Emotional Avatar Animation on Smart Phone SMS

Hae Won Byun* and Jung Suk Lee**

*School of Information Technology, Sungshin Women's University**
*Incross Company***
hyewon@sungshin.ac.kr

Abstract

Recently people do not only use mobile phone to call, but also use it to send a SMS. However, it is difficult to express own complicated emotion with text and emoticon of exited SMS service. We pay attention to express user's emotion interesting and correct, we use character animation. This paper suggests emotion based gesture animation generation system that uses character's facial expression and gesture to delivery emotion excitably and clearly. Michael[1] investigated interview of two people who has stylized gesture. They suggested gesture generation graph for stylized gesture animation. In this paper, we focus to analyze and extract emotional gestures of Disney animation characters and create 3D models of extracted emotional gestures. To express emotion of person, we use emotion gesture generation graph that import the emotion flow graph that expresses emotion flow for probability. We investigated user reaction for research proprieties of suggested system and the alternation propriety.

Keywords: *Mobile phone 3D animation, Emotional gesture generation, Emotion flow, Gesture profile, Mobile phone 3D engine*

1. Introduction

Now mobile phone is more than a creature of high-technology for people. Mobile phone makes people connecting to communicate with other person. People do not only use mobile phone to call, but also use it to send a SMS. However, SMS is formed of only text and emoticon, so it has limitation to express atmosphere and nuance. In modern societies, expressing individual character is very important; therefore the desire to express emotion became stronger. But, current SMS services standard and limited tools to express detail of emotion perfectly. In this paper, we suggest using character animation for SMS service to communicate user's idea dynamical and correctly. We tried to create character animation that express emotion by facial expression and body gesture. Michael et al suggest character animation recreation system that annotated interview video to build stylized gesture animation such as standard gesture of an announcer.

Michael et al focused on characteristic gesture; on the other hand we focused on that SMS service is more used to expressing emotion than to send informative contents, we designed system that generates expressing emotion gesture animation to notify user's emotion at a glance. We annotated and analysis character's gesture of Disney animation to use exaggerated gesture. Based on this analysis, we recreated gesture animation that represents SMS message. Experimental results demonstrate the effectives of our approaches.

*Corresponding author: hyewon@sungshin.ac.kr

2. Related Work

To generate and animate facial animation from an input of speech track is recent research endeavor. Pearce and Parke developed 3D speech animation system based on early suggested parameter facial expression models [2]. It provided way to change voice signal in law to control parameter sequence for generating animation.

Water et al developed 3D speech animation system based on test, and named it DECface [4]. Kalberer et al suggested similar system, but they focused on more realistic speech animation; therefore they used 3D digitizer to capture the talking head. These researches focused on lip sync.

Recent research about gesture animation based on recreating gesture for input text. Cassell et al developed the Behavior Expression Animation Toolkit. In this system, gestures are generated using hand-made rules and are selected using priority values [7].

However, it is destitute of speaker's style and hand-made rules are subjective to make the system automatic. To overcome this limitation, Micheal et al triggered gestures probabilistically and planed both the gestures' internal structure (phases, timing, shape). Also they annotated interview of two people to get stylized gesture of speaker and used graph structure to generate animation based on new text.

In contract with computer system, mobile phone system should be fast and not heavy. Vrabec et al suggest parallel system to generate animation for real-time. Server system and mobile phone system handle animation data at the same time. It is possible to generate animation for input text faster. Emura et al used all kind of data such as GPS, settlement and weather to generate animation. This system can animate realistic animation and animate abundant information.

3. System Overview

We suggest the gesture generation system to express emotion of SMS messages as character gesture and facial expression animation and to service on a mobile phone. The system is constructed of emotion recognition, gesture profile, emotion gesture generation and mobile phone engine (see Figure 1). When user put the text to mobile phone, the system recognizes emotion of text at server and generates final gesture and facial expression based on recognized emotion. Generated data from server system is transmitted to mobile phone 3D engine and play animation on the mobile phone.

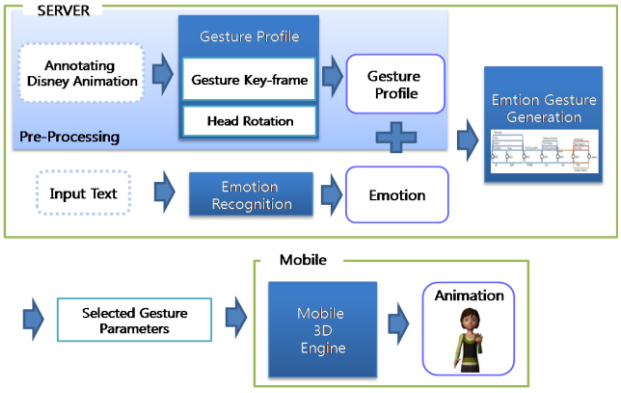

Figure 1. System Architecture

Emotion Recognition System: This system recognizes 8 emotions (neutral, happy, sad, angry, surprise, afraid, disgust) from English text. We used open source Synesketch that recognize emotion from text and animate colors and shapes represented the recognized emotion. We defined emotion words that decided emotion of entire sentence from Disney animation and add emotion recognition system's database.

