

DOES IMPRESSION MANIPULATION OF AVATAR FACES AFFECT OBSERVERS' FACE RECOGNITION PERFORMANCE AND DISPLACEMENT OF THEIR GAZE?

Ryoko Yamada[†], Momoko Hada[†], Yuiko Sakuta^{††}, Shigeru Akamatsu[†]

[†] Housei University

3-7-2 Kajino-cho, Koganei-shi, Tokyo 184-8584, Japan

^{††} Jissen Women's University

4-1-1 Osakaue, Hino-shi, Tokyo, 191-8510 Japan

ABSTRACT

We investigated the effects of the appearance transformation of 3D faces for social impression manipulation on the cognitive behaviors of observers represented in terms of face identification performance and eye movement patterns during the recognition phase. We examined whether our participants could identify people whose face's social impression were slightly changed between the encoding and testing phases in a face recognition procedure, and found that facial impression manipulation in semantically negative directions reduced the recognition performance of 3D faces. We also found that eye movements are not affected by the dimension of social impressions but mainly by personal differences in memory strategies.

1. INTRODUCTION

In person-to-person visual communication, faces play an important role as media that convey a variety of cognitive information that involves not only personal identity but also the person's emotional state and social impressions. This also applies to visual communication between humans and synthesized avatar faces based on virtual reality technologies. However, it remains unclear whether the transformation of the appearance of 3D faces for transmitting social impressions affects people's identification of faces. We investigated whether the manipulation of the social impressions transmitted by 3D faces affects observer's cognitive behaviors in terms of face identification performance as well as disparities in eye movement patterns during the recognition phase.

In this experiment, we examined whether people can be identified correctly when their face's social impressions are slightly changed between the encoding and recognition phases and whether the observer's eye movements are affected by the dimensions of social impressions. Such findings are expected to contribute to reference data for designing avatar faces with higher communicative competence for prospective VR communication systems.

2. SYNTHESIZED FACES AS VISUAL STIMULI

2.1. Statistical model for face representatio

By learning from many examples of 3D faces, a morphable 3D face model [1] was introduced and sets of computer-generated faces that transmit various dimensions of social impressions were created by systematically controlling the model's parameters [2].

High-dimensional face representations, such as 3D points that define the shape and the color of a skin surface, are reduced to lower dimensional parameters using PCA, as described in Eq. (1):

$$\mathbf{x}_i = \bar{\mathbf{x}} + \sum_{j=1}^k \mathbf{u}_j \cdot f_{ij} = \bar{\mathbf{x}} + \mathbf{U} \cdot \mathbf{f}_i \quad \cdots(1)$$

Here,

\mathbf{x}_i : high-dimensional vector representing shape/reflectance of the i-th face

\mathbf{f}_i : low-dimensional weight vector (parametric representation of the i-th face)

$\bar{\mathbf{x}}$: average shape/reflectance

$\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]$: matrix with k orthogonal principal components \mathbf{u}_j .

We used the face space model implemented by Facegen Modeller [3], a derivative of Blanz and Vetter's work [1], based on 271 faces laser-scanned in 3D.

2.2. Modeling social impressions of faces

Several works[2][4][5] have built data-driven statistical models linking parameters in the PCA-based face representation model and the social impressions affected by synthesized faces. We used a set of computer-generated faces (Todorov Lab. Face Database[4][6]), which includes three levels (i.e., weak, original, and strong) of impression manipulation for each of three impression dimensions (i.e., trustworthy, dominant, and threatening) as shown in Fig.1.

