An Empirical Analysis on Emotional Body Gesture for Affective Virtual Communication

Prabath Weerasinghe[1], R.P.C. Janaka Rajapakse[2] and Ashu Marasinghe[1]

[1] Kansei Information Engineering Laboratory, Nagaoka University of Technology
Nagaoka, Niigata 940-2188, Japan
[2] Graduate Institute of Animation and Film Art, Tainan National University of the Arts
Guantian , Tainan 72045, Taiwan

Abstract

Development of the engineering on virtual environment over last decades have been attempted to establish a metaphor of human communication scheme, utilizing avatar models as user proxies. Many of the avatar models in present Collaborative Virtual Environments have still built on the universality hypothesis, which tends to ignore the multiplicities of human expressions. In such context, this research attempted to conduct an empirical study on the emotional body gestures of avatar models considering the differences in terms of age, gender, body size and humanoid. The observable differences in average emotional intensity levels implied that the age, gender, body size and humanoid of avatar effect on expression of emotions, especially the differences of gender and humanoid are highly significant in the results revealed. The findings of the study factually demands future development of Avatar models based on demographic and physical features to improve overall-effectiveness of virtual communication.

Keywords: *Collaborative Virtual Environment, Effective Communication, Emotional Expressivity, Kansei Engineering, 3D Model*

1. Introduction

Collaborative Virtual Environments (CVEs) have been ever more popular over last few decades in several domains such as distributed simulations, 3D multiplayer games, collaborative engineering software, and e-learning. Many research works have constructively contributed to improve the effectiveness of virtual communication agents or avatar models. Ability of these virtual agents to be connected with the user on a personal and emotional level is a crucial factor in enhancing the effectiveness of virtual communication environment. 'Thus, a life- like human face can enhance interactive applications by providing straightforward feedback to and from the users and stimulating emotional responses from them. Numerous applications can benefit from employing believable, expressive characters since such features significantly enhance the atmosphere of a virtual world and communicate messages far more vividly than any textual or speech information' [4]. "Understanding how emotional expressions and 'body language' relate to and reflect emotional state, and developing avatars capable of projecting that state, is an extremely important aspect of enabling rich social interaction in virtual environment" [5].

The understanding that the agents who exhibit personality, livability and emotions through affective behaviors can be better related with users has caused to grow the research attempts on developing virtual avatar models with strong emotion sensitivity. "The interest on the degree of universality of emotions arose actively more than half a century ago (Ekman & Friesen, 1971) and since then several hypothesis on the universality of emotions have been presented (Izard, 1994). The vast majority of works done on automatic recognition of emotions assume that emotions are universal and the cultural variable of each individual is ignored [whereas] current works in psychology shows that there is no agreement on the universality hypothesis" [6].

Many of the emotion databases use in present Collaborative Virtual Environments have still built on the universality hypothesis, which tends to ignore the multimodal behavior of social-beings. There are some databases currently available for this purpose (Gunes et al., 2011 cited in [6]) but many of them do not address the differences of avatar models based on physical and demographic variables such as age and gender. Furthermore, the very limited applications available also have considered the facial expressions but not adequately addressed the multiplicity of gestures and postures.

In such context, this research attempts to conduct an empirical study on the multiplicity of emotional body gestures of avatar models based on physical and demographic variables. The study aims to compare emotional body gestures among selected virtual corpses, supporting the limited domain of emerging studies to solve the problem of universality while facilitating to improve

the overall-effectiveness of virtual communication agents.

2. Background of the study

2.1 Expression of emotions in Avatar Models

Expression of emotions plays a vital role in communication between social-beings. 'In everyday life, people communicate using speech as well as their face and their body to express their emotions' [4]. 'For decades now, emotions and their expression and understanding have been studied in several fields and the evidence shows the important role they play in our daily lives' (Gratch et al., 2009 cited in [6]). Collaborative Virtual Environment is one such domain where applied researches have been extensive carried out over last few decades.