Gesture Profile: Gesture profile is gestures data classified by noun and emotion to generate animation that express emotion of text. We analysis and annotate Disney animation's gesture and extract gestures that has each emotion. Also head rotation, feature of animation and time information are saved to gesture profile for animation generation.

Emotion Gesture Generation: Emotion gesture generation proceed by emotion gesture generation graph that selects final gestures to animate from gestures that transmitted form gesture profile. Create node with inputted text's noun and make link with final gesture classified by emotion. For the last, linked gestures are priority sorted by emotion that emotion recognition system recognized.

Mobile Phone 3D Engine: The role of mobile phone 3D engine is to create 3D character animation based on final gesture that selected by emotion gesture generation graph. Final selected gesture parameters are transmitted to mobile phone 3D engine and do the blending with facial animation. Finally, creates emotion based character animation and service on the mobile phone.

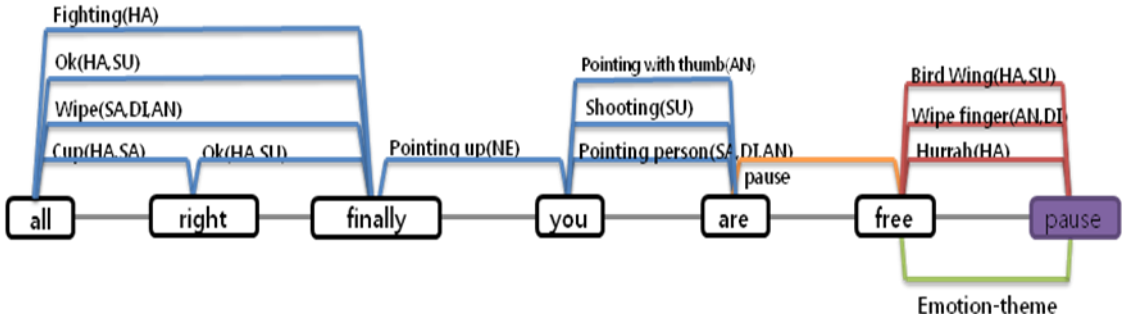

Figure 1. Emotion Gesture Generation Graph

4. Gesture Profile

Our system made analysis of exaggerated gesture of Disney animation to get characterized gesture. We annotate gestures that represent emotion and classify by noun, time and emotion.

Cassell et al extracted gestures from iconic sentences (this, there, you, I). In contrast, Michel et al approach iconic and also metaphor (number, long, all, open). They expanded range of nouns that can be extracted and generate animation similar with real communication. In this paper, we also use iconic, metaphor and add emotion (gloomy, fighting, upset) to express emotion with gesture more efficiently. Extracted gestures are classified by noun, emotion, time and add head rotation, time information to generate the animation (see Figure 3). After these procedures, we save the data as gesture profile.

NOUN	TIME	EMOTION GESTURE	HEDA ROTATION
'yes'	00.02.00 00.01.60	Shoulder down(Sad) Figthing(Happy)	Head noding (Neutral)
'ok'	00.01.30	Ok(Happy, Neutral)	
'sure'	00.02.22 00.02.80	Cup(Happy, Neutral) Shoulder down(Sad)	Head noding
'you'	00.00.98 00.01.00 00.01.00	Point with thumb(Anger) Shooting(Surprise, Happy) Pointing you(Neutral, Sad, Disgust, Anger)	Forward person
'long', 'far'	00.20.15	Hand on eyebrows (Neutral)	

Figure 3. Gesture Profile

5. Emotion Gesture Generation

Based on recognized emotion and gestures of text extracted by gesture profile, we use graph structure (see Figure 2). We express text noun as node and extracted gestures as arcs. We also consider idioms, so we unify arcs for idioms.

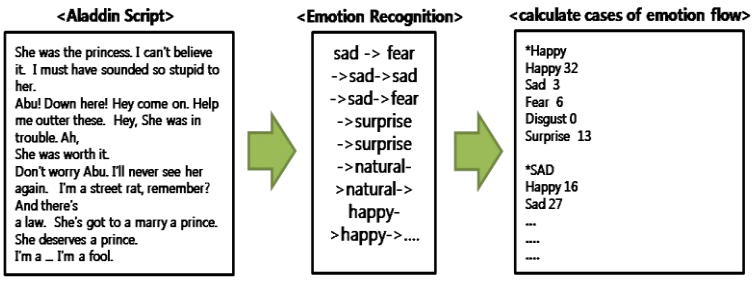

Figure 4. Cases of Emotion Flow

However there can be a node does not have gesture arc represents emotion that extracted for whole sentence. For example, 'free' is key noun to recognize 'Happy' emotion of sentence, but 'you' and 'finally' nodes do not have gesture arcs matched with 'Happy' emotion. We define and import 'Emotion Flow graph' to choose gesture arcs that has other emotion. 'Emotion Flow Graph' is based on theory that emotion flows with certain rules [8] and it composed of probability.

