$ and HIP to determine guidance force magnitude. The rule for magnitude computation, which uses the guidance range d, is essentially the same as that of the elementary rendering algorithm described in Section II-A. The force direction is set to be from HIP to $\mathbf{v}$ for attractive guidance. If $\mathbf{v}$ is not on the line segment between $\mathbf{v_1}$ and $\mathbf{v_2}$, the output force is computed using the distance between $\mathbf{v_1}$ and HIP, pointing from HIP to $\mathbf{v_1}$. Finally, the output force is determined by multiplying the stiffness K to the force magnitude.

B. Methods

This experiment was to examine the extent to which our haptic facial feature rendering algorithm transfers the emotion-related information embedded in photographs to users.

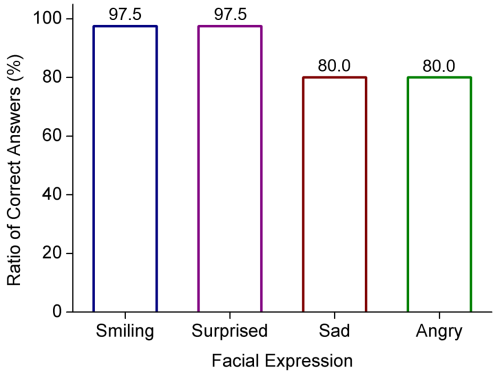

Fig. 8. Emotion recognition ratios.

1) Experimental Conditions: We collected 16 photographs showing frontal faces of all different persons. Each photograph represented one of the four emotional states of smiling, surprised, sad, and angry (4 photographs for each emotion). The emotions were labeled using the automatic emotion recognition function in Face++ [14]. Their facial features were processed as explained earlier for haptic rendering.

2) Task and Procedure: Participants had a training session. They scanned four faces, one each for the four emotional states, rendered using our facial feature rendering algorithm. We explained how we rendered faces to represent emotions and which emotion each face expressed. The training was done in about 1 min for each photograph. The photographs used for training were not tested in the main session.

In main trials, the 16 facial images were presented in random order to participants. Participants perceived the facial features using the PHANToM while freely scanning on the plane and then answered the emotion represented by the facial features. They were not allowed to explore each face for longer than 1 min. They usually finished haptic scanning and said answers in 20-30 s. The experiment took around 30 min. Participants were blindfolded during the experiment.

3) Participants: Ten volunteers (nine male and one female; 24–30 years old, average 25.0) participated in this experiment.

C. Results

Fig. 8 shows average correct recognition ratios in percent for the four emotional states. The recognition ratios were very high (97.5%) for the smiling and surprised faces. They were still high (80%) for the sad and angry faces. The average was 88.75%, indicating that our guidance-based facial feature rendering algorithm has high potential in expressing the emotional states represented in portrait photographs.

Looking at the raw data, most recognition errors occurred between the sad and angry faces. The eye contours were generally different between sad and angry photographs. However, we noticed that the mouth contours were sometimes similar (see the two bottom photographs in Fig. 6). The participants also tended to spend more time in perceiving the mouth area than the eye area, presumably because the mouth area was considerably larger. This seems to have made the confusion between the sad and angry faces worse.

V. Conclusions

In this paper, we have described part of our on-going work about multimodal accessible photography. A focus is on 2D haptic contour rendering algorithm designed to express the facial features in portrait photographs without visual cues. The main idea is an appropriate application of kinesthetic haptic guidance to help users' exploration on the surface. Three user studies are also reported in order to maximize and verify the performance of our algorithm. The findings of this study can contribute to the design and implementation of the assistive technologies that enable similar or "equal-level" interactions between visually-impaired and sighted users [15].

We plan to design a light, portable 2D force-feedback system to be mounted on touchscreen devices and then apply our algorithm for the haptic comprehension of photographs.

References

[1] J. Lim, Y. Yoo, H. Cho, and S. Choi, "TouchPhoto: Enabling independent picture taking and understanding for visually-impaired users," 2019, submitted to IEEE World Haptics Conference 2019.

[2] M. Hachet and A. Kulik, "Elastic control for navigation tasks on pen-based handheld computers," in *Proc. of the IEEE Symposium on 3D User Interfaces*, 2008, pp. 91–96.

[3] T. P. Way and K. E. Barner, "Automatic visual to tactile translation. I. Human factors, access methods and image manipulation," *IEEE Trans. on Rehabilitation Engineering*, vol. 5, no. 1, pp. 81–94, 1997.

[4] J. C. Bliss, M. H. Katcher, C. H. Rogers, and R. P. Shepard, "Optical-to-tactile image conversion for the blind," *IEEE Trans. on Man-Machine Systems*, vol. 11, no. 1, pp. 58–65, 1970.

[5] J. P. Fritz and K. E. Barner, "Design of a haptic data visualization system for people with visual impairments," *IEEE Trans. on Rehabilitation Engineering*, vol. 7, no. 3, pp. 372–384, 1999.

[6] S. Panëels and J. C. Roberts, "Review of designs for haptic data visualization," *IEEE Trans. on Haptics*, vol. 3, no. 2, pp. 119–137, 2010.

[7] M. W. Asghar and K. E. Barner, "Nonlinear multiresolution techniques with applications to scientific visualization in a haptic environment," *IEEE Transactions on Visualization and Computer Graphics*, vol. 7, no. 1, pp. 76–93, 2001.

[8] D. Lareau and J. Lang, "Haptic rendering of photographs," in *Proc. of the IEEE International Workshop on Haptic Audio Visual Environments and Games*, 2012, pp. 107–112.

[9] K. Kahol, J. French, L. Bratton, and S. Panchanathan, "Learning and perceiving colors haptically," in *Proc. of the International ACM SIGACCESS Conference on Computers and Accessibility*, 2006, pp. 173–180.

[10] G. Jansson, "Can a haptic force feedback display provide visually impaired people with useful information about texture roughness and 3d form of virtual objects," in *Proc. of the European Conference on Disability, Virtual Reality and Associated Technologies*, 1998, pp. 105–111.

[11] W. Yu, R. Ramloll, and S. Brewster, "Haptic graphs for blind computer users," in *Haptic Human-Computer Interaction*. Springer, 2001, pp. 41–51.

[12] J. Lee and S. Choi, "Effects of haptic guidance and disturbance on motor learning: Potential advantage of haptic disturbance," in *Proc. of the Haptics Symposium*. IEEE, 2010, pp. 335–342.

[13] H. Lee, G. Han, I. Lee, S. Yim, K. Hong, H. Lee, and S. Choi, "Haptic assistance for memorization of 2D selection sequences," *IEEE Trans. on Human-Machine Systems*, vol. 43, no. 6, pp. 643–649, 2013.

[14] F. Inc., "Face++ cognitive services," 2015. [Online]. Available: https://www.faceplusplus.com/

[15] H. Cho, K. Park, and S. Choi, "Equal-level interaction: A case study for improving user experiences of visually-impaired and sighted people in group activities," in *Proc. of the IEEE International Symposium on Haptic Audio-Visual Environments and Games*, 2018, pp. 45–50.