An Analysis of Self-expressions Based on Synthesized Facial Expression in Japan and the United States

Yuki Nakagawa[1], Kyoko Ito[2, 1], Shogo Nishida[1], Rajesh K. Gupta[3]

[1]Graduate School of Engineering Science, Osaka University, Osaka, Japan
(Tel: +81-6-6850-6382, E-mail:{nakagawa, ito, nishida}@nishilab.sys.es.osaka-u.ac.jp)
[2]Center for the Study of Communication-Design, Osaka University, Osaka, Japan
(Tel: +81-6-6850-6382, E-mail: ito@cscd.osaka-u.ac.jp)
[3]Department of Computer Science and Engineering, University of California, San Diego,
California, United States of America (Tel: +1-858-822-4391, E-mail: gupta@cs.ucsd.edu)

Abstract: Our project focused on clarifying the differences in synthesized facial expressions between monocultural and multicultural societies, and on supporting the communication between both types of cultures. In this research, as a first step, we analyzed and compared the synthesized facial expressions created by the members of a monocultural society and the members of a multicultural society, and clarified their features (similarities and differences) and the future issues that need to be considered in more detail. From the experiment in which Japanese and American students participated, we found that participants from both cultures tend to create more positive facial expressions. We also found that most AU intensities of the American students are stronger than those of the Japanese students and that the SD of most of the AU intensities of American students is bigger than that of the Japanese students. This may mean that it is important for people in a multicultural society to express feelings etc. by nonverbal information such as facial expressions more clearly and exactly than for those in a monocultural society. Based on the results, we decided what issues will need to be considered in more detail in the future.

Keywords: Monocultural and multicultural communication, Synthesized facial expression

1. INTRODUCTION

Face to face communication is important when we want to convey exact information to others, because the face contains a lot of important non-verbal information (e.g., facial expression, hair-style and eye-contact) [1, 2]. Recently communication through e-mail and audio chat i.e., which does not include non-verbal information, is becoming more and more widespread. Especially, facial expression plays an important role in communication, which is conveying our emotions to others [3]. Ekman proved that the facial expressions of 6 basic emotions ("joy", "wonder", "sadness", "anger" and "fear") are expressed universally in all cultures [3]. On the other hand, many researchers have investigated the cultural differences of various facial expressions. For example, Ekman and Frisesen conducted the experiment in order to survey the display rules of facial expressions of Japanese and American people, and showed that Japanese express less negative facial expressions than American people in social situations [3]. The definition of the social situation in this research is that where other people exist with the participant of the experiment.

How can we define "culture"? Although we often use the word, it is a difficult concept to define. With the rapid internationalization, various cultures have started interacting with one another, leading to the appearance of "multicultural societies". We think it is important to take multiculturism into consideration when discussing intercultural communication. This study introduces "monocultural" and "multicultural" societies as two different types of culture. In a monocultural society, since all people in it share similar views and values, their verbal and non-verbal expressions may be homogeneous. People with different cultural backgrounds are mixed in a multicultural society, so we can say that it also contains a mix of verbal and non-verbal expression types.

It is more and more important to consider communication between monocultural and multicultural societies. It is supposed that, as this type of communication is going to increase in face-to-face interaction, some troubles might arise, as mentioned above, because of the misunderstanding of non-verbal information such as facial expressions. In order to communicate with people from different cultures, we are offered many chances to learn the language of the others, but there are few opportunities to learn the appropriate facial expressions of the other culture. Therefore, it is generally difficult to use the correct facial expressions, so as not to be misunderstood by people from other cultures.

In this research, we focused on a facial-expression synthesizing technology for smooth monocultural and multicultural communication. For example, we think it possible in the future for people to use facial-expression synthesizing technology to change their facial expression automatically into one that is

appropriate for the other culture, when communicating through a video call. With the development of informational technology, we can say that the use of facial-expression synthesizing technology will increase. Our project will focus on clarifying the differences in synthesized facial expressions between monocultural and multicultural societies, and on supporting the communication between both types of cultures. In this research, as a first step, we will analyze and compare the synthesized facial expressions created by the members of a monocultural society and the members of a multicultural society, and clarify their features (similarities and differences). Based on this analysis, we can also decide what issues will need to be considered in detail in the future. This research will prove useful from now on when we communicate with people from different cultures.