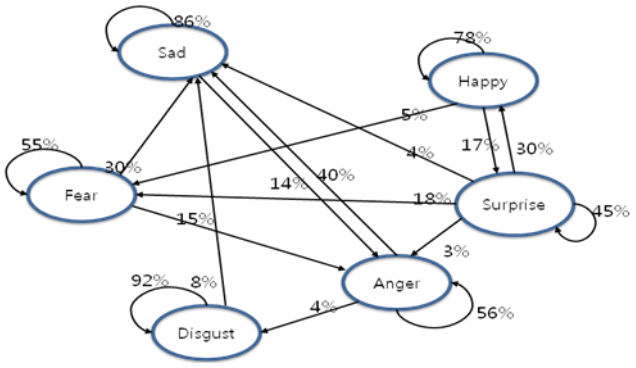

Figure 5. Emotion Flow Graph

To define emotion flow based on character's characteristic; recognize emotion form Disney Aladdin script and calculate cases that emotion transits to other emotion (see Figure 5). Through importing defined emotion flow graph, it is possible to priority sort gesture arcs even though the node does not have gesture arc for extracted emotion.

6. Mobile Phone 3D Engine

The role of mobile phone 3D engine is to create the final animation by transmitting final gesture data from emotion gesture generation system at server and composing with facial animation. To generate 3D character animation, do modeling gesture key-frames that transmitted from gesture profile by MAYA to OBJ files. For real-time playing at the mobile phone, we do modeling low polygon models.

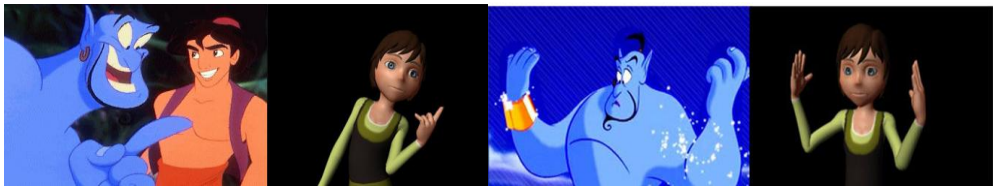

Figure 6. Disney Animation 3D Modeling

During the models, we load the models chosen by emotion gesture graph to mobile phone engine and compose with facial animation. Arrange a low with key-frame and blending key-frame to generate animation. We use Google Android as emulator and our system is manufactured by java and OpenglES.

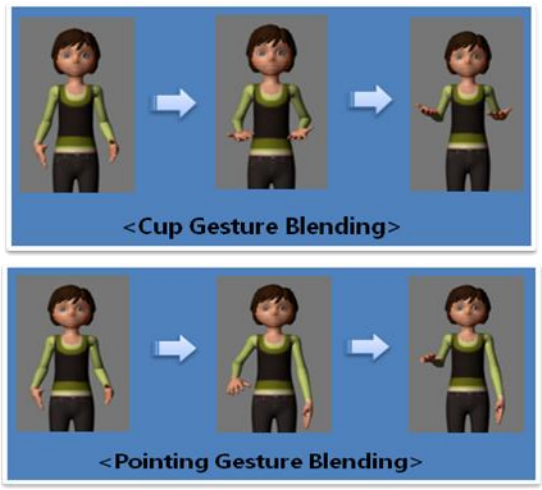

Figure 7. Blending the Gestures

7. Experiments

We did experiment to prove our system's feasibility, user satisfaction and alternation propriety with objective validity. We separate group to people in their twenties and people in their forties for contrasting user ability for mobile phone. After they experiment our system, they should answer the question that ask satisfaction, interest, feasibility and check their emotion is recognized well or not.

User can test the system for 50 times, we research the number of times that gestures are generated agreeable to inputted text and whether express user's emotion well or not.

