

## Cross-cultural Study of Avatar Expression Interpretations

Tomoko Koda and Toru Ishida

*Department of Social Informatics, Kyoto University  
Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501 JAPAN  
koda@lab7.kuis.kyoto-u.ac.jp, ishida@i.kyoto-u.ac.jp*

### Abstract

*Avatars are increasingly used to express our emotions in our online communications. Such avatars are based on the assumption that avatar expressions are interpreted universally among all cultures. This study aims to elucidate the following two issues: 1) Identifying cultural differences in interpreting avatars' facial expressions. This is done by applying psychological findings on cultural differences in human facial expression recognition to the case of avatar expressions. 2) Identifying avatar facial expressions that are recognized differently across cultures. We conducted an open web experiment to gather users' interpretations of various avatar facial expressions from eight countries within Asia, North and South America, and Europe. The results showed: 1) Cultural differences do exist in interpreting avatar facial expressions, which confirms the psychological findings that physical proximity affects recognition accuracy. Japan had the highest recognition accuracy for avatar expressions designed by Japanese designers, followed by Korea. 2) There are wide differences among cultures in interpreting positive expressions, while negative expressions had higher recognition accuracy regardless of culture.*

### 1. Introduction

Since instant messenger and chat services are frequently used in our daily communication beyond nationality and languages, emoticons and expressive avatars are widely used to provide nonverbal cues to text-only messages [1, 2, 3, 4]. Studies on emoticons and avatars report positive effects on computer-mediated communication. Those studies indicate that emoticons and avatars improve user experiences and interactions among participants [5, 6, 7] and build enthusiasm toward participation and friendliness in intercultural communication [8, 9].

However, these avatars are used based on an implicit assumption that avatar expressions are interpreted universally across cultures. Since avatars work as graphical representations of our underlying emotions in online communication, those expressions should be carefully designed so that they are recognized universally. We need to closely examine cultural differences in the interpretation of expressive avatars to avoid misunderstandings in using them.

However, few studies have compared the cultural differences in interpreting avatars. One of those studies compared interpretations of avatars' animated gestures between the Netherlands and Japan [10]. Their results showed that there are cultural differences in perceived valance in animated characters between the two countries. Japanese women perceived stronger emotions in some animated gestures of an avatar, i.e., bowing, than the Dutch subjects, although there were no overall differences in interpretation of the presented gestures. In our previous study, we conducted a cross-cultural experiment in the form of a series of discussions on a multilingual BBS with expressive avatars between China and Japan [8]. The results show some facial expressions used in the experiment were interpreted completely differently and used for different purposes between Chinese and Japanese. Those "misinterpreted" expressions are "sweat-on-the-face," "wide-eyed," and "closed-eyes." For example, the "wide-eyed" expression was interpreted as "surprised" by the Japanese subjects, while the Chinese subjects interpreted it as "intelligent" and used it when presenting a novel idea or asking questions. We observed that the Japanese subjects tried to confirm the meaning of the Chinese subject's message with the "wide-eyed" expression. This is one example of communication gaps caused by different interpretations of avatar expressions between the two countries.

The above two studies were each conducted between only two countries. We need to conduct an evaluation experiment among multiple countries in

order to investigate cultural differences in avatar expression interpretation and what kinds of expressions are interpreted universally and what kinds are not. We believe the results would serve as a design guideline for universal avatar expression that would not lead to miscommunication.

In this study, we apply findings from psychological studies on human facial expressions, since there has been a much wider variety of studies in psychology on human expressions than on avatar expressions. The most widely accepted findings come from the work of Ekman. He states that seven emotions, namely, anger, fear, disgust, surprise, sadness, happiness and contempt, are universally expressed by all cultures [11]. However, he also argues the implications and connotations of those facial expressions are culturally dependent, and the degree of allowance in showing or perceiving those expressions socially differs across cultures [12]. Recent psychological research found evidence for an “in-group advantage” in emotion recognition. That is, recognition accuracy is higher for emotions both expressed and recognized by members of the same cultural group. Elfenbein et al. state, “This in-group advantage, defined as extent to which emotions are recognized less accurately across cultural boundaries, was smaller for cultural groups with greater exposure to one another, for example with greater physical proximity to each other [13].” Also, the decoding rule implies that we concentrate on recognition of negative expressions, since misinterpretation of negative expressions leads to more serious social problems than misinterpretation of positive expressions would cause [14].

