FaceShare: Mirroring with Pseudo-Smile Enriches Video Chat Communications

Keita Suzuki[1], Masanori Yokoyama[1, 2], Shigeo Yoshida[1], Takayoshi Mochizuki[2], Tomohiro Yamada[2], Takuji Narumi[1], Tomohiro Tanikawa[1], Michitaka Hirose[1]
[1]Cyber Interface Laboratory, The University of Tokyo, Tokyo, Japan
[2]NTT service evolution laboratories, Kanagawa, Japan
Email: [1]{ksuzuki, yokoyama, shigeodayo, narumi, tani, hirose}@cyber.t.u-tokyo.ac.jp
[2]{yokoyama.masanori, mochizuki.takayoshi, yamada.tomohiro}@lab.ntt.co.jp

ABSTRACT
"Mirroring" refers to the unconscious mimicry of another person's behaviors, such as their facial expressions. Mirroring has many positive effects, such as enhancing closeness and improving the flow of a conversation, which enriches the quality of communication. Our study set out to devise a means of evoking these positive effects in a video chat without any conscious effort of participants. We constructed a videophone system, called *FaceShare*, which can deform the user's face into a smile in response to their partner's smiling. That is, our system generates mirroring by producing a pseudo-smile through image processing. We conducted an experiment in which pairs of participants had brief conversations via *FaceShare*. The results implied that mirroring using the pseudo-smile lets the mimicker, whose face is deformed according to the expressions of their partner, feel a closeness, and improves the flow of the conversation for both the mimicker and the mimickee, who sees the mimicker's deformed face.

Author Keywords
Mirroring; Facial Expression; Telepresence; Transcendent Telepresence; CSCW

ACM Classification Keywords
H.5.4. Group and Organization Interfaces: Computer-supported cooperative work

INTRODUCTION
"Mirroring" is a phenomenon whereby people unconsciously imitate others' limb movements [7], gestures [1], and facial expressions [4]. Mirroring plays a crucial role in facilitating communication by showing empathy, enhancing the feeling of closeness, and encouraging cooperative behavior [1,12,14]. However, mirroring sometimes does not occur, or is weak, depending on the situation and the relationship between the persons involved [9]. Therefore, conversations could be effectively improved by drawing on the positive effects of mirroring by making an effort to display behavior that is congruent with that of one's partner [12]. While doing so currently requires a conscious effort, by using telepresence, the effort can be reduced.

Telepresence refers to technologies that enable communication between remotely located people. Previous studies have focused on the presence and the accuracy of the transmission of information. "Transcendent telepresence", which has been proposed recently [8,13], is based on the idea that it is not always effective to convey information exactly. The ability to transmit information in a modified form is an advantage of telepresence. Transcendent telepresence achieves psychological effects that do not occur or which would be limited in face-to-face communication.

In this study, we focused on facial mirroring as part of our research into transcendent telepresence. We devised a technique to convey the image of a user that has been deformed into a smile to match that of their partner, rather than their original image, in a video chat. We assumed that "mirroring with a pseudo-smile" would evoke the positive effects of actual mirroring more than it tends to occur in natural face-to-face communication.

Our main contribution is proposing the idea of "artificial mirroring through pseudo-smiles" and verifying its effectiveness. Our system will be utilized in both formal and casual situations, for example, to support and strengthen emotional labor in work and to help shorten the mental distance between friends or families who are physically far from each other. We believe that it adds new insights into traditional Computer-Mediated Communication research.

