KEER2010, PARIS | MARCH 2-4 2010
INTERNATIONAL CONFERENCE ON KANSEI ENGINEERING AND EMOTION RESEARCH 2010

POSE GENERATION SYSTEM EXPRESSING FEELINGS AND STATE

Tomoko EMI*, Masafumi HAGIWARA

Faculty of Science and Technology, Keio University,
3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan

ABSTRACT

A pose generation system expressing feelings and the state is proposed in this paper. The proposed system generates a pose in accord with the image that the user wants to make, and it outputs avatars with the poses. The operation and the usage of the system are very simple, and high quality avatars with poses can be obtained. First, the user chooses onomatopoeias which are close to the image of the pose that the user wants to create. Onomatopoeia is a word, or occasionally, a grouping of words, with a sound imitating the sound it is describing, such as "buzz" for a bee, "meow" for a cat, and "moo" for a cow. Then the system determines the initial values of the parameters to constitute a pose from the chosen onomatopoeias. Various poses are generated by interactive genetic operations based on the initial values, and they are displayed to the user. Then the user selects and evaluates favorite poses, and the system learns preference of the user based on the evaluations. In this way, interactive genetic operations are performed and more and more suitable poses are generated. Furthermore, addition of the expression to the pose is possible: an expression and various feelings/state expression by the pose are realized. A viewpoint can be moved freely on the screen. Expression by the pose and the effectiveness of the proposed system are confirmed through evaluation experiments.

Keywords: *Pose generation, Kansei, Onomatopoeia, avatar*

1. INTRODUCTION

Owing to rapid development of communication and Internet, studies on media contents and communication tools become active. By introducing information processing technology into contents production, new kinds of works and scientific analysis of arts become possible [1]. Although human's kansei, sensibility, are objective and qualitative, if they can be treated subjectively and quantitatively, computers will be able to process human's kansei. A lot of studies have been carried out to treat contents by kansei engineering approach [1]-[7].

* **Corresponding author**: hagiwara@soft.ics.keio.ac.jp

E-mail and PC communication were popular in the first half of 1990's as communication tools. Various kinds of Web applications have been widely used nowadays. As for bulletin boards in Internet, we can use not only plain texts but also decorated characters and images because of development of broad band Internet. In addition to e-mail and bulletin board, blog, social network services, and messengers are appeared and have obtained large popularity. Since these communication tools are mainly based on characters, it is difficult to express delicate emotional and kansei expressions. Icons, decorated letters, are popular in

Japanese cellar phones. Avatar is another form to express user's emotion and kansei. The avatar is a character expressing the other self of the user on the Web. A lot of studies on avatar have been performed [8]-[12]. There are various kinds of avatars such as animals, mascots, fish, and humans. Among them, especially human of whole body avatars are the most popular. Avatars can be widely used in such situations:

1) Blog: Avatars contribute to visualizing the writer.

2) Game& chat: Avatars can be used as images of players.

In addition, ref.[13] reports that usage of avatars can boost participation sense of users and affinity among users.

On the other hand, avatars themselves are difficult to express emotions and states because of lack of poses. If poses and viewpoints are added to avatars, they can express more abundant information.

In this paper, a pose generation system expressing feelings and the state is proposed. The system generates poses in accord with the image that the user wants to make, and it outputs avatars with the poses. In Sec.2, 3-dimentional characters are explained. A pose generation system is proposed in Sec.3. Experimental results are shown in Sec.4. Sec.5 concludes the paper.

2. 3D CHARACTERS

2.1. Construction of a character

In the proposed system, a 3D character made by 3D modeling software is used as a fundamental avatar. We treat avatars to express whole size of human, from head to foot. A fundamental 3D character is constructed by 17 parts based on a human joint model and needs the setting of 14 joint coordinate data.

Fig.1 shows joint positions and their names of the 3D character. A 3D character is composed of the following independent parts: hair, a head, a face, a chest, an upper arm (right and left), an arm (right and left), a hand (right and left), a waist, a leg (right and left), shin (right and left), and a foot (right and left). These parts consist of tree structure joined by joints.

2.2. Mechanism of the pose generation

Each part of 3D character above-mentioned has angle of rotation $arg(x, y, z)$. This angle of rotation $arg(x, y, z)$ is composed of angle $arg(x)$ which turned around x-axis, angle $arg(y)$ turned around y-axis, and $arg(z)$ turned around z-axis. Angle of rotation $arg(x, y, z)$ is the parameter that is necessary for pose generation, and a 3D character has 17 parts in total.

One pose is defined by 51 parameters in total. Each part turns $arg(x, y, z)$ around the coordinate of upper-side joint in a tree structure. For example, the upper arm turns around a joint coordinate of shoulder. Each part has its own movable range to create natural poses.

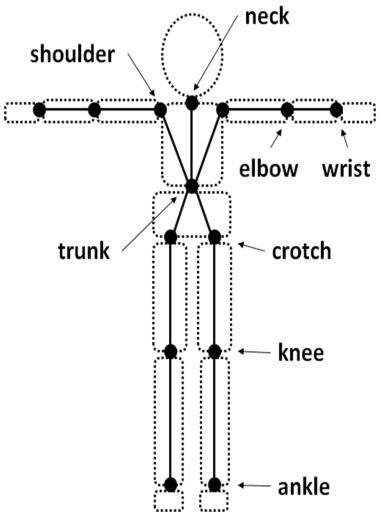


Figure 1: Joint positions and their names.

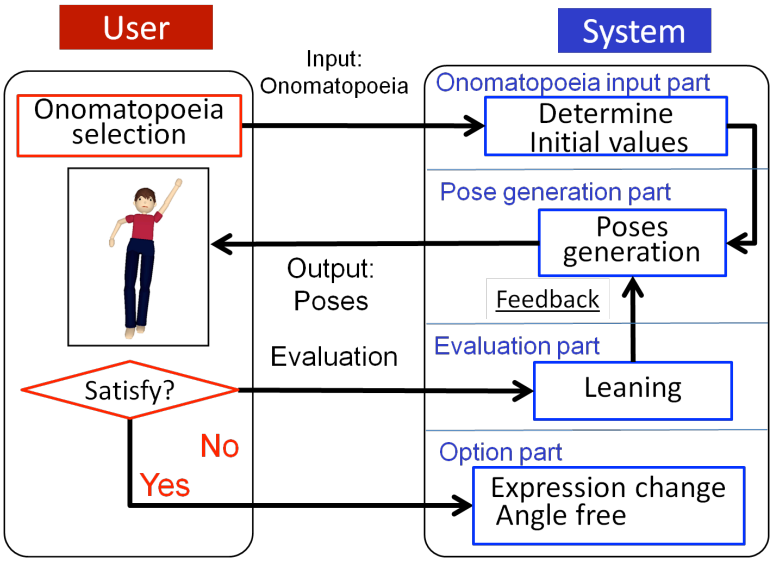


Figure 2: Flow of the proposed system.

3. POSE GENERATION SYSTEM

3.1. Construction of a character

Fig.2 shows the flow of the proposed system to create a pose. The system is composed of an onomatopoeia input part, pose generation part, evaluation part, and an option part. First, a user determines the image of the pose to create.

In the onomatopoeia input part, the image of the pose that the user wants to make is input to the system as onomatopoeias. The user can choose at most three from the prepared onomatopoeias. The selected onomatopoeia is input into the system, and the system generates the initial values of the parameters to constitute the pose.

The pose generation part has two processes: the initial generation and the following generation. In the initial generation, the system generates an initial pose based on values a from the onomatopoeia input part. In the following generation, the next poses are generated using the evaluation of the former pose obtained at the evaluation part.

The proposed system displays ten avatars to the user. The user evaluates the poses using the evaluation part. The evaluation is fed back to the pose generation part. And the system learns the image of the pose that the user wants to make. If the pose the user satisfies is generated, the genetic operations are ended and the optional operations, expression change function and angle-free function, are carried out if the user wants. The final avatar with pose is stored.

In the following, details of each part are explained.

3.2. An onomatopoeia input part

In the onomatopoeia input part, the image of the pose that the user wants to make is input to the system as onomatopoeia. The user can choose prepared onomatopoeia to three. The chosen onomatopoeia is input into the system, and the system generates the initial values of the parameters to constitute a pose.

Onomatopoeia has the following features:

- It is direct sensible expression;
- There is the commonality of the meaning with forms of words;
- There is common sense as the Japanese in a sound.

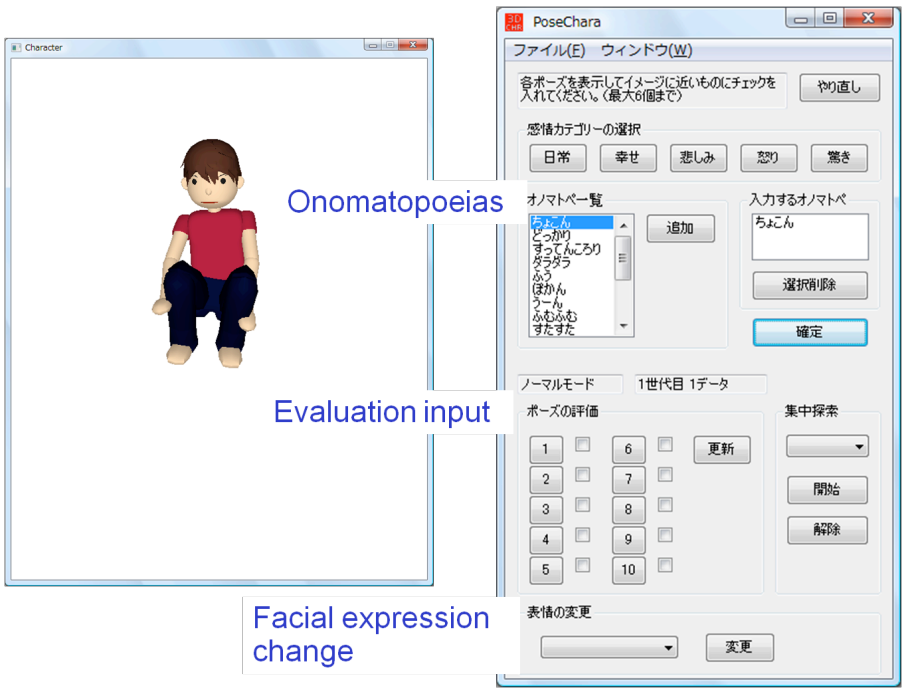


Figure 3: Interface window.

Table 1: List of onomatopoeias used in the proposed system.

Daily		Happiness		Sorrow		Angry		Surprise	
ちょこん	cyokon	るんるん	run run	とぼとぼ	tobotobo	イライラ	iraira	ぎくっ	gikku
どっかり	dokkari	やったー	yatta-	ぐす	gusu	プンプン	pun pun	はっ	hattu
すってんころり	suttennkorori	わーい	wa-i	しくしく	sikusiku	むすっ	musu	ぎゃー	gya-
ダラダラ	daradara	うふふ	ufufu	うわあん	uwaan				
ふう	fuu	わくわく	wakuwaku	がっかり	gakkari				
ぽかん	pokan			ガーン	ga-n				
うーん	u-n			へなへな	henahena				
ふむふむ	fumufumu			ずうん	zuin				
すたすた	sutasuta			ぽつん	potun				
タッタッタッ	tattatta								
ぐったり	guttari								
ぜえぜえ	zeizei								
ばたんきゅう	batanncyu								
すやすや	suyasuya								

Fig.3 shows the interface window. The input of the onomatopoeia in the proposed system is the followings:

Step 1:
The user chooses one emotional category from five emotions categories: "happiness", "sorrow", "anger", "surprise", and "daily " .

Step 2:
From a list of displayed onomatopoeias, the user chooses onomatopoeia in accord with the image and adds it.

Step 3:
By repeating Step 1 and Step 2, the onomatopoeias at most three are added.

Step4:
The user pushes the decision button.

Table 1 shows a list of onomatopoeias used in the proposed system. They are selected based on the opinions of 42 subjects and are classified.

3.3. Pose generation part

This system operates by a framework of interactive genetic algorithm. The user performs a subjective evaluation, and the system estimates parameter values to create a pose from the evaluation. These parameters are arranged in a line to form an individual in genetic algorithm.

Fig.4 shows the gene expression of the parameters to constitute a pose. A flow of interactive genetic algorithm in the proposed system is the followings:

Step 1: Generation of initial individuals
Step 2: Genetic operations
(selection, crossover, mutation)
Step 3: Judgment to stop or not
Step4: Go to Step 1 (new generation)

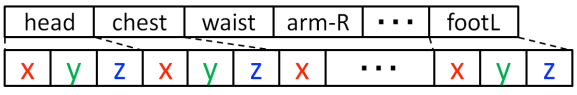


Figure 4: Gene expression to constitute a pose.

3.4. Evaluation part

In the evaluation part, each displayed pose is evaluated. The user evaluates the pose and the system estimates the fitness value of the pose. Individuals to survive to the next generation are decided by the fitness value. In this way, the evaluation is fed back to the pose generation part, and the system learns the image that the user wants. The interactive genetic algorithm based operation is repeated until a pose to be satisfied is generated. The fitness value of each individual in the above-mentioned genetic algorithm is calculated by the value of each chromosome in the individual $p^{(g)}$. The fitness value of the individual $p^{(n)}$ is calculated as,

$$Fitness_n = E - \sum_{i=1}^{51} \left(\frac{p_i^{(n)} - p_i^{(g)}}{l_i} \right)^2 \quad (1)$$

$$l_i = \max_i - \min_i$$

Here, E is the constant, l_i is the range of chromosome i which is used to normalization.

3.5. Option part

The proposed system has two kinds of options: an expression change function and an angle-free function. The user uses these functions as needed after obtaining a preferable pose. The expression change function is the function that can change the expression of the avatar. When the user wants to change "a plain expression" of the initial setting, the user can select a facial expression from 11 kinds of expressions: "plain expression", "happy", "delight", "sad", "tears", "anger" ,"surprise", "hatred", "fear", "sweat", and "sleep"

11 kinds of expressions are shown in Fig.5. These are determined by referring 6 basic feelings (happiness, sorrow, anger, surprise, hatred, fear) by Ekman [14] and handwritings and decorated letter expression [13], which are popular in Japan.

Fig.5 11 kinds of expressions.

4. EXPERIMENTS

We carried out the following two kinds of experiments to evaluate the proposed system.

1) Expression by facial expression and pose.

2) Usefulness of the proposed system.

4.1. Methods of experiments

Experiment 1

First, we prepared 3 patterns of avatars, "facial expression", "pose", and "facial expression and pose" for 5 kinds of themes, happy, angry, sorrow, surprise, and depression. There were 23 subjects to cooperate for the experiment 1. The subjects evaluated the expression with five ranks ([1-5], 1: bad, 5: good) to see the avatars of 3 patterns in each theme.

Then we prepared the various variations of avatars using the expression change function. Eight patterns of expressions, six fundamental expressions proposed by Ekman [20], plain, and additional one pattern such as sleepy, were used. The subjects were asked to see the 8 patterns of avatars and select the suitable behavior.

One of the examples is the following:

Mmm

Expression: plain, happy, sad, angry, surprise, hatred, fear, sweat

Behavior: deliberate, worry, plan, brood, flash, recall, miss, at a loss, get trouble, hurry, fear, others

Experiment 2

In the experiment 2, we evaluate the reflection degree of user's image in order to evaluate the effectiveness of the proposed system.

In this experiment, we assume that the user submits text and avatar to a blog. At first, subjects determined a theme to submit a blog and then the subjects were asked to make an avatar suitable to the contents. And the subject evaluated the created avatars from the following 6 view points with five ranks. In the experiment 2, 16 subjects were cooperated.

1. Was a pose in accord with an image generated?

2. The variety of the generation pose.

3. Do you feel that you were able to express your feelings and state by the generated avatars?

4. Was the onomatopoeia input useful to convey the image of the user?

5. How about the operability of the system?

6. How about the entertainment of the system?

4.2. Results

Figs.6-10 show 3 patterns of avatars used in the experiment 1. Table 2 shows the averaged value of the evaluation. It can be seen that addition of pose is effective to express the onomatopoeia.

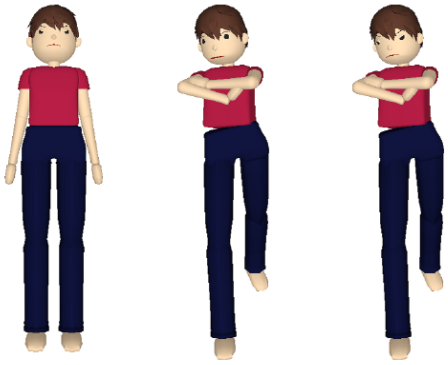
Figs.11-12 show 8 patterns of avatars with fixed one pose and different 8 facial expressions and the selection results by the subjects. We can see that various states can be expressed by changing facial expressions even if the pose is the same.

Fig.13 shows some examples of created poses for various situations.

Table 3 summarizes the averaged value for each question.

(1) Facial expression (2) Pose (3) Both

Fig.6 Happiness.



(1) Facial expression (2) Pose (3) Both

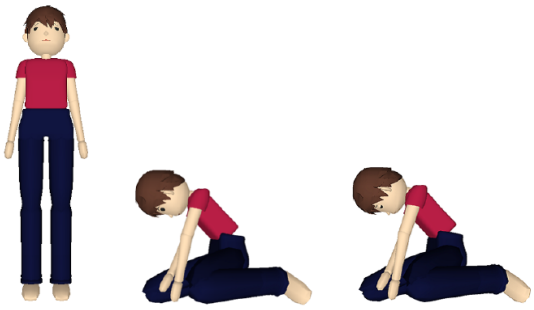
Fig.7 Angry.

(1) Facial expression (2) Pose (3) Both

Fig.8 Sorrow.

(1) Facial expression (2) Pose (3) Both

Fig.9 Surprise.



(1) Facial expression (2) Pose (3) Both

Fig.10 Depressed.

Table 2 Evaluation of Expression (averaged).

Theme	Facial Expression	Pose	Both
Happiness	2.83	1.74	4.48
Angry	3.26	1.22	4.26
Sorrow	3.83	1.48	4.30
Surprise	2.91	1.65	3.91
Depressed	2.83	2.26	4.61

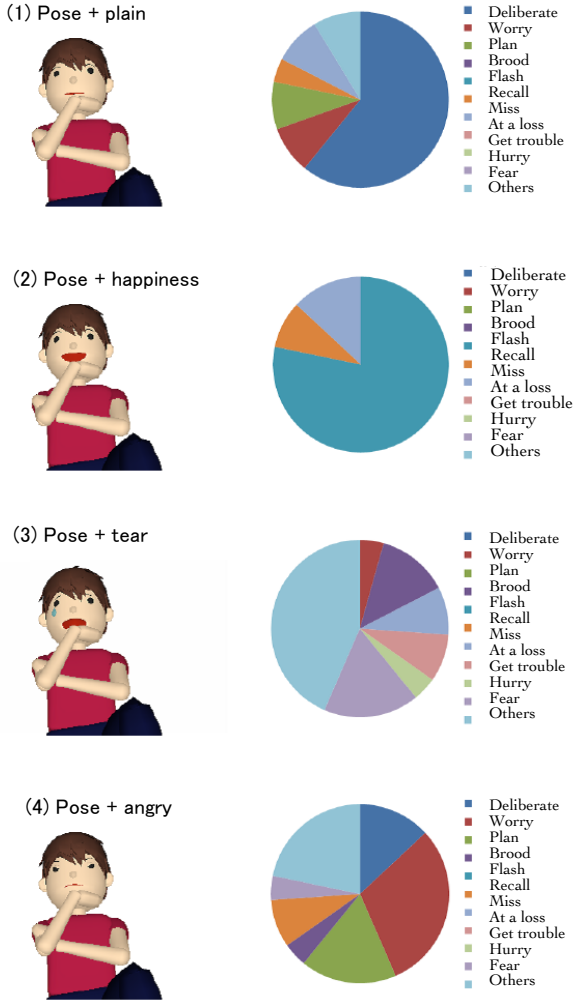


Fig.11 8 patterns of avatars with fixed one pose and different 8 facial expressions and the selection results. (1/2)

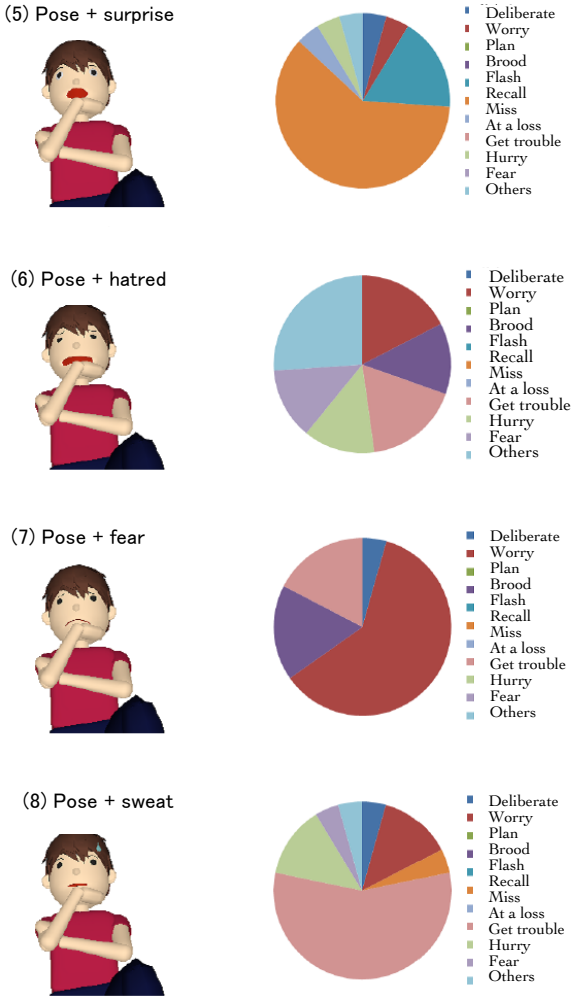


Fig.12 8 patterns of avatars with fixed one pose and different 8 facial expressions and the selection results. (2/2)

Fig.13 Some output examples for different scenes.

Table 3 Summary of evaluation results.

Questions	Ave.
1. Was a pose in accord with an image generated?	3.9
2. The variety of the generation pose.	4.4
3. Do you feel that you were able to express your feelings and state by the generated avatars?	4.4
4. Was the onomatopoeia input useful to convey the image of the user?	4.1
5. How about the operability of the system?	4.1
6. How about the entertainment of the system?	4.7

5. CONCLUSIONS

We have proposed a pose generation system expressing feelings and the state. The system generates a pose in accord with the image that the user wants to make, and it outputs avatars with the pose. The operation of the system is very simple, and the generation of the high quality avatars with poses is possible.

Expression by the pose and the effectiveness of the proposed system have been confirmed through evaluation experiments.

REFERENCES

[1] Katayose Haruhiro, Kasao Atsushi: Multimedia Contents and Its Evaluation, Transactions of Information Processing Society of Japan, Vol.44, No.2, pp. 222-226, 2003.

[2] Yoshitaka Atsuo, Matsui Ryoji, Hirashima Tsukasa: Extracting Movie Rendition Based on Camera Work Detection and Classification, Transactions of Information Processing Society of Japan , Vo.47, No.6, pp.1696-1707, 2006.

[3] Ishizuka Kenkichi, Kato Shigeru, Onisawa Takehisa: Transformation of Theme Music Based on Impressions of Stories Generated From Pictures, Journal of Japan Society for Fuzzy Theory and Intelligent Informatics , Vo.19, No.4, pp.378-391, 2007.

[4] Takanobu Obata, Masafumi Hagiwara: A Color Poster Creating Support System to Reflect Kansei, Transactions of Information Processing Society of Japan , Vol.41, No.3, pp. 701-710, 2000.

[5] Hirokazu Hirata, therole and organization of an emotion design, Journal of Japan Society of Kansei Engineering, Vol.8, No.2, pp. 183-188, 2009.

[6] Korenaga Motoki, Masafumi Hagiwara: An Interior Layout Support System with Interactive Evolutionary Computation, Transactions of Information Processing Society of Japan, Vol.41, No.11, pp. 31152-3160, 2000.

[7] Nishiyama Haruhiko, Ohkubo Tatsuma, Saito Shinsuke, Matsushita Yutaka: Virtual Space Produce Based on Composition Knowledge, The transactions of the Institute of Electronics, Information and Communication Engineers, Vol.J81-D-2, No.1, pp. 146-155, 1998.

[8] Hisashi Ito, Masafumi Hagiwara: Character-agent automatic creating system, Journal of Japan Society of Kansei Engineering, Vol.5, No.3, pp.11-16, 2005.

[9] Ogawa Takefumi, Tsukamoto Masahiko, Nishio Shojiro: A Communication Support System on WWW Using Images, Transactions of Information Processing Society of Japan , Vol.43, No.11, pp. 3364-3374, 2002.
[10] Kusumi Takashi, Komeda Hidetsugu, Kojima Takatsugu : Improving Communication in 3D-MUD (Multi User Dungeon) by Using Avatar's Facial Expression Features, Japan journal of educational technology,Vo.31, No.4, pp. 415-424, 2008.
[11] Ito Kyoko, Kozuki Masanori, Ishii Hirotake, Yoshikawa Hidekazu: Experiment on a Networked Community Using Character Agents as Avatar and Navigator, Transactions of Information Processing Society of Japan , Vo.44, No.7, pp. 1812-1827, 2003.
[12] Inoue Masayuki, Mochizuki Takayoshi, Tsuboi Toshiaki: A Study and Evaluation of the Operability and Avatar-Design for Three-Dimensional Cyber Space, The journal of the Institute of Image Information and Television Engineers, Vol.60, No.8, pp. 1296-1306,2006.
[13] Tomoko Kanda, Toru Ishida: Cross-cultural Comparison of Interpretation of Avatars' Facial Expressions, Transactions of Information Processing Society of Japan ,Vol.47, No.3, pp. 731-738, 2006.
[14] P.Ekman , W. V.Friesen, (Tikara Kudo translate) : Hyojyo Bunseki Nyumon (Introduction of Facial Expression Analysis), Shisei syobo, 1987.