This study investigates the following two issues: 1) Verifying cultural differences in interpreting avatars’ facial expression; this is done by applying the above psychological findings on cultural differences in human facial expression recognition to the case of avatar expressions. 2) Identifying avatar facial expressions that are recognized differently across cultures.

We conducted an open web experiment to investigate the above research issues by comparing interpretations of avatar expressions from multiple countries. We expect the result to serve as an avatar design guideline for online communication tools.

## **2. Experiment overview**

### **2.1. Experimental procedure**

The experiment’s web site is open to the public [15]. People from all over the world can access this web site

and freely participate in the experiment. Participation is voluntarily.

The experiment itself was developed using the application of Macromedia Flash. Subjects first answer a brief questionnaire on their background profile. The main experiment starts after the questionnaire, which is presented as a matching puzzle game (Figure 1). Subjects are requested to match 12 facial expressions to 12 adjectives. The 12 facial expressions are displayed in a 4 x 3 matrix and the 12 adjectives as buttons below the matrix. As shown in Figure 1, subjects can drag/drop the adjective buttons to/on the 12 expressions and continue changing the location of each button until they are satisfied with their answer. One avatar representation is chosen randomly from 40 avatars, and facial expression images are randomly placed in the 4 x 3 matrix. The adjective buttons are always displayed in the same order, and the 12 adjectives are always the same (see sec. 2.3 for the adjectives used in the experiment).

Subjects’ answers to the puzzle game and questionnaire, as well as their background profile including gender, age, county of origin, and native language, are logged in the server for later analysis. Subjects can continue the experiment with another set of avatars until they finish evaluating all 40 avatar designs or can stop at any time. Each avatar design is displayed only once to the same subject.

The adjectives can be shown in English, French, German, Italian, Spanish, Chinese, Korean, and Japanese (all validated by native speakers). Subjects from countries where the above languages are primarily spoken can see the adjective selections in their native language according to the background profile. The default language is set to English.

### **2.2. Avatar design**

Commercially used avatars are represented not by photo-realistic images but as caricatures or comic figures. We prepared 40 avatar representations drawn by three Japanese designers using Japanese comic/anime drawing style. By using avatars drawn with techniques from one culture, we can use those avatars as “expressers” and subjects as “recognizers” as in [13]. Accordingly, comparing the answers between Japanese users and those of other countries made it easier to validate the in-group advantage.

Avatars are categorized into five groups, namely, human figures, animals, plants, objects, and imaginary figures (culture-dependant). Figure 2 shows examples from the 40 avatar representations.

### 2.3. Facial expression design

The 12 expressions used in the experiment are “happy,” “sad,” “approving,” “disapproving,” “proud,” “ashamed,” “grateful,” “angry,” “impressed,” “confused,” “remorseful,” and “surprised” as shown in Figure 3. Those expressions are selected from Ortony, Clore and Collins’ global structure of emotion types, known as the OCC model [16]. These are commonly used expressions in chat and instant messenger systems [1, 2, 3], and they reflect those emotions desired by the subjects for intercultural communication in [8].

These 12 expressions are paired as valanced expressions as defined in the OCC model, that is, negative/positive emotions that arise in reacting to an event or person. “Happy,” “approving,” “proud,” “grateful,” and “impressed” are positive expressions, while “sad,” “disapproving,” “ashamed,” “angry,” “confused,” and “remorseful” are negative expressions, leaving “surprised” as a neutral expression.



Subjects can drag/drop the adjective buttons to the matching facial expressions.