RELATED WORK

Facial Mirroring
When people are (even unconsciously) exposed to happy and angry facial expressions, their emotion-relevant facial muscles react congruently [4]. There have been several discussions of the effects of this facial congruent response. For example, the intentional mimicry of facial expressions makes both the mimickers and mimickees become more effectively attuned to each other, resulting in a smoother interaction and a feeling of closeness [12]. Although previous

CHI 2017, May 06-11, 2017, Denver, CO, USA

DOI: http://dx.doi.org/10.1145/3025453.3025574

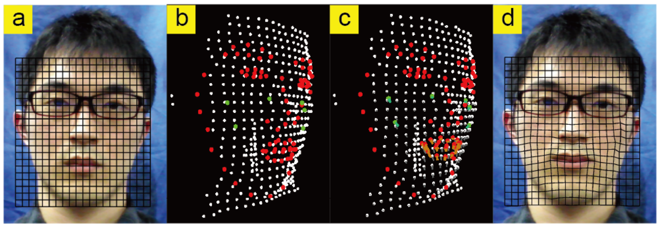

Figure 2. Flow of facial deformation

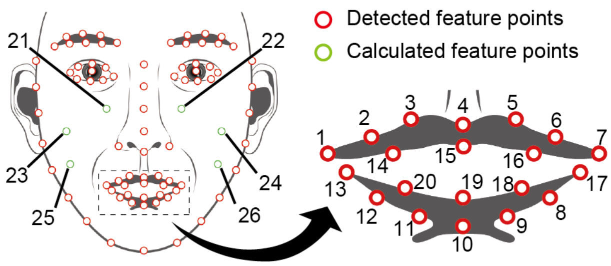

Figure 1. Feature points

experiments required the conscious efforts of the mimickers to mimic their partner, we believe that automatic facial deformation by image processing would enable users to invoke these benefits without any effort.

Ichikawa et al. revealed that the congruent facial response of a receiver maintains the sender's facial expression longer. Moreover, the likeability ratings given by senders were found to be higher for smiling receivers than for those who frowned, regardless of the congruency of the receiver's response. Furthermore, the receiver's congruent smiling expression was rated as being more likeable than an incongruent smiling expression [6]. Therefore, we can assume that a person's congruent response to a smile increases that person's likeability, and causes the partner to smile for a longer duration. It is thought that a longer-duration smile evokes positive emotions because smiling tends to enhance positive emotions even when viewed by a person who is unaware of any emotion [2]. Therefore, in this study, we focused on the smile.

Facial Deformation in Video Chat

Smart Face uses facial deformation in a video conference [10] to enhance a users' creativity during brainstorming sessions by deforming their faces into pseudo-smiles by using image processing. *Smart Face* succeeds in evoking positive effects, which implies that the use of a pseudo-smile would also be effective in our study. However, there are some situations wherein the use of *Smart Face* would be inappropriate because the users are always smiling. We believe that our proposed method could be applied to a wider range of scenarios because our method deforms a user's face congruently with that of their partner. In addition, people's congruent facial expressions tend to be stronger when they see a dynamically changing facial expression, transitioning to a smile, as opposed to viewing a static image of a smiling person [11]. Therefore, we believe that the dynamic

point No.	left	up	rear	point No.	left	up	rear
1	-3.6	5.7	3.7	14	-1.5	3.5	1.8
2	-1.7	3.4	2.6	15	0.0	3.0	1.3
3	-0.5	1.9	1.5	16	1.5	3.5	1.8
4	0.0	1.8	1.5	17	3.2	5.3	3.4
5	0.5	1.9	1.5	18	1.5	3.5	1.5
6	1.7	3.4	2.6	19	0.0	3.0	1.1
7	3.6	5.7	3.7	20	-1.5	3.5	1.5
8	2.5	5.0	2.3	21	-0.6	2.1	0.6
9	1.3	4.5	0.4	22	0.6	2.1	0.6
10	0.0	4.3	-0.2	23	-2.2	3.9	-0.1
11	-1.3	4.5	0.4	24	2.2	3.9	-0.1
12	-2.5	5.0	2.3	25	-3.9	3.7	0.1
13	-3.2	5.3	3.4	26	3.9	3.7	0.1

Table 1. Amount of deformation of each point handle in each direction (mm)

deformation used in our method can evoke stronger positive emotions than those evoked by static deformation.