'In Collaborative Virtual Environments, the real world actions of a human has been transferred into the virtual environment through a representative (avatar), while the virtual world perceives these actions and corresponds through respective system avatars who can express their emotions using human-like expressions and gestures' [4]. In this domain, a textual characterization of an agent's mood or a body expression is generally referred an emotion. However, emotions are very complex in nature and it is extremely difficult to be isolated with clearly defined limits. In some situations, emotions mix with each other it is a well-taken challenge to model such emotions in virtual environments. Many of the recent Kansei engineering researches also have been attempted to synchronize human expression into virtual agents in order to familiarize the robot with human society [8].

Initial attentions of transforming emotions to the avatar models have been more focused on facial expressions. "Since publication of Darwin's *The Expression of the Emotions in Man and Animals* (Darwin, 1872/1965), a great deal of attention has focused on how emotions are communicated through facial expression, and much has been learned ranging from knowledge of the ways in which individual muscles contribute to each expression (e.g. Ekman & Friesen, 1984; Izard, 1979; Rinn, 1984) to the similarities and differences between cultures in emotional perception (e.g. Boucher & Carlson, 1980; Ekman *et. al*., 1987; Haidt & Keltner, 1999; Mesquita & Frijda, 1992; Russell, 1991). A similar although smaller literature exists on the perception of emotion from the voice (e.g. Banse & Scherer, 1996; Murray & Arnott, 1993; Scherer, 1986, 1995; van Bezooijen & Boves, 1986; Wallbott & Scherer, 1986)" [7].

Incorporating basic facial expression was capable to add a little human touch to avatar models but 'human bodies are large objects possessing multiple degrees of freedom which would appear to be ideal channels for emotional communication (Montepare, Koff, Zaitchik, & Albert, 1999), especially at distances where the recognition of emotion from facial expression is difficult (Walk & Walters, 1988). As an avatar model being considered as a real projection of a human user in collaborative virtual environments, its ability to reflect one's emotional state is crucial. Therefore, understanding how emotional expressions and "body language" relate to and reflect emotional state, and developing avatar models capable of projecting that state, is an extremely important aspect of enabling rich social interaction in virtual environments.

Hence, in the line of gradual improvements of the avatar models, researchers paid more attentions to improve the signals of emotional body language, which explains how to express the emotions through the body pose and motion of avatar models. 'Recently, researchers have also turned to emotional body language, i.e. the expression of emotions through human body pose and/or body motion (de Gelder, 2006; Grezes, Pichon, & de Gelder, 2007; Meeren, van Heijnsbergen, & de Gelder, 2005; Peelen & Downing, 2007) [1]. Among a number of commendable research attempts on this regard, many still tend to focus on general properties and fail to represent the dynamic human emotional status in avatar models. In fact, character usage for emoticons varies across communities in terms of their cultural, demographic and physical properties.

This research explains how users perceive the differences of emotional body gestures among avatar models of different age, gender, body size and humanoid. The results of the study will contributed to further develop avatar models through incorporating distinct human emotions and make virtual agents more communicative in their applications of any virtual domain.

3. Method of the study

3.1 Participants

A group of thirty-one respondents, who were graduate students of the Nagaoka University of Technology,