**Figure 1. Experiment screen: Matching puzzle game between facial expressions and adjectives**



**Figure 2. Examples of avatar representations**



From top left, happy, sad, approving, disapproving, proud, ashamed, grateful, angry, impressed, confused, remorseful, and surprised, drawn in Japanese comic style.

**Figure 3. Twelve facial expressions using one of the avatars**

### 3. Results

#### 3.1. Subjects and participating countries

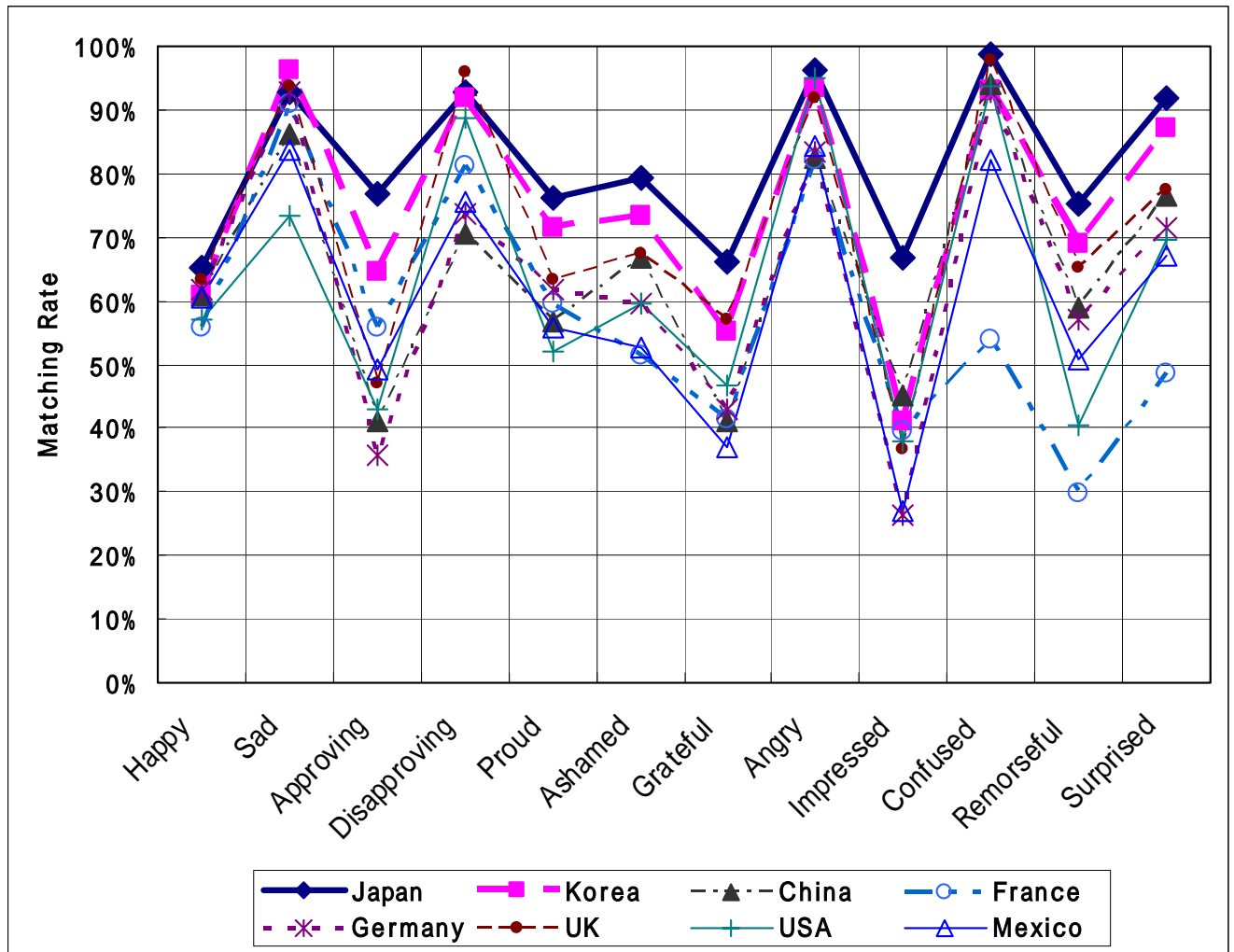
We have had 1,240 participants from 31 countries. Subjects’ gender ratio is roughly male:female = 1:1 (676 male subjects and 561 female). Subjects’ age ranges are: 6% are in their 10s, 43% in their 20s, 35% are in their 30s, 12% are in their 40s, and 4% in their 50s.