2. RELATED STUDIES

In this section, we shall introduce several related studies supporting cross-cultural communication.

Fujii et al. developed a chat system, AnnoChat, which can support cross-cultural communication [4]. AnnoChat has two functions, (1) eliminating language barriers by using machine translation and (2) sharing cultural information by annotating words and phrases. These functions were evaluated in Fujii's research. Koda et al. compared the interpretation of avatars' facial expressions among countries [5]. They investigated the cultural differences that appear when interpreting the avatars' facial expressions and proposed guidelines for avatar design universally interpreted within most cultures.

In our present research, we focused on supporting the communication between "monocultural" and "multicultural" societies by using facial expression synthesized technology. This is where our study will likely bring something new and different from the above-mentioned approaches.

3. SYNTHESIZED FACIAL EXPRESSION

3.1 Synthesized facial expressions

The FACS (Facial Action Coding System) is a system which makes it possible to code human facial expressions [6]. FACS defines 44 Action Units (AUs) which represent the facial muscular activity. The Duchenne Smile is, for example, made by using AU6 (Cheek Raiser) and AU12 (Lip Corner Puller) .This means that any facial expression can be created by adjusting the AU intensity.

Using several AUs simultaneously allows us to make naturally-looking synthesized facial expressions with a computer. However, it is sometimes difficult for users to select and adjust the necessary AU. Therefore, in this study, we will use the facial expression creating interface developed by Ito et al. to simplify the procedures for creating the target facial expression [7]. Ito et al. showed that the average time to make the target facial expression

Table 1 The list of AUs used in the interface and their corresponding facial movements

Parts	AU No.	Facial Movements
Eyebrows	1	Inner Brow Raiser
	2	Outer Brow Raiser
	4	Brow Lowerer
Eyes	5	Upper Lid Raiser
	7	Lid Tightener
	41	Lid droop
	42	Slit
	43	Eyes Closed
Cheeks	6	Cheek Raiser
	35	Cheek Suck
Mouth	8	Lips towards Each Other
	10	Upper Lip Raiser
	12	Lip Corner Puller
	15	Lip Corner Depressor
	16	Lower Lip Depressor
	18	Lip Puckerer
	20	Lip stretcher
	22	Lip Funneler
	23	Lip Tightener
	28	Lip Suck
Chin	17	Chin Raiser
	26	Jaw Drop
	27	Mouth Stretch
	29	Jaw Thrust
	30	Jaw Sideways

using this interface is between 3 and 5min, and the questionnaires suggested that many users think it easy to use the interface and create the intended facial expression.