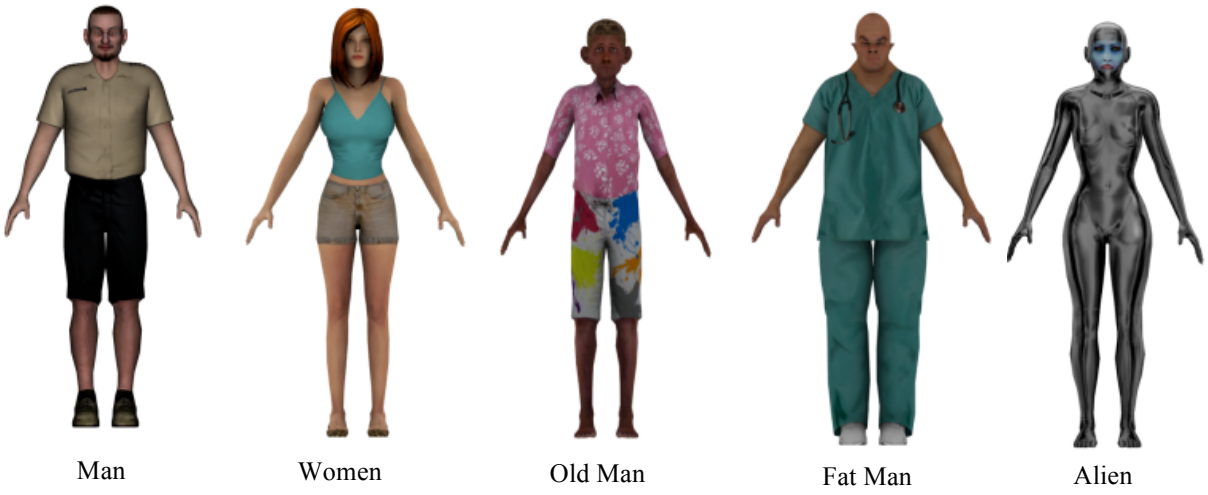

Fig.1 Five Different Virtual Agents

participated in this experiment. All of the respondents were healthy subjects who had been certified by the university medical officer as enduring eligible psychophysical conditions to undertake higher education without any recorded neurological disorders and with normal or corrected-to-normal vision. The group of thirty-one respondents belongs to the age group between twenty-six and thirty-five.

3.2 Stimulus Generation

"In everyday life people communicate using speech as well as their face and their body to express their emotions" [4]. There had been many studies focused on expression of emotions and its role in effective communication. "Broadly accepted studies on the hypothesis of universality of emotions (Ekman & Keltner, 1997) suggest there is a group of basic and universal emotions that are equally expressed and recognized in all the different cultures around the globe. New break through studies in psychology (Jack et al., 2012) neurology (Chiao et al., 2008) (Adams et al., 2010) and affective computing (Quiros-Ramirez & Onisawa, 2013) (Kamaruddin et al., 2012) have presented new evidence to support the specificity of emotions: culture seems to play an important role on the expression and recognition of human emotions" [5]. Accordingly, in this experiment, we have assessed a set of eight suggested emotional status such as anger, despire, interest, irritation, joy, pleasure, pride and sadness.

"Research into the visual perception of human emotion has traditionally focused on the facial expression of emotions. Recently researchers have turned to the more challenging field of emotional body language, i.e. emotion expression through body pose and motion" [1]. "Although much has been written about body posture and emotion, little systematic research has been carried out, and that which exists has tended to focus on emotional expression through movement rather than static posture. This lack of research is surprising when one considers that bodily expression is widely recognized as an important diagnostic tool in therapeutic practice (e.g. Berger, 1994; Dosamontes-Beaudry, 1997; Flack, Laird, & Cavallaro, 1999; Fuchs, 1996; Hirsch, 1994), and is a crucial feature of artistic forms such as acting and dance" [7]. Therefore, in this study, we have focused on motions of virtual agents rather than static images. There was a set of forty-five motions. This includes forty motions of eight suggested emotional status and five unclassified motions expressed by five different virtual agents (Fig.1) such as woman, man, old man, fat man and alien considering the physical, demographic and humane variations. Inclusion of an alien is objected to test the un-canny-valley hypothesis.

Forty-five FBX motion files were captured at 3-D motion capture studio of the Graduate Institute of Animation and Film Art, Tainan National University of the Arts, Taiwan (Fig.2-a). As shown in Fig.2-b, the captured motion data were re-targeted to virtual agents. Forty-five correspondent .MOV files with a pixel size of 1920 X 1080 were rendered through motion builder 2015 software.