We have analyzed answers from eight countries having more than 40 participants, namely, Japan (n=310), South Korea (n=322), China (n=50), France (n=111), Germany (n=62), United Kingdom (n=49), United States (n=75), and Mexico (n=149). The subjects from those eight countries saw the adjectives in their mother tongue. We used answers only in the cases where the subject’s native language and the official language of his/her country matched.

#### 3.2. Overall cultural differences in interpreting expressions

Subjects’ answers to the puzzle game are analyzed by calculating matching rates between expressions and adjectives. There is no correct answer to the matching puzzle, but the avatar designers’ original intention can be used as an expresser’s “standard” answer. Each expression and adjective is assigned a number (1-12) within the system. The designer’s intended pairs are described as (1,1), (2,2), (3,3), (4,4) reflecting (expression number, adjective number). We calculated each country’s number of “expression-adjective” pairs that are the same as the designers’ pairs. Consequently, here, “matching rate” means the percentage of pairs of expressions and adjectives that match the avatar designer’s intentional pairs. For example, the matching rate of answer pairs (1,5), (2,1), (3,3), (4,9) is 25%.

The matching rate for each facial expression by country is shown in Figure 4 and Table 1. The



Matching rate means the percentage of pairs of expressions and adjectives that match the avatar designer's intentional pairs. Numbers of answers by each country are: Japan: n=310, Korea: n=322, China: n=50, UK: n=49, France: n=111, Germany: n=62, USA: n=75, Mexico: n=149.

**Figure 4. Matching rate of each expression by country**

**Table 1. Matching rate of each expression by country**

Adjectives with grey background are positive expressions.

	Happy	Sad	Approving	Disapproving	Proud	Ashamed	Grateful	Angry	Impressed	Confused	Remorseful	Surprised
<b>Japan</b>	65%	93%	77%	93%	76%	79%	66%	96%	67%	99%	75%	92%
<b>Korea</b>	61%	96%	65%	92%	71%	73%	55%	93%	41%	93%	69%	87%
China	61%	86%	41%	71%	57%	67%	41%	82%	45%	94%	59%	76%
France	56%	91%	56%	81%	59%	51%	41%	82%	40%	54%	30%	49%
Germany	62%	93%	36%	74%	62%	60%	43%	83%	26%	93%	57%	71%
UK	63%	94%	47%	96%	63%	67%	57%	92%	37%	98%	65%	78%
USA	57%	73%	43%	89%	52%	59%	47%	95%	38%	94%	41%	70%
Mexico	61%	84%	49%	76%	56%	53%	37%	84%	27%	82%	51%	67%

matching rate of Japanese is significantly higher for all expressions except “sad” and “disapproving” (by chi-squared test and Scheffe’s method of multiple comparison,  $p < 0.01$ ), followed by Korea. Nevertheless, Japan maintains high matching rates for “sad” and “disapproving” expressions. There are no significant cultural differences in these matching rates among the countries other than Japan and Korea.

As stated in 2.2, avatars are designed by Japanese designers using Japanese comic/anime drawing techniques. Thus we can regard the designers as expressers and the subjects as recognizers. Japanese subjects’ recognition accuracy of the avatar expressions is significantly higher than that of other countries, while Korean subjects’ accuracy is the second highest. This verifies that there is an in-group advantage within the same country (within Japan) and one between neighboring countries (Japan and Korea).