VIDEOPHONE SYSTEM CAPABLE OF MIRRORING BY GENERATING PSEUDO-SMILES

We developed *FaceShare*, a videophone system, which can deform a user's facial expression into a smile, congruently with that of their partner's smile, in real time. *FaceShare* uses facial deformation and congruent deformation techniques, which are described in detail in this section.

Facial Deformation Technique

The facial expressions are deformed in real time by means of 3D deformation using the feature points of the face including the contours, eyes, nose, and mouth. We used an Intel RealSense Camera (F200) and its SDK to capture an image of the user's face to detect the feature points. We applied Zhu and Gortler's 3D deformation method [15] to the facial deformation. This method interactively changes the positions of the mesh vertexes of 3D objects by manipulating a set of point handles.

First, the 2D image surrounding the user's face is divided into a mesh (Figure 2a). Second, the real-world coordinates of the grid points are calculated from their depths and assumed to be the mesh vertexes. Furthermore, the detected feature points are added as the point handles (Figure 2b). Third, the point handles are moved according to predetermined parameters, and the mesh vertexes are then moved accordingly (Figure 2c). Finally, the moved mesh vertexes are projected onto the original 2D image, and the texture is re-pasted (Figure 2d). Our technique is robust to changes in facial rotation and distance from the camera, because it uses depth information.

We deform a user's face into a smile by lifting the lower eyelids, both ends of the mouth, and the cheeks, in line with the results of the research by Ekman et al. [5]. The positions of the lower eyelids and cheeks are calculated from the positions of the nose, mouth, eyes, and eyebrows because the Intel RealSense SDK could not detect them. Figure 1 shows the detected and calculated feature points. The feature points that are lifted to form a smile are numbered.

A parameter design experiment was conducted to acquire the appropriate amount of deformation for the point handles (feature points) needed to form a natural smile. We asked 13 participants (8 males and 5 females) to mimic an image of a smiling person and measured the amount of movement of each point handle, relative to a neutral face. Table 1 lists the results. The amounts of deformation on the left and right sides of the face were averaged to attain symmetry.

Congruent Deformation Technique

In our system, the deformed user's image is generated and fed to a video stream. Users can video chat on their web browser via WebRTC. Furthermore, the intensity of a user's smile as obtained by Intel RealSense SDK is sent to the browser via WebSocket. The partner's smile intensity can thus be obtained in real time via WebRTC. When *FaceShare* detects that the partner is smiling, it deforms the user's face into a smile after a predetermined stand-by time (400 ms). At this time, the amount of deformation increases linearly taking a predetermined transition time (300 ms) because instantaneous deformation would feel unnatural. When the system detects that the partner is no longer smiling, the amount of deformation decreases linearly taking a transition time (300 ms) after a stand-by time (400 ms). We set the stand-by time based on the fact that the facial muscles react about 300 to 400 ms after a person sees a specific facial expression [3]. We also set the transition time according to the experiment conducted by Ichikawa et al. [6]. The latency is about 170 ms and the frame rate is about 11 fps. A Surface Pro 4 TH4-00014 (CPU: Intel Core i7, GPU: Intel Iris Graphics 540, memory: 16 GB) was used to run the facial deformation and videophone programs. Tests of our system confirmed that *FaceShare* could detect smiles and generating pseudo-smiles correctly even during speech.

EXPERIMENT

We conducted an experiment to investigate whether mirroring with a pseudo-smile generated by *FaceShare* can reproduce the effects of actual facial mirroring. The effects of actual facial mirroring, as described in a related work, are improving the mimickee's impression (especially closeness) of the mimicker, making the mimickee feel that the conversation is progressing more smoothly, and maintaining the mimickee's smile longer.

Six pairs of participants were asked to have conversations via *FaceShare*, after which they completed questionnaires about their impressions of their partner and the conversation. The duration of the smiles of participants during the conversation were measured.