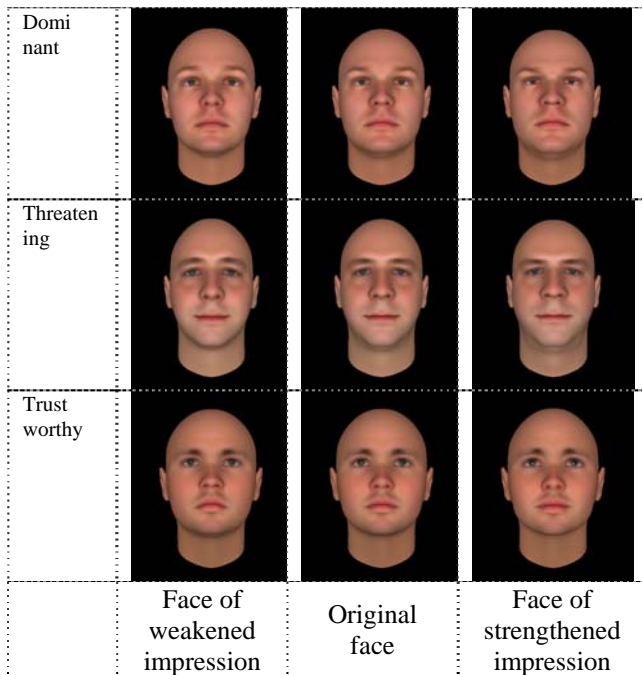


Fig.1 Examples of impression-manipulated faces

3. EXPERIMENTAL METHODOLOGY

3.1. Presentation of visual stimuli

As stimuli, we used the face images prepared in the database described in 2.2. Each stimulus was displayed on a 20-inch monitor with a resolution of 1050×1680 pixels at a 60-cm viewing distance

3.2. Apparatus for eye movement measurement

An EyeLink-CL eye tracker and an Experiment Builder (SR Research products) monitored the participants' eye

movements. The tracking mode was the pupil and corneal reflections at a sampling rate of 500 Hz. A standard nine-point calibration procedure was carried out before the experiment started. A chin rest was used during the experiment.

3.3. Concurrent procedures of encoding/recognition and eye movement measurement

Our experiment consists of encoding and recognition phases in both of which we concurrently measured the eye movements. The experimental procedure for the encoding phase is shown in Fig. 2, where 11 face images (including two dummy stimuli that appear at the test phase's beginning and ending) are presented one by one for seven secs. All of the face images are neutral and no impressions were manipulated: we asked the participants to memorize them.

The experimental procedure for the recognition phase is shown in Fig. 3. In the recognition phase, we conducted the old-new recognition test and made judgments by a 6-point scale. In this test, we presented 18 face images one by one for seven secs each. Three of the images were the same neutral faces presented in the encoding phase, three had weakened impressions, and three had strengthened impressions. The remaining nine were distracters.

3.4. Participants

We recruited 18 Japanese undergraduate and graduate students (ten males and eight females; mean age = 22.1 years old; SD = 4.27) who reported normal or corrected to normal vision.