### 3.3. Differences between positive/negative expressions

When we focus on the matching rate of each expression, the result shows that positive expressions in valanced expression pairs (happy-sad, approving-disapproving, and grateful-angry) have lower matching rates than the negative expressions in the same pair. Negative expressions (sad, disapproving, angry, and confused) have significantly higher matching rates regardless of country (by analysis of variance and Scheffe’s method of multiple comparison,  $p < 0.01$ ), while positive expressions (happy, approving, proud, grateful, and impressed) have significantly lower matching rates regardless of country. The matching rate of the “impressed” expression is significantly lower than that of any other expression (by analysis of variance and Scheffe’s method of multiple comparison,  $p < 0.01$ ).

This indicates that the subjects’ interpretation of positive expressions (sad, disapproving, angry, and confused) are similar to the designers’ intentions regardless of country and that the subjects’ answers to those expressions are similar across countries. On the contrary, the subjects’ interpretation of positive expressions (happy, approving, proud, grateful, and impressed) varies across countries.

We further analyzed the answers for the twelve expressions by principal component analysis. The results show that positive expressions (happy, approving, proud, grateful, and impressed) get mixed up ( $p < 0.01$ ). In other words, the reason for the positive expressions’ low matching rate is that each of those four expressions is not distinguished from the others.

Subjects’ comments support this result. Both Japanese and non-Japanese commented that they had difficulty in selecting the expressions matching “approving,” “grateful,” and “impressed.”

### 3.4. Differences in interpreting confusing expressions

We next analyzed the answers to the “impressed” expression, which has the lowest recognition accuracy. Figure 5 shows a breakdown of the answers to the “impressed” expression by country. Analysis by chi-squared test and Scheffe’s method of multiple comparison indicates that Japanese answers to the “impressed” expression are significantly different from those of other countries ( $p < 0.01$ ). In particular, the answers from Germany are most different from those of Japan ( $p < 0.01$ ), followed by the United Kingdom and Mexico ( $p < 0.01$ ).

As seen in Figure 5, 70% of Japanese interpreted the “impressed” expression as “impressed,” while the majority of participants in each of the other countries did not make this association. We can assume that this is further evidence for the in-group advantage within a country.

Statistical Analysis (chi-squared test and Scheffe’s method of multiple comparison) indicates that the “impressed” expression is mixed up with “happy,” “approving,” “proud,” and “grateful” ( $p < 0.01$ ). All of those expressions are positive, supporting the finding in sec. 3.3 that positive expressions get mixed up with each other.

We then conducted principal component analyses for the answers for other expressions. The results show that the subjects tend to mix up “ashamed” with “remorseful,” “confused” with “surprised,” and “disapproving” with “angry” ( $p < 0.01$ ), although negative expressions have higher recognition accuracy as stated earlier.

## 4. Discussion

The results of overall recognition accuracy show that avatar expressions designed by one culture’s drawing techniques are recognized with significantly higher accuracy by subjects from the same country than by those of other countries. The recognition accuracy of a neighboring country (Korea in this experiment) is the second highest. The in-group advantage mainly occurs within the same country where the expresser and recognizer belong to the same culture, and the degree of recognition accuracy is next highest between