We hypothesized that the closeness, smoothness, and the amount of smiling would be greatest in the case of the mimickees, because the pseudo-smile would reproduce the functions of mirroring. We also hypothesized that these measurements would be second highest for the mimickers, because they would receive feedback to their own pseudo-smile; their pseudo-smile would maintain their partner's smile longer, which would give rise to the feedback.

Conditions

This study was conducted as a within-participant experiment under the following conditions:

- *Normal*: the participant/partner use a normal videophone.
- *Mimicker*: the participant's face is deformed congruently with that of the partner.
- *Mimickee*: the participants are exposed to their partner's congruently deformed face.

We set up three situations according to the conditions to be assigned to the participants (*Normal-Normal*, *Mimicker-Mimickee,* and *Mimickee-Mimicker*), so that all of the participants experience all three conditions. The order in which the different situations were introduced was random.

Procedure

We asked 6 pairs of participants (4 male pairs and 2 female pairs; all pairs were acquainted with each other) to have three 5-minute conversations via *FaceShare*. The experiment was performed in one room, and the participants' positions were not in the same line of sight. Earphones and earmuffs were used to prevent the participants from hearing their partners directly. The themes of each conversation, which were assigned in random order, were: "meals they recently ate," "places they recently visited," and "books or comics they recently read." The participants completed the questionnaires after each conversation. In addition, we asked participants to freely describe what they noticed and cared about during the experiment. We did not reveal the purpose of the study or the mirroring with pseudo-smiles to participants in advance.

Measurements

Closeness: We asked the participants "How much did you enjoy the conversation with your partner?" and "How much did you want to talk more with your partner?" in the questionnaire using a 7-point Likert scale, where 1 was the most negative and 7 was the most positive, to evaluate their closeness.

Smoothness: The participants responded on a 7-point Likert scale to the question "How lively was the conversation?", and the subjective smoothness of the conversation was evaluated. We also measured their utterance amount by measuring the number of frames in which the volume exceeds a certain threshold to provide an objective measurement of the smoothness.

Smile Amount: We measured the participants' average intensity of smile I and smiling rate R during their conversations, using Intel RealSense SDK, to obtain the intensity of the smile s_i in each frame i within a range of 0 to 100. As the duration t_i of each frame is not constant, I and R were calculated as follows:

$$I = \frac{\sum_i s_i t_i}{\sum_i t_i}, \qquad R = \frac{\sum_i sgn(s_i) t_i}{\sum_i t_i},$$

where $sgn(x)$ is the sign function.

Results

Closeness : Figure 3 shows the subjective evaluations of closeness. For the question "How much did you enjoy the conversation with your partner?" a Kruskal-Wallis test showed no significant differences between conditions (p = 0.15). The analysis of “How much did you want to talk more with your partner?” showed a marginally significant difference between conditions (p = 0.056) and Wilcoxon signed-rank post-hoc tests with the Holm method revealed a significant difference: *Mimicker* > *Normal* (p = 0.041).

Smoothness : Figure 4 shows the subjective evaluation of the smoothness and the amount of utterance. A Kruskal-Wallis test for the question “How lively was the conversation?” showed a significant difference between conditions (p = 0.030). Wilcoxon signed-rank post-hoc tests with the Holm method revealed significant differences: *Mimickee* > *Normal* (p = 0.023), *Mimicker* > *Normal* (p = 0.039). A one-factor ANOVA for the utterance amount showed no significant difference between conditions (F(2,22) = 0.29, p = 0.75).

Smile Amount : Figure 5 shows the average intensity of the smile and smiling rate. A one-factor ANOVA for the average intensity of the smile showed a marginally significant difference between conditions (F(2,22) = 2.69, p = 0.090). An analysis of the smiling rate showed no significant differences between conditions (F(2,22) = 1.14, p = 0.34).