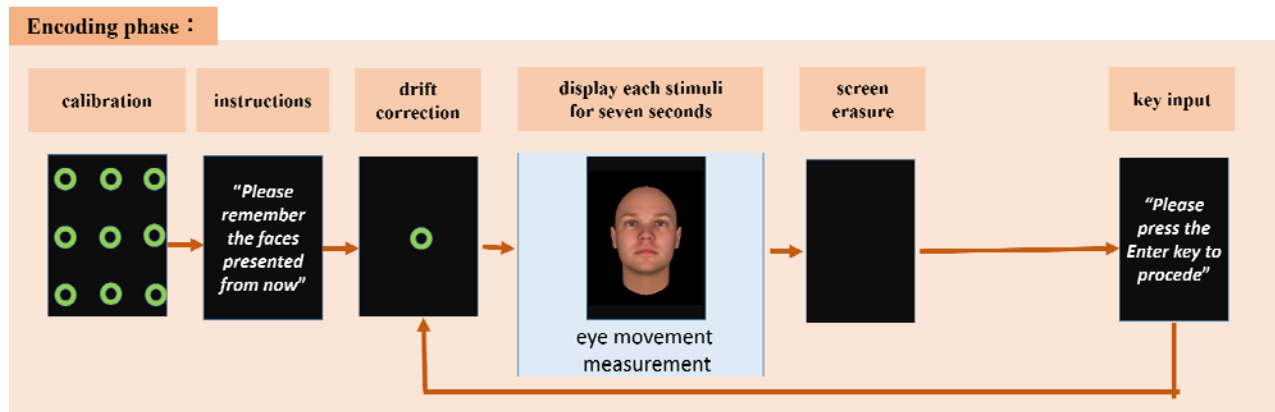


Fig.2 Experimental procedure for encoding phase

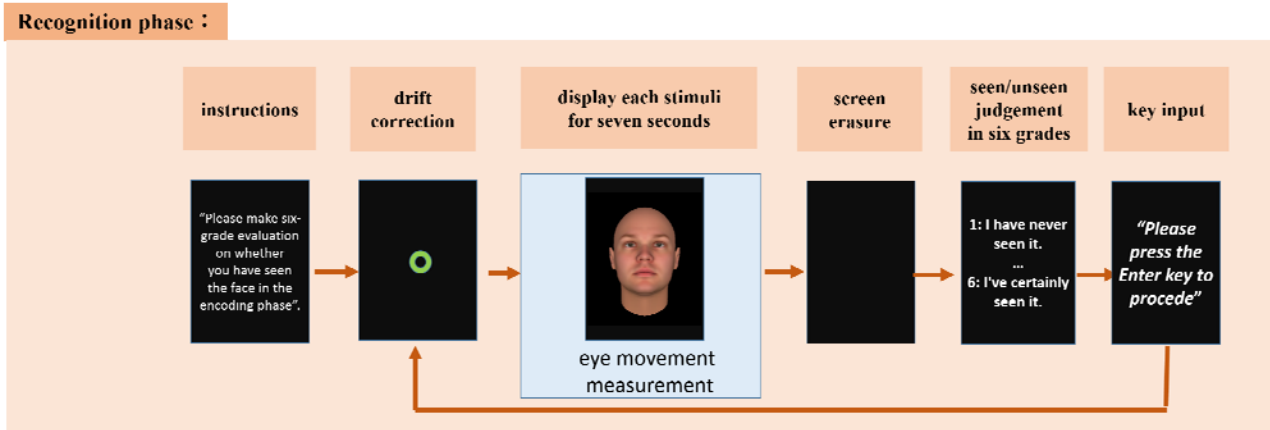


Fig.3 Experimental procedure for recognition phase

4. EXPERIMENTAL RESULTS

4.1. Recognition performance

We categorized three intensity levels of manipulation (i.e., strengthened, without change, and weakened) with social impressions into three semantic directions (i.e. positive, neutral, and negative), as shown in Table 1, depending on the supposed favorability of the observers.

Figure 4 shows the hit rates under the manipulation of the original faces with respect to three types of social impressions in either positive/negative directions. Participants generally failed to properly recognize the encoded faces as previously “seen” ones when the impression manipulation was made in a semantically negative direction (e.g., less trustworthy).

Table 1 Positive/negative categorization of social impression manipulation

Social impressions	Strengthened	Without change	Weakened
Trustworthy	Positive	Neutral	Negative
Dominant	Negative	Neutral	Positive
Threatening	Negative	Neutral	Positive

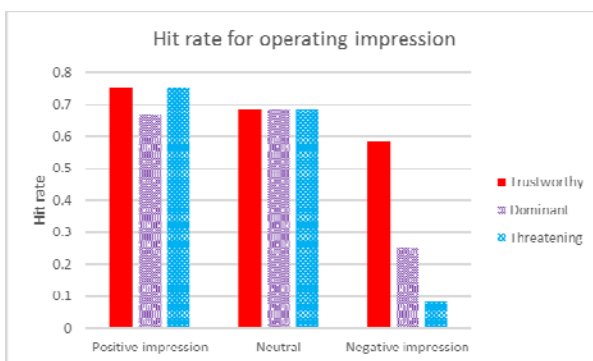


Fig.4 Hit rate for operating impression

4.2. Eye movement pattern: gaze fixation position and duration

We defined the Regions of Interests (ROI) [7] in the facial pattern shown in Fig. 5 and introduced the following numerical features to characterize the gaze fixation position and the eye movement duration.

The following are the evaluation indexes of the gaze fixation in region i :

$$\text{relative area of region } i : S_i = \frac{\text{area of region } i}{\text{total area}}$$

fixation duration rate of region i :

$$T_i = \frac{\text{duration of fixation in region } i}{\text{total fixation area}}$$

$$\text{normalized fixation duration rate of region } i : t_i = \frac{T_i}{S_i}.$$

The graph in Fig.6 illustrates normalized fixation duration rates, measured for each region of interest, when the faces were manipulated with respect to the dominant impressions in either the positive/negative directions and presented as stimuli. In the positive, neutral, and negative impression operations, no significant difference was confirmed in the degree of normalized fixation duration rates.