3.2 Interface for creating facial expression

Let us introduce the facial expression creating interface used in this research. This interface is implemented on a PC and is operated with a mouse. The module making the synthesized facial expressions uses FaceFit [8]. A synthesized facial expression is made with 25 kinds of AUs (shown in Table 1). A photograph of the user's face is registered before the experiment. There are 2 steps to create target facial expressions: "Basic setting" (shown in Fig. 1) and "Detailed setting" (shown in Fig. 2). In the Basic setting, the user can select a rough facial expression. On the left, there is the facial expression the user is selecting at that moment. The right circle is used for selecting the facial expression. Any point in the circle can be selected and the facial expressions of 6 basic emotions corresponding to each direction are displayed as labels. After choosing one point in
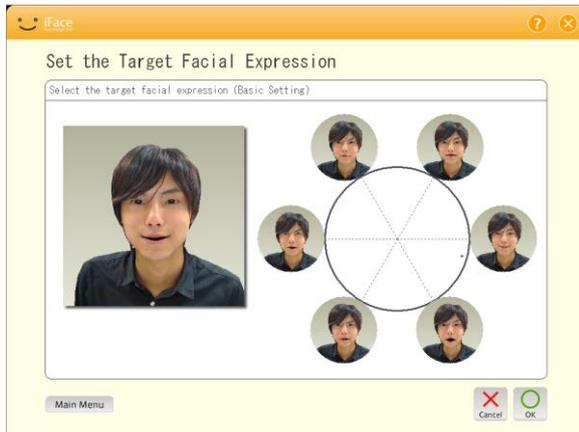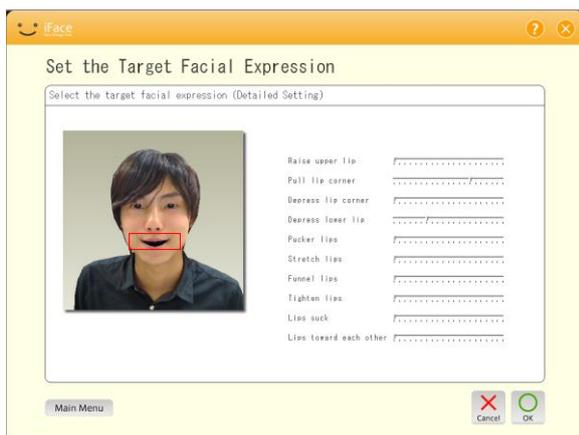
Fig. 1 Interface for Basic Setting

Fig. 2 Interface for Detailed Setting

the circle, the user can go to the Detailed setting. In the Detailed setting, the details of the facial expression in each part of the face can be adjusted. When the cursor is placed on the synthesized facial expression on the left, the parts of the face (eyebrows, eyes, cheeks, mouth and chin) are highlighted and the part that needs adjusting can be selected. By moving the slider on the right, the synthesized facial expression on the left changes. Each slider shows the facial muscular activity, which corresponds to the AU. After the 2 steps mentioned above, the synthesized facial expression is named and saved as a file and the process is completed. After that, we obtain the values of the 25 degrees of AU intensity for each facial expression.

4. EXPERIMENT

4.1 Purpose

The purpose of this experiment is to analyze and compare the synthesized facial expressions created by the members of a monocultural society and the members of a multicultural society, and clarify the similarities and differences. Based on the results, we can also decide the issues to be considered in detail in the future.

4.2 Experimental design

Participants

In this experiment, we considered Japan as a monocultural society. Even though Ainu and foreigners live in Japan, Japan can be defined as a monocultural society because the biggest part of the population is made up by Japanese. As a multicultural society, we choose the United States of America. The U.S. has not adopted multiculturalism officially. However, in the state of California, for example, people can choose one language from several alternatives when getting their driver's license; therefore, we can say that the U.S. tends to encourage multiculturalism. We choose university students as participants from both cultures. This is because in the U.S., it is considered that there are many people with various cultural backgrounds, especially in universities. Based on the above considerations, we will conduct experiments on Japanese college students (Japanese students) and college students living in the U.S. (American students).

Experiment flow

With our experiment, we intend to establish the basis for clarifying the similarities and differences of synthesized facial expressions between members of a monocultural society and members of a multicultural society. It is desirable that we heuristically analyze the features of the facial expressions created by participants without limiting conditions in order to establish what issues will need to be considered in detail in the future. We will not tell the participants to make a specific facial expression in a specific situation, but ask them to freely create a facial expression of their choice. In order to identify what meaning the synthesized facial expression carries, we will also ask them to describe it after interacting with the interface using 35 words expressing various emotions [9].

4.3 Method

Experiment flow

(1) Participants answer questions about themselves (i.e., age, sex, *the country where they have lived the longest and *native language) (5min).

(2) The experimenter explains the participants how to use the facial expression setting interface and then they practice by themselves (4min). During practice, they interact with the synthesized facial expression of the experimenter (author). This interface has 2 editions; Japanese and English. Japanese students use the Japanese edition and American students use the English edition.