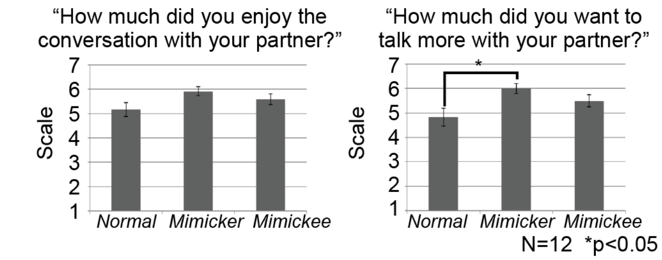

Figure 3. Mean (with SE) ranks of questionnaires about closeness

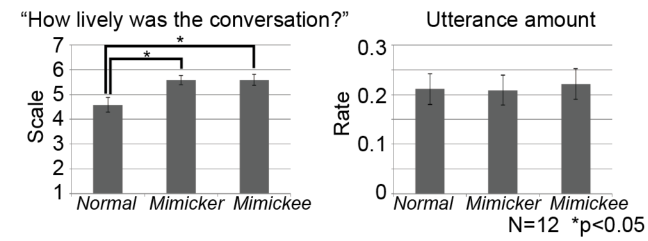

Figure 4. Mean (with SE) ranks of questionnaire about smoothness (left) and mean (with SE) rate of utterance amount (right)

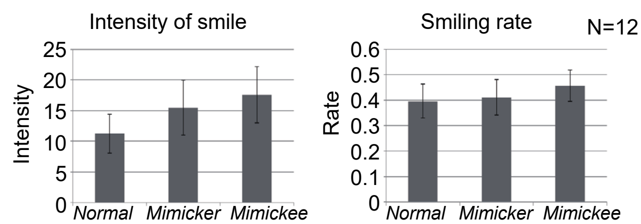

Figure 5. Mean (with SE) intensity of smile (left) and mean (with SE) smiling rate (right)

Discussion

None of the participants in the experiments noticed that their partner’s face was deformed artificially according to the free descriptions. This implies that *FaceShare* does not give the participants the impression of an unnatural image.

Contrary to our expectation that the subjective evaluations about closeness would be highest under the *Mimickee* condition, we found that the *Mimicker* condition produced the highest result. We believe that the mimicker’s mirroring with the pseudo-smile provoked the mimickee’s stronger real-smile, which made mimicker feel “My partner is very interested in my talk!” Furthermore, unconscious recognition of artificial facial deformation might make the effects of mirroring weaker than those of real smiles.

Moreover, we found that mirroring with the pseudo-smile enhances the smoothness of a conversation for both the mimicker and the mimickee. However, there was no significant difference in the amount of utterance. We believe that the improvement in the subjective smoothness was not reflected in the utterance amount because there was hardly any silence during any of the conversations in the trials. Moreover, any increase in the utterance amount under the *Mimickee* and *Mimicker* conditions could offset each other.

Both the intensity of the smile and smiling rate were highest under the *Mimickee* condition and second highest under the *Mimicker* condition, as expected, although there was no statistically significant difference. Furthermore, the difference in the intensity of the smile was greater than that of the smiling rate, which implies that mirroring with a pseudo-smile enhances the intensity of a smile rather than the duration of a smile. We noted that, when laughing, some participants tended to look down or hide their face with their hand. This might result in the smile being impossible to detect. In a follow-up study, we would instruct the participants not to hide their face during the conversations.

CONCLUSION

We developed *FaceShare*, a telepresence system that evokes the positive effects of mirroring by deforming a user’s face into a smile that is congruent to that of their partner’s. The system employs facial deformation and congruent deformation techniques. Our experiments revealed that *FaceShare* enhances the mimicker’s feeling of closeness toward the mimickee and improves the subjective smoothness of a conversation between them. Surprisingly, *FaceShare* was found to have a greater effect on closeness for the mimickers than for the mimickees.

In this study, we focused on smiling rather than frowning, crying, and so on as the target of mirroring. However, further investigation is needed to examine the effects of negative expressions. We are also considering utilizing flows of expressions, conversation contents and voice states to realize more natural mirroring rather than the present “always mimic” approach.