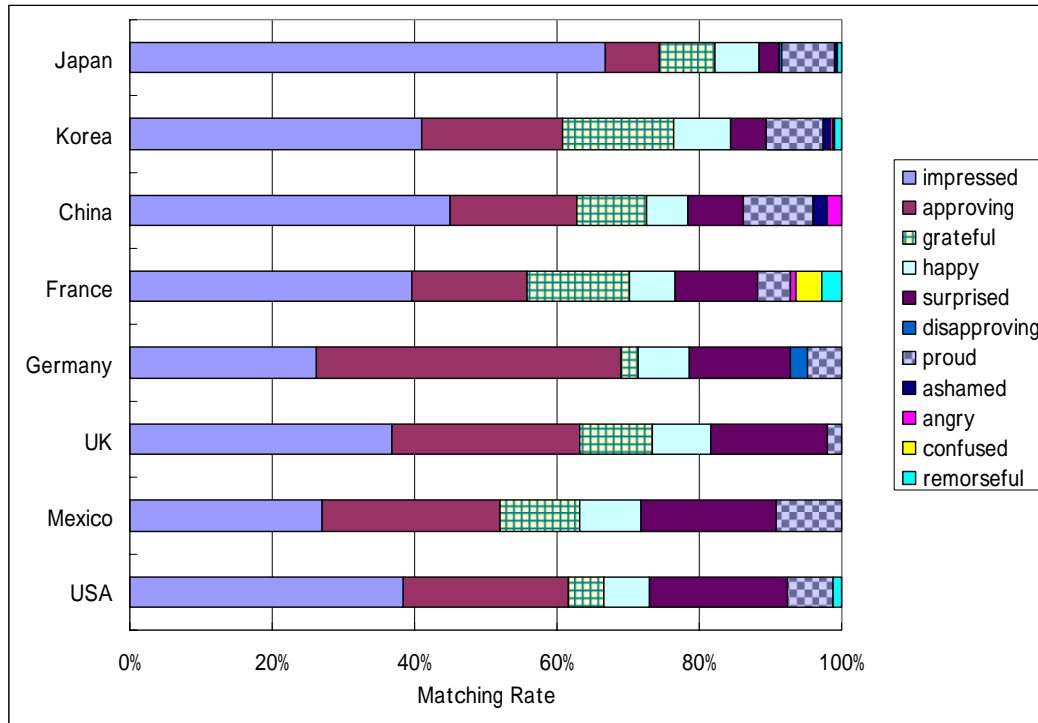


Figure 5. Answers to "impressed" expression by country

neighboring countries [13]. This result confirms that the in-group advantage that occurs in human expression recognition is applicable to avatar expression recognition within a country and between neighboring countries.

The results of breaking down the answers to the most confusing expression, "impressed," show there are cultural differences in interpreting the expression. The expressions that are mixed up with "impressed" within Japan are significantly different from those of other countries. This result provides further evidence of the in-group advantage within a country in interpreting avatar expressions.

The results of the negative expressions having significantly higher recognition accuracy than the positive expressions may indicate that the "decoding rule" in psychological studies is applicable to avatar expressions. Mixing up expressions occurs within positive/negative expression groups other than "confused" and "surprised." Accordingly, we can be less concerned about misunderstanding positive emotions as negative ones or vice versa. However, connotations and implications of each expression, for example, whether one is approving or grateful within the positive expression group, are not recognized accurately across cultures. For example, the communication gap between China and Japan caused

by different interpretations of the "big-eyed" expression in our earlier experiment [8] is one example of a confusing experience for the subjects, although it did not lead to a serious misunderstanding.

One of the reasons that there are in-group advantages within Japan and between Korea and Japan may arise from characteristics of Japanese comics. References on the study of comics point out that the Japanese comics market has expanded in a unique manner [17], and during this expansion, Japanese comics have developed various drawing styles, i.e., gestural line styles to show certain movements [18]. We used avatar designs by Japanese designers only in this experiment to limit the expresser to one country. Further study should be done to evaluate avatars designed by artists of other cultures, e.g., European or American.

## 5. Conclusion

There have been many psychological studies on human emotion recognition and related cultural differences. However, avatars are used based on the implicit assumption that avatar expressions are interpreted universally among all cultures in online communication across cultures.

In this study, we conducted an experiment comparing cultural differences in recognizing avatar



expressions to test this implicit assumption. The experiment was conducted as an open web experiment to gather various interpretations of avatar expressions from all over the world. We pursued two research issues in this study:

1) Identifying cultural differences in interpreting avatars' facial expressions by applying psychological findings on cultural differences in recognition of human facial expressions to the case of avatar expressions.

2) Identifying avatar facial expressions that are recognized differently across cultures.

The results show the following.