(3) The participants create the desired facial expressions starting from their own face pictures using the interface. There is no time limit, so they can work for as long as they want.

Table 3 Creation rate of facial expressions including each word for emotion

Categories	Words for emotion	Creation rate	
		Japanese	American
positive	joy*	27%	34%
	pleasure	47%	21%
	happiness	27%	38%
	love	20%	10%
	satisfaction	13%	17%
	warmness	30%	21%
	longing	3.3%	3.4%
	respect	0%	6.9%
	tenderness	40%	6.9%
	curiosity	6.7%	31%
neutral	doubt	3.3%	14%
	pride	6.7%	3.4%
	patience	3.3%	3.4%
	vacant	3.3%	0%
	pity	0%	0%
	shame	3.3	0%
	wonder*	3.3	24%
negative	ennui	3.3%	0%
	jealously	0%	0%
	repentance	0%	0%
	intolerance	3.3%	6.9%
	unpleasantness	6.7%	6.9%
	hesitation	0%	3.4%
	anxiety	0%	10%
	sadness*	0%	3.4%
	loneliness	0%	3.4%
	painfulness	0%	0%
	dislike*	3.3%	14%
	disappointment	0%	10%
	hate	0%	0%
	humiliation	0%	3.4%
	indecent	0%	0%
	anger*	6.7%	3.4%
	contempt	0%	3.4%
	fear*	0%	3.4%

* Basic emotions

(4) The participants select from the 35 words available to describe the expression they created (multiple answers possible).
*Only for American students

The experiment is conducted in a specially prepared room, and the experimenter is present all through the experiment. The experimenter takes pictures of the participants before the experiment.

Method of analysis
・The words for emotion describing the synthesized facial expressions:
We categorize them into "positive", "neutral" and "negative" and investigate what kind of facial expressions the participants tend to create.

・The AU intensities of the synthesized facial expressions:
We compare the AU intensities between the two cultures, and analyze the characteristics of the intensity and the standard deviation (SD) of each AU.

We analyze the various synthesized facial expressions obtained, by focusing on Ekman's facial expressions of the 6 basic emotions in order to establish basic knowledge about the cultural differences.

4.4 Results
Participants
An experiment using the interface was conducted on 30 Japanese students (19 men and 11 women, with an average age of 22.4 years) and 29 American students (15 men and 14 women, with an average age of 23.8 years). 19 American students answered "America" to the question, "What is the country where you have lived the longest". Therefore 30 Japanese students (19 men and 11 women, with an average age of 22.4 years) and 19 American students (9 men and 10 women, with an average age of 22.5 years) were chosen for analysis in this research. The experiment for the Japanese students was conducted at Osaka University; they were all students of Osaka University. The experiment for the American students was conducted at the University of California, San Diego and the participants were all students of the University of California, San Diego.

Synthesized facial expressions created by the participants
Table 3 shows the creation rate of each type of facial expression (including the word used to describe it). It shows that many participants tend to create positive facial expressions in both cultures. We can say that there are few differences about the creation rate of positive and neutral facial expressions in both cultures. On the other hand, as far as negative facial expressions are concerned, many Japanese students hardly created those facial expressions, while the American students created various types of negative facial expressions. The facial expression of "joy", which is one of the facial expressions of the 6 basic emotions, was created by many students from both cultures. However, the other basic facial expressions (especially, "sadness" and "fear") were created by very few the Japanese students.