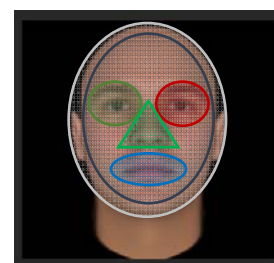


Fig.5 ROIs defined in facial pattern

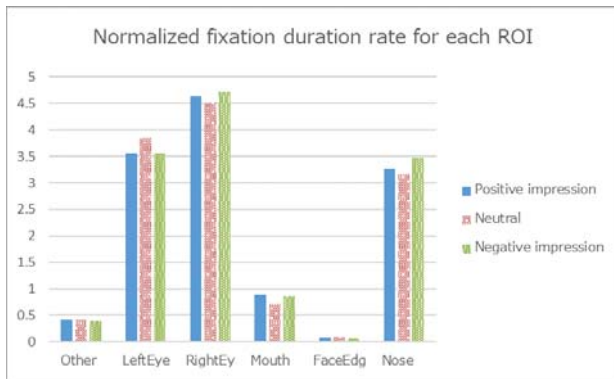


Fig.6 Comparison of ROIs in terms of gaze fixation position and duration

4.3. Eye movement pattern: gaze fixation sequences

Figure 7 shows the eye movement transition patterns measured from subjects who achieved different levels of hit rates in the recognition performance. The average hit rate in the face recognition experiment was 56.8%. Images on left and in middle show the gaze fixation sequences of the participants whose recognition performance scores, 41% and 47%, are lower than the average. Although both subjects have difficulty identifying persons when their facial impressions were manipulated, their eye movement patterns differed greatly. On the other hand, the image on the right shows the gaze fixation sequence of a participant who achieved a recognition performance score of 64%, which exceeds the average. Compared with the middle image, the two participants showed similar eye movement patterns that were greatly concentrated in the center of the face, but their ability to identify people was rather different when impressions of the faces were manipulated.

Analysis of the eye-tracking data suggests that, as with the results obtained in our previous study [7], eye movements are not affected by the dimension of social impressions but mainly by personal differences in memory strategies.

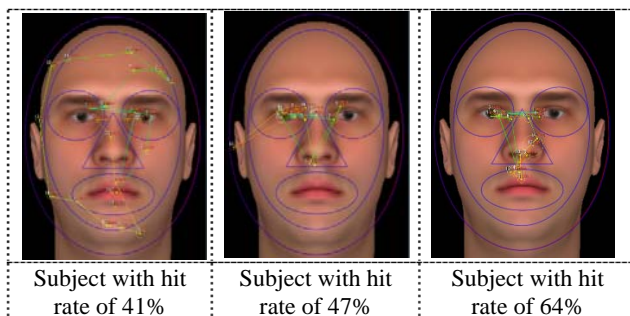


Fig.7 Transition patterns of eye movement measured with different subjects

Observers tended to recognize the faces as encoded ones when these impressions were manipulated in semantically positive directions (i.e., more trustworthy, less threatening), while they didn't recognize them when these impressions were manipulated in semantically negative directions (i.e., less trustworthy, more threatening). We also found that the eye movements of the observers were not affected by the dimension of social impressions but mainly by personal differences in memory strategies.

We believe that some of our findings are beneficial as reference data for designing avatar faces with higher communicative competence for prospective VR communication systems.

ACKNOWLEDGMENT

This work was partly supported by JSPS KAKENHI Grant Number 16K00378.

REFERENCES

- [1] V. Blanz and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," SIGGRAPH'99 Conference Proceedings, pp. 187-194, 1999
- [2] M. Walker and T. Vetter, "Portraits made to measure: Manipulating social judgments about individuals with a statistical face model," Journal of Vision, 9 (11), pp. 1-13, 2009
- [3] Facegen Modeller (<http://facegen.com>)
- [4] N. N. Oosterhof and A. Todorov, "The functional basis of face evaluation," Proceedings of the National Academy of Sciences, 105 (32), pp. 11087-11092, 2008
- [5] Y. Inaba, R. Kobayashi, H. Ishi, J. Gyoba, and S. Akamatsu, "Impression Transformation of 3D Face Based on Morphable 3D Model of Face and Semantic Differential Method," Journal of IIEEJ, vol. 40, no. 1, pp. 96-104, 2011
- [6] A. Todorov and N. N. Oosterhof, "Modeling Social Perception of Faces," IEEE Signal Processing Magazine, pp. 117-122, March 2011
- [7] A. Maruyama, T. Iida, Y. Saitou, Y. Inaba, H. Ishi, J. Gyoba, and S. Akamatsu, "Analysis of eye fixation points in different facial impression judgments," Proc. of IWAIT 2015, PS. 2, 117, Jan. 2015

5. CONCLUSION