1) Cultural differences do exist in interpretation of avatar facial expressions, which confirms the psychological findings that physical proximity affects recognition accuracy. The in-group advantage was found within Japan and between Korea and Japan.

2) There are wide differences among cultures in interpreting positive expressions, while negative expressions had higher recognition accuracy.

Although misinterpretation is less likely to occur between positive and negative expressions, we need to design avatar expressions carefully to convey accurate emotions regardless of culture. We expect that further investigation of avatar representation and interpretation would serve as a design guideline for universal avatar expressions that could avoid the risk of miscommunication.

## Acknowledgements

This research was supported by the Universal Design of Digital Cities Project of CREST/Japan Science and Technology Agency and a Grant-in-Aid for Scientific Research (A) (15200012, 2003-2005) from the Japan Society for the Promotion of Science (JSPS). Naomi Yamashita, researcher at NTT Communication Science Laboratories, gave tremendous support to our analysis of the data.

## References

- [1] MSN Messenger: <http://messenger.msn.com>
- [2] Yahoo! Messenger: <http://messenger.yahoo.com/>
- [3] Smiley Central: <http://www.smileycentral.com/>
- [4] Damer, B., *Avatars: Exploring and Building Virtual Worlds on the Internet*. Berkeley: Peachpit Press, 1997.
- [5] Kurlander, D., Skelly, T., and Salesin, D., *Comic Chat*. Proceedings of Computer Graphics and Interactive Techniques, ACM Press, New York, 1996, pp. 225-236.
- [6] Smith, M.A., Farnham, S.D., and Drucker, S.M., *The Social Life of Small Graphical Chat Spaces*. Proceedings of CHI, ACM Press, New York, 2000, pp. 462-469.
- [7] Pesson, P., *ExMS: an Animated and Avatar-based Messaging System for Expressive peer Communication*. Proceedings of GROUP, ACM Press, New York, 2003, pp. 31-39.
- [8] Koda, T., *Interpretation of Expressive Characters in an Intercultural Communication*, KES2004, LNAI 3214, Part II, Springer-Verlag, Berlin, 2004, pp. 862-898.
- [9] Isbister, K., Nakanishi, H., and Ishida, T., *Helper Agent: Designing and Assistant for Human-Human Interaction in a Virtual Meeting Space*. Proceedings of Human Factors in Computing Systems (CHI2000), ACM Press, New York, 2000, pp. 57-64
- [10] Bartneck, C., Takahashi, T., and Katagiri, Y. *Cross Cultural Study of Expressive Avatars*. Proceedings of the Social Intelligence Design 2004.
- [11] Ekman, P., *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. Henry Holt and Company, 2003.
- [12] Ekman, P., *About Brows: Emotional and Conversational Signals*, Cranach, M.V., Foppa, K., Lepenies, W., and Plog, D (eds.), *Human Ethology: Claims and Limits of a New Discipline: Contributions to the Colloquium*, Cambridge University Press, Cambridge, 1979, pp. 163-202.
- [13] Elfenbein, H. A. and Ambady, N. *On the Universality and Cultural Specificity of Emotion Recognition: A Meta-Analysis*. Psychological Bulletin, Vol. 128, No. 2, American Psychological Association, Inc., 2002, pp. 203-235.
- [14] Elfenbein, H. A. and Ambady, N. A., *Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition*. Journal of Cross-Cultural Psychology, 34, 2003, pp. 92-110.
- [15] *The Universal Character Experiment* <http://character.kuis.kyoto-u.ac.jp/>
- [16] Ortony, A., Clore, G.L., and Collins, A., *The Cognitive Structure of Emotions*. Cambridge Univ. Press, Cambridge, 1998.
- [17] McLoud, S., *Reinventing Comics: How Imagination and Technology Are Revolutionizing an Art Form*, Perennial, New York, 2000, pp. 118,122-123.
- [18] McLoud, S., *Understanding Comics: The Invisible Art*, Harper Perennial, New York, 1993, pp. 131-133, 210.