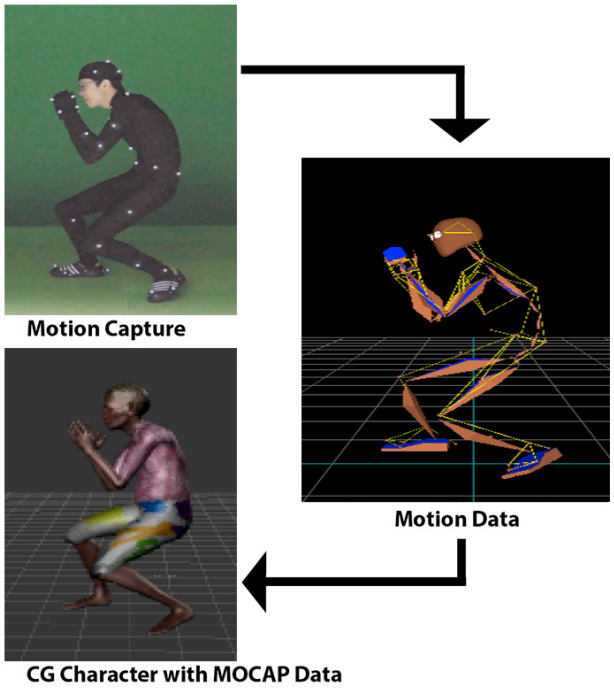

Fig. 2-a -Motion Capture Process.

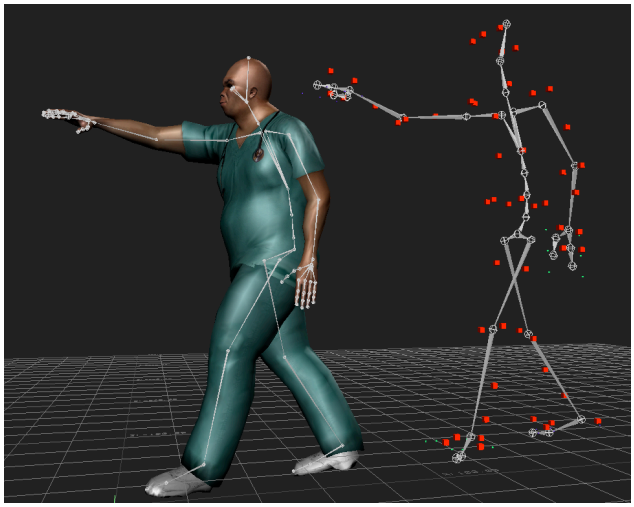

Fig. 2-b- Retargeting Motion Data to Virtual Agents.

3.3 Apparatus

All subjects were comfortably seated in a ventilated, light-controlled room for about 15-30 minutes keeping a reasonable time lap of five seconds between two images. Average duration of a motion clip was 8 seconds. The forty-five motion clips were uploaded into web interface and screened on a calibrated display at a video size of 1024 in width and 786 in height. Forty-five clips were randomly retrieved when each subject operate the series of clips.

Fig. 3 Experimental setup, the subject was comfortably seated in front of the LCD monitor in an experimental booth.

As shown in Fig.3, a white coloured experimental booth under controlled office light conditions was employed in performing the experiment. The experimental setup was executed on a standard PC equipped with a 2.9 GHz Intel Core i5 processor and NVIDIA GeForce GT 750M graphics processor with 1GB of GDDR5 memory.

3.4 Procedure

All subjects were briefed with a short introduction about the objective of the research and the procedures of the experiment prior to the commencement. Then, each subject was requested sit in the experimental booth in front of the specified monitor to fill a basic information questionnaire. Once confirmed that the subjects are ready with a calm mind, the web based online motion file series were opened on the screen. The subjects were asked to watch the clip and record the recognized emotion with the correspondent intensity. Five seconds later the first image, the second one was appeared and the process continued repeatedly for all forty-five images. Participants were asked to obtain an unlimited time as per their convenience but the maximum time they really spent did not exceed

fifty minutes for the experiment. Thirty-one participants went through the whole procedure individually recognizing all forty-five images.

