

# Introduction of Face Expression into Dialog System for A Sense of Affinity

Hiroki Midorikawa<sup>1,\*</sup>, Takatsugu Suzaki<sup>1</sup>, Masahiro Kawamura<sup>1</sup> and Masayuki Numao<sup>1</sup>

<sup>1</sup>The University of Electro-Communications 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

## Abstract

For dialogue system to be accepted into people's lives, it is important to give people sense of affinity to them. In order to do that, we introduce face expression into dialogue system and suggest the architecture of face expression generation. Face expression express emotion, personality, and wiling. It can reduce inorganic and mechanical impression of systems and make affinity. In this architecture, emotion which is estimated or specified in scenario file is converted to face expression. Emotion is expressed two dimensional number based on Russell's circumplex model. Therefore, this is easily applied to other emotional actions such as voice tone. Finally we implemented the system and integrated to the dialog robot "WANCO". We evaluated that the user's affinity is enhanced by adding facial image.

## Keywords

dialog system, facial expression, emotion recognition

## 1. Introduction

For AI such as dialogue system to be accepted as a human partner, it is essential to make people feel familiar with AI. Especially, in areas such as health care, it is important to motivate users to use the system frequently. One way to give users a sense of affinity to systems is to express emotions. This would visually impress users the personality of the system, and motivates them to talk to the system. In this study, we will improve the sense of affinity felt by users by implementing facial graphics that change face expressions in the interface of the spoken dialogue system. Furthermore, we will suggest architecture of face expression generation. To connect dialogue manager system to face expression, we also consider a markup language to write emotional action. The architecture of the dialogue system is figure1. This is a rule-based system in which the content of dialogue is described in advance as a scenario file.

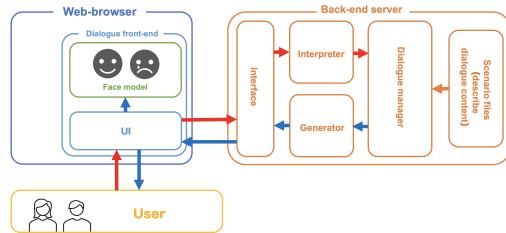


Figure 1: System architecture

AAAI 2023 Spring Symposia, Socially Responsible AI for Well-being, March 27-29, 2023, USA

\*Corresponding author.

hiroki.midorikawa@uec.ac.jp (H. Midorikawa)

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

To map responses generated by systems to facial expressions, the developer can either write the correspondence between facial expressions and speech in advance or automatically estimate emotions from text of speech. In this study, we examine two methods: one is for the developer to specify facial expressions in advance in rule-based dialogues, and the other is to create an emotion dictionary to estimate rough emotions in order to generate facial expressions from words in the text.

## 2. Socially Responsible AI for Well-being

Familiarity with AI integrates it into human society and gives it a role as a component of society. This is very important for the success of AI as more than just a tool.

## 3. Russell's circumplex model

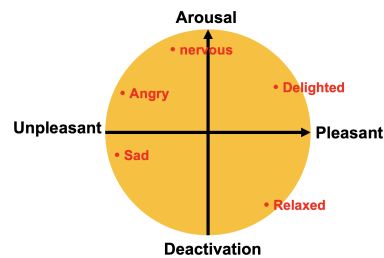


Figure 2: Rusell's circumplex model

Russell's circumplex model represents emotions in a circular model with two axes, as shown in the figure2.

The horizontal axis represents pleasantness and unpleasantness, and the vertical axis represents arousal level. The red dots show examples of typical emotions. In the proposed method, emotions are specified by 2D coordinates in Russell's circumplex model.

## 4. Face-expression Generation

Face expressions express emotions, and therefore emotions can be converted to them. It should be possible to estimate suitable face from conversation text. However, using emotion to decide face expression make some advantages. First, the emotional state of system becomes clear. Furthermore, the emotions are able to be expressed other means, such as tone of voice, gestures, and wording. The implementation of face model can be changed to others. When we decide the mapping between emotions and them, the system performs multiple emotional expression consistently. On the other hand, the mapping must be obtained by other means. In the current implementation, emotional coordinates are mapped to facial movements, e.g., eyebrow and mouth angle positions, eyelid openings, etc., for continuous facial expression generation.

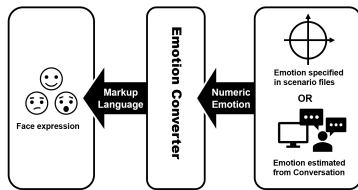


Figure 3: flow of face-expression generation

As figure3, emotions estimated or specified in scenario files is holded in