Facial expression of "joy"
In this subsection, we analyze the facial expression of "joy" in more detail. Fig. 3 shows the AU intensities of the facial expressions that include "joy" for 8 Japanese students (5 men and 3 women) and 8 American students (3 men and 5 women). Ekman stated that the "joy" facial expression is formed using AU6 and AU12 [6]. We found that the intensity of
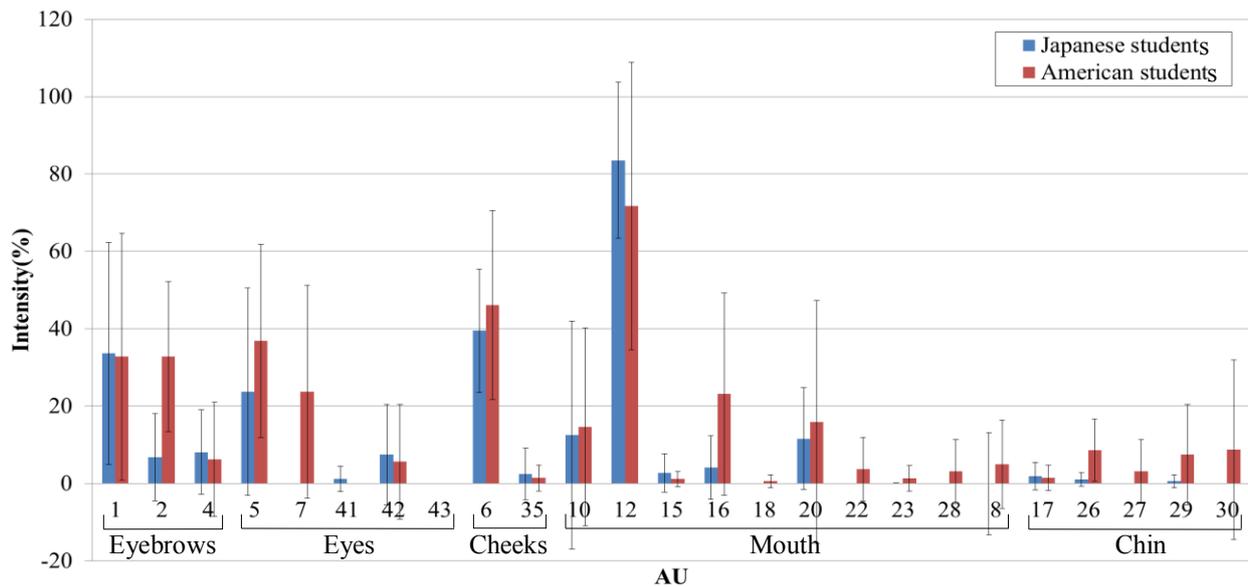
Fig. 3 AU intensities of "joy" facial expression

AU12 is the strongest and that of AU6 is the second strongest of all. We also noticed that most AU intensities of the American students are greater than those of the Japanese students and, what is more, the SD of most of the AU intensities of American students (18 AUs in 25) is bigger than that of the Japanese students.

Other facial expressions ("wonder", "dislike" and "anger")

As far as other facial expressions are concerned, there was only 1 Japanese student and 2 American students who created a facial expression and described it using, among other words, the word "wonder"; 1 Japanese student and 4 American students created facial expression describing it with the word (among others) "dislike"; and 2 Japanese students and 1 American student created a facial expression for which they chose "anger" as one of the describing words. Therefore, because of the small amount of data, it is difficult to find the differences between the two cultures with regard to the facial expressions for "wonder", "dislike" and "anger".

4.5 Discussion

Synthesized facial expressions created by the participants

This experiment was conducted immediately after the author met each participant, in the room prepared by the author and in the attendance of the author. Therefore, we can say that this experiment was performed in a social situation, which is mentioned in chapter 1. The display rules of facial expression, stating that, Japanese people are apt to express less negative facial expressions than American people in social situations are described in Ekman, as previous mentioned [3]. In this experiment, most Japanese

students created positive facial expressions and fewer negative facial expressions than the American students. Therefore, it is safe to say that previous knowledge about the display rules of expression may be adapted to the synthesized facial expressions, too.

Facial expression of "joy"

We found that most AU intensities of the American students are stronger than those of the Japanese students and, what is more, the SD of most of the AU intensities of American students (18 AUs in 25) is bigger than that of the Japanese students. It is considered that in a monocultural society, people in it do not have to express feelings etc. by nonverbal information such as facial expressions very clearly and exactly. This may be because, people in a monocultural society share non-verbal information well, and are easily able to understand the meaning of the facial expressions of other people in the same society. On the other hand, in a multicultural society, since people with different cultural backgrounds are mixed in it, people likely tend to depend on non-verbal information more heavily than for those in a monocultural society. Therefore, we can say that it is important for people in a multicultural society to express facial expressions clearly and exactly in order to what they want to say to others.