REFERENCES

1. Tanya L Chartrand and John A Bargh. 1999. The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology* 76, 6 (1999), 893. DOI: http://dx.doi.org/10.1037/0022-3514.76.6.893
2. Ulf Dimberg and Sven Söderkvist. 2011. The voluntary facial action technique: A method to test the facial feedback hypothesis. *Journal of nonverbal behavior* 35, 1 (2011), 17–33. DOI: http://dx.doi.org/10.1007/s10919-010-0098-6
3. Ulf Dimberg and Monika Thunberg. 1998. Rapid facial reactions to emotional facial expressions. *Scandinavian journal of psychology* 39, 1 (1998), 39–45. DOI: http://dx.doi.org/10.1111/1467-9450.00054
4. Ulf Dimberg, Monika Thunberg, and Kurt Elmehed. 2000. Unconscious facial reactions to emotional facial expressions. *Psychological science* 11, 1 (2000), 86–89. DOI: http://dx.doi.org/10.1111/1467-9280.00221
5. Paul Ekman and Wallace V Friesen. 2003. *Unmasking the face: A guide to recognizing emotions from facial clues*. Ishk.
6. Hiroko Ichikawa and Junshiro Makino. 2007. Function of congruent facial responses to smiling and frowning. *Perceptual and motor skills* 105, 3 (2007), 838–851. DOI: http://dx.doi.org/10.2466/pms.105.3.838-851
7. James M Kilner, Yves Paulignan, and Sarah-Jayne Blakemore. 2003. An interference effect of observed biological movement on action. *Current Biology* 13, 6 (2003), 522–525. DOI: http://dx.doi.org/10.1016/S0960-9822(03)00165-9
8. Yuki Kinoshita, Masanori Yokoyama, Keita Suzuki, Takayoshi Mochizuki, Tomohiro Yamada, Sho Sakurai, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. 2016. Transcendent Telepresence: Tele-Communication Better Than Face to Face Interaction. In *International Conference on Distributed, Ambient, and Pervasive Interactions*. Springer, 429–438. DOI: http://dx.doi.org/10.1007/978-3-319-39862-4_39
9. Jessica L Lakin, Valerie E Jefferis, Clara Michelle Cheng, and Tanya L Chartrand. 2003. The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of nonverbal behavior* 27, 3 (2003), 145–162. DOI: http://dx.doi.org/10.1023/A:1025389814290
10. Naoto Nakazato, Shigeo Yoshida, Sho Sakurai, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. 2014. Smart Face: enhancing creativity during video conferences using real-time facial deformation. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 75–83. DOI: http://dx.doi.org/10.1145/2531602.2531637
11. Wataru Sato and Sakiko Yoshikawa. 2007. Spontaneous facial mimicry in response to dynamic facial expressions. *Cognition* 104, 1 (2007), 1–18. DOI: http://dx.doi.org/10.1016/j.cognition.2006.05.001
12. Mariëlle Stel and Roos Vonk. 2010. Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction. *British Journal of Psychology* 101, 2 (2010), 311–323. DOI: http://dx.doi.org/10.1348/000712609X465424
13. Keita Suzuki, Masanori Yokoyama, Yuki Kinoshita, Takayoshi Mochizuki, Tomohiro Yamada, and Sho Sakurai. 2016. Enhancing Effect of Mediated Social Touch between Same Gender by Changing Gender Impression. In *Proceedings of the 7th Augmented Human International Conference* 2016. ACM, 17. DOI: http://dx.doi.org/10.1145/2875194.2875239
14. Rick B Van Baaren, Rob W Holland, Kerry Kawakami, and Ad Van Knippenberg. 2004. Mimicry and prosocial behavior. *Psychological science* 15, 1 (2004), 71–74. DOI: http://dx.doi.org/10.1111/j.0963-7214.2004.01501012.x
15. Yuanchen Zhu and Steven J Gortler. 2007. 3D deformation using moving least squares. *Harvard Univ.: TR-10-07, Tech. Rep* (2007).