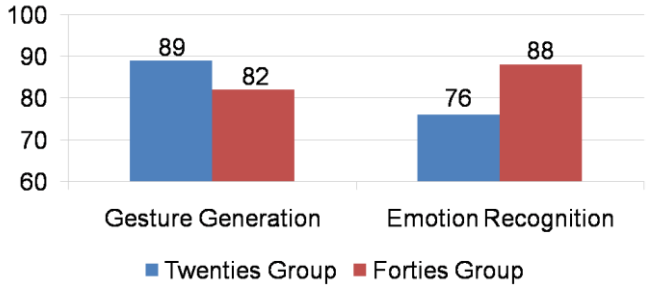

Figure 8. Result of Survey for Gesture Generation and Emotion Recognition

You can the result of experiment at Figure 7, by and large; it shows that result of our system is well matched with user's purpose. Twenties group answered result animation is generated well for 89%. It's higher than 82% that answer of forties group. This fact tells us exaggerated animation is more right for people in their twenties SMS substance.

Percentage of well recognition of emotion comes out 76% for people in their twenties and 88% for forties group. It shows that emotion of forties group simpler in general, so it is easier to recognize.

Another survey, we ask user to evaluate 1 to 5 point for user's satisfaction, interest and alternation propriety. Figure 9 represents average point for the answer and it shows that high scores for every clauses. This result proves that our system furnishes the high satisfaction and interest. Also our system is appropriate to alternate with current SMS service.

Figure 9. Result of Survey for Satisfaction, Interest, and Alteration propriety

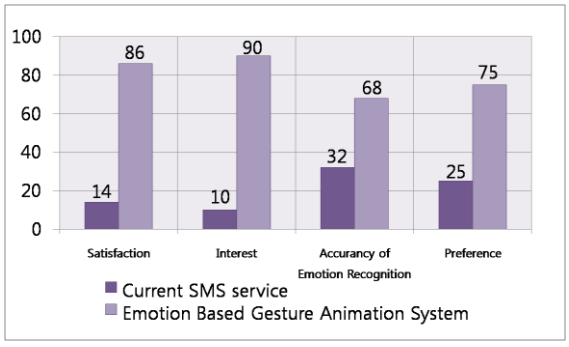

Figure 10. Result of Compare Survey

To contrast with current SMS service, we make 50 people compare our system and current SMS service for satisfaction, accuracy of emotion recognition and preference. The result, Figure 10, shows that our system's preference is higher than current system. For satisfaction and interest, 90% for our system shows difference with 10% for existed system. Besides for accuracy of emotion recognition comes out 68% which is higher than exited system

8. Conclusion

This work presents a system for generating gesture expressing emotion for mobile phone SMS service. We use Disney animation to extract exaggerate gesture for noun, emotion. We organized data as gesture profile and use graph structure to select gesture based on recognized emotion.

To select the best suitable gesture for emotion, we define and import 'Emotion Flow Graph' through Disney animation character's emotion. Selected gesture's parameters are sanded to mobile phone and are blended with facial animation to generate animation. We concentrate to enhance user's satisfaction and interest by providing more realistic and emotional conversation animation.

Figure 11. Emulate at Google Andriod

Acknowledgements

This work was supported by the Sungshin Women's University Un-Jeong Global Project research Grant of 2012.

References

[1] N. Michel, K. Michel, A. Irene and S. Hans-peter, "Gesture Modeling and Animation Based on a Probabilistic Re-Creation of Speaker Style", ACM Transactions on Graphics, vol. 27, **(2008)**.

[2] Computer generated animation of faces. In Proceedings of the ACM Annual Conference, ACM Press.

[3] D. W. Massaro, M. M. Cohen, "Visible Speech and Its Potential Value for Speech Training for Hearing-Impaired Perceivers", Bibliographic reference, In Still, **(1998)**, pp. 171-174.

[4] G. A. Kalberer, J. Luc and G. Van, "Realistic face animation for speech", Journal of Visualization and Computer Animation, **(2002)**.

[5] A. Garcia-rojas, F. Vexo, D. Thalmann, A. Raouzaiou, K. Karpouzis and S. Kollias, "Emotional Body Expression Parameters In Virtual Human Ontology", IEEE org, **(2005)**.

[6] J. Cassell, H. Vilhj Almsson and T. Bickmore, "BEAT: The Behavior Expression Animation Toolkit", In Proceedings of SIGGRAPH, **(2006)**, pp. 477-486.

[7] C. L. Lisetti and P. Gmytrasiewicz, "Can a Rational Agent Afford to Be Affectless?", Applied Artificial Intelligence, vol. 16, no. 7-8, **(2002)**.

Authors

Hae Won Byun

She received the BS degree in computer science from Yonsei University, Korea, in 1990 and MS and PhD degrees in computer science from Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1992 and 2004, respectively. She is currently a professor in the School of Information Technology at Sungshin Women's University, Seoul, Korea. She's primary research interests are facial motion capture and expression retargeting, performance-based animation, non-photorealistic rendering, and texture synthesis.

Jung Suk Lee

She received the BS and MS degree in computer science from Sungshin Women’s University, Korea, in 2007 and 2009. She is currently working at Incross company. Her research interests include facial animation, gesture animation, and mobile phone service.