4.6 Subjects of future investigation

Experiment instructions

In this experiment, the participants created facial expressions freely by using photographs of their own faces. However, we were not able to obtain sufficient data about neutral and negative facial expressions. When people misunderstand each other's negative facial expressions, this may give rise to serious social problems. Therefore, negative facial expressions are considered to be more important than positive facial

expressions (Decoding rules) [10]. Thus, it is desirable that not only positive facial expressions such as "joy", but also neutral or negative facial expressions should be revised by using facial-expressions synthesizing technology. We also need to consider the diversity of facial expressions in a multicultural society. Therefore, we plant to ask the participants to create a certain facial expression, in a certain context (e.g., "Please show you are "angry" to a child").

Psychological burden of the experiment with a synthesized facial expression

When we told the participants the purpose of our experiment and exactly what they were expected to do it, a few of them, from both cultures declined to participate. In the communication using the synthesized facial expressions, the psychological burden of the user might be big. In order to reduce it, we need to investigate this aspect in more detail in the future.

5. CONCLUSION

In this paper, we compared the synthesized facial expressions created among members of a monocultural society and members of a multicultural society, and clarified their characteristics and the future issues that need to be considered in more detail. From the experiment in which Japanese and American students participated, we found that participants from both cultures tend to create more positive facial expressions, and thus we conducted that the previous knowledge about display rules of expression may be adapted to the synthesized facial expressions, too. We also found that most AU intensities of the American students are stronger than those of the Japanese students and that the SD of most of the AU intensities of American students is bigger than that of the Japanese students. This may mean that it is important for people in a multicultural society to express feelings etc. by nonverbal information such as facial expressions more clearly and exactly than for those in a monocultural society.

As the next stage of our research, we plan to conduct experiments where the participants will be asked to create a specific facial expression; also, we intend to investigate methods for reducing the psychological burden on the user when communicating via synthesized facial expressions. We are confident that this research will prove useful in the future when we communicate with people from different cultures.

REFERENCES

[1] A. Mehrabian, Silent messages, Implicit Communication of Emotions and Attitudes, 2nd Ed., Wadsworth Pub. Co., 1981.

[2] M. F. Vargas, Louder Than Words: An Introduction to Nonverbal Communication, Iowa State Pr, 1986.

[3] P. Ekman and W. V. Friesen, "Unmasking the Face," Englewood Cliffs, N. J. : Prentice Hall, 1975.

[4] K. Fujii, T. Shigenobu and T. Yoshino, "Evaluation of Annotation in Intercultural Communication Chat Tool Using Machine Translation," Journal of Information Processing, Vol. 48, No. 1, pp. 63-71, 2007(in Japanese).

[5] T. Koda and T. Ishida, "Cross-cultural Comparison of Interpretation of Avatar's Facial Expression," Journal of Information Processing, Vol. 47, No. 3, pp.731-738, 2007(in Japanese).

[6] P. Ekman and W. V. Friesen, The Facial Action Coding System, Consulting Psychologists Press, 1978.

[7] K. Ito, H. Kurose, A. Takami and S. Nishida, iFace: Facial Expression Training System, Affective Computing Chapter 17, Austria, 2008.

[8] Galatea Project : http://hil.t.u-tokyo.ac.jp/~galatea/index-jp.html (in May, 2012).

[9] M. Yoshida, R. Kinase, J. Kurokawa and S. Yoshikawa, "Multi-dimensional scaling of emotion," Japanese Psychological Research, Vol.12, No. 2, pp.45-61, 1970.

[10] H. A. Elfenbein, and N. A. Ambady, "Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition," Journal of Cross-Cultural Psychology, Vol.34, pp.92-110, 2003.