3. Results

Emotions	Average of Emotion recognition rate
Anger	95%
Despair	90%
Interest	99%
Irritation	97%
Joy	90%
Pleasure	90%
Pride	84%
Sadness	91%
Average	**92%**

Table 1: Average of Emotion Recognition Rate

In the experiment, subjects were asked to judge all stimuli, which were randomly presented. Table 1 depicts the average emotion recognition rate of the subjects. Average emotional recognition rate of all studied emotions were higher as 92%. The highest emotion recognition rate was recorded for 'interest' (99%) whereas the lowest for 'pride' (84%). All of the other emotions were recognized by over 90% of the subjects. Differences between emotion discrimination thresholds were tested with one-way analysis of variance (ANOVA). The mean and standard deviation for average accuracy of responses were: Anger (mean =6.21875, sd: 2.1512), Despair (mean =5.36666, sd: 0.80871), Interest (mean =5.548387, sd: 1.3865), Irritation (mean =6.321581, sd: 1.92158), Joy (mean =5.366667, sd: 1.44993), Pleasure (mean =5.193548, sd: 1.42406), Pride (mean =6.655172, sd: 2.12596), and Sadness (mean =5.354839, sd: 1.81747). The one-way ANOVA demonstrated that there is a statistically significant (at the 1% level) difference (F (7, 237= 3.26) P= 0.002476) between emotions.

Fig.4 depicts the comparison of emotion expression intensity levels between female and male avatar models. The mean and standard deviation of female (mean =2.43, sd: 1.822) and male (mean =2.82, sd: 1.880) were versatile enough to produce a significant effect in ANOVA. Therefore, one-way ANOVA demonstrated that the effect of gender is significant (at the 1% level) for average intensity deference on emotions (F (1, 1407= 14.517 Sig= 0.000)). As visually expresses in Fig.4, the average intensity level of 'anger' is recognizably different as 2.10 for female avatar models whereas 2.65 for male avatar models. Similarly, average intensity level of other emotions also showed a difference according to the gender; for instance for the 'interest' emotion, the intensity levels for female and male were 2.15, 2.10 respectively.

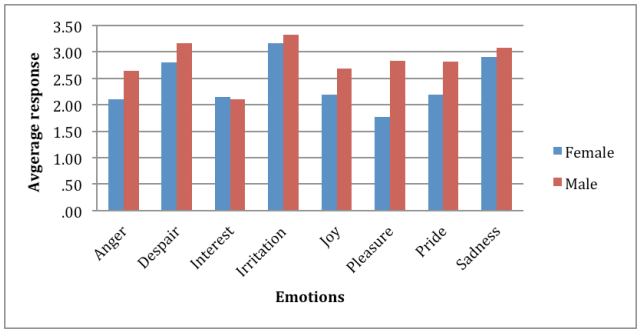

Fig. 4 Average response towards different emotion models: Female / Male

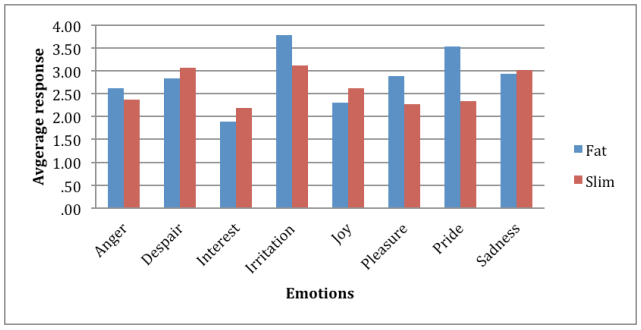

Fig. 5 Average response towards different emotion models: Fat / Slim

Fig.5 illustrates the comparison of emotion expression intensity levels between fat and slim avatar models. The mean and standard deviation of fat (mean =2.86, sd = 1.875) and slim (mean =2.63, sd= 18.63) were not enough to produce a significant effect in ANOVA. Hence, one-way ANOVA demonstrated that the effect of body size is not significant (at the 1% level) for average intensity deference on emotions (F (1, 1407= 3.446 Sig = 0.064)). As visually expresses in Fig.5, the average intensity level of 'anger' is notably different as 2.62 for fat avatar models whereas 2.38 for slim avatar models. Similarly, average intensity level of other emotions also revealed a slight difference according to the body size; for instance for the 'interest' emotion, the intensity levels for fat and slim were 1.89, 2.19 respectively.

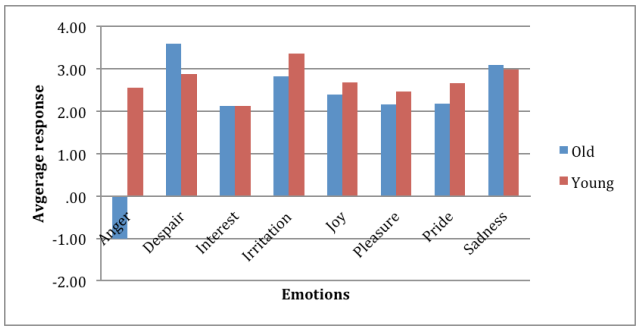

Fig. 6 Average response towards different emotion models: Old / Young

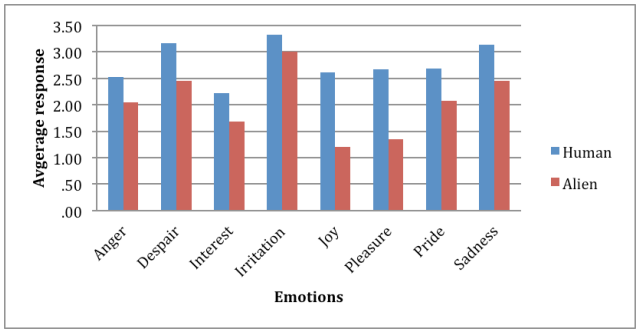

Fig. 7 Average response towards different emotion models: Human / Alien

Fig.6 illustrates the comparison of emotion expression intensity levels between old and young avatar models. The mean and standard deviation of old (mean =2.50, sd: 2.002) and young (mean =2.72, sd: 1.831) were not enough to produce a significant effect in ANOVA. Hence, one-way ANOVA demonstrated that the effect of age is not significant (at the 1% level) for average intensity deference on emotions (F (1, 1407= 3.043 Sig. = 0.081)). As visually expresses in Fig.6, the average intensity level of 'anger' is notably different as -1.00 for old avatar models whereas 2.55 for young avatar models. Similarly, average intensity level of other emotions also revealed a considerable difference according to the age; for instance for the 'irritation' emotion, the intensity levels for old and young were 1.82, 3.36 respectively.

Fig.7 depicts the comparison of emotion expression intensity levels between human and alien avatar models. The mean and standard deviation of human (mean =2.79, sd: 1.858) and alien (mean =2.15, sd: 1.823) were versatile enough to produce a significant effect in ANOVA. Therefore, one-way ANOVA demonstrated that the effect of humanoid is significant (at the 1% level) for average intensity deference on emotions (F (1, 1407= 24.366 Sig= 0.000)). As visually illustrated in Fig.7, the average intensity level of 'anger' is recognizably different as 2.52 for human avatar models whereas 2.05 for alien avatar models. Similarly, average intensity level of other emotions also showed a difference according to the humanoid; for instance, for the 'interest' emotion, the intensity levels for female and male were 2.22, 1.68 respectively.

The observable difference in average emotional intensity levels between male and female models implies that the gender of avatar effect on expression of emotions, thereby, affects the effectiveness of communication.

3. Conclusions

This study experimented about body gesture based emotion expressions of avatar systems in virtual environment with an aim to explain how users perceive the differences of emotional body gestures among avatar models of different age, gender, body size and humanoid. Study tested eight types of emotions expressed by four human and one alien Avatar models through body gestures.

Emotion recognition was tested through thirty-one subjects and average emotion recognition rate was 92%. Emotion expression is a combined process of body gestures and facial expressions. In a situation, where only tested body gestures, the rate of 92% was versatile enough to proceed to the next level yet future research works, the test can be improved further by adding facial expressions to emotion, which is highly recommended.

The observable difference in average emotional intensity levels implies that the age, gender, body size and humanoid of avatar effect on expression of emotions, especially the difference of gender and humanoid is highly significant in the results revealed. This study strongly emphasizes the difference of emotion expression of avatar models based on age, gender, body size and humanoid, thereby, contribute to future development of avatar models through incorporating distinct human emotions and make virtual agents more communicative in their applications in virtual domains.

References

[1] K. Schindler, L.V. Gool, V.de Gelder, "Recognizing emotions expressed by body pose: A biologically inspired neural model", Vol.21, 2008, pp.1238-1246.

[2] G. Cridaikis, A. Raouzaiou, F. Ingenelser, K. Karpouzis, E. Andre "cross-cultural, multimodal, affective corpus for gesture expressivity analysis", Vol.7, 2012, pp. 121-134.

[3] C. Peters, G. Cstellano, M. Rehm, , G. Andre, A. Raouzaiou, K. Rapantzikos, K. Karpouzis, G. Volpe, A. Camurri, A. Vasalou, " Fundamentals of Agent Perception and Attention Modelling" , Emotion-Oriented Systems, Cognitive Technologies , 2011, pp. 293-319
[4] K. Karpouzis, A.Raouzaiou, S. Kollias, "Moving Avatars: Emotion Synthesis in Virtual World", Image, Video and Multimedia Systems Laboratory, National Technical University of Athens.
[5] P. Weerasinghe, A. Marasinghe, R. Ranaweera, S. Amarakeerthi and M. Cohen, "Emotion Expression For Affective Social Communication," International Journal of Affective Engineering, Vol.13 (2014) No.4 Special Issue on 2013 ICBAKE, Part 1 p.261-268
[6] M. Quiros- Ramirez, S. Polikovsky, Y. Kameda, T. Onisawa, "A Spontaneous Cross-Cultural Emotion Database: Latin-America vs. Japan", International Conference on Kansei Engineering and Emotion Research, 2014 pp. 1127-1133
[7] M. Coulson, "Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence", Journal of Nonverbal Behavior, Vol.28, 2004, pp. 117- 133.
[8] T. Usui, K. Kume, M. Yamano and M. Hashimoto "A Robotic KANSEI Communication System Based on Emotional Synchronization", International Conference on Intelligent Robots and Systems, France, Sept, 22-26, 2008 pp. 3344-3349.

Prabath Weerasinghe is a Doctoral Student in the Department of Management and Information Systems Engineering, Nagaoka University of Technology, Nagaoka, Japan. He received a master in Computer Science & Engineering from the University of Aizu, Japan. He received a B.Sc. degree in Computer Science and Mathematics from the University of Colombo, Sri Lanka. He has worked as a Teaching Assistant, The Graduate Department of Information Systems**,** The University of Aizu, Fukushima, Japan. He has worked as a system engineer at NS Service CO., LTD. in JAPAN from 2007-2010. He was a Research Assistant, The Graduate Department of Information Systems**,** The University of Aizu, Fukushima, Japan. His current research interests include mobile Computing, HCI and empowering collaborative virtual environments with emotions.

R. P. C. Janaka Rajapaksha is an Associate Professor at Graduate Institute of Animation and Film Art, Tainan National University of Arts, Taiwan. And he is also a visiting researcher in Department of Media and Image Technology at Faculty of Engineering, Tokyo Polytechnic University. He received his Ph.D. in Knowledge System Science from Japan Advanced Institute of Science and Technology in 2008. His research interests include computer graphics, computer animations, haptic interface, virtual reality, kansei engineering, and biometrics. He is a member of the IEEE, IEICE and SIG-Design Creativity

Ashu Marasinghe is an Associate Professor in the Department of Management and Information Systems Engineering, Nagaoka University of Technology, Nagaoka, Japan and visiting Professor at International University of Japan. Prof Ashu obtained his B.Sc. degree in Physics and Mathematics from the University of Colombo, Sri Lanka, M.Sc. in Computer Science & Engineering from the University of Aizu, Japan and Ph.D. in Computer Science & Engineering from the University of Aizu in 1997, 2001, and 2004 respectively. He has worked in the University of Aizu, Japan. Prof Ashu`s research interests are: Engineering Management, Construction Management, Intelligent Control System, Multi Agent System, Ubiquitous System, Social Informatics, and Kansei Engineering. Prof. Ashu served as a General Chair in several International Conferences include ICBAKE 2009, HC-2008, KEAS`08, CDG`08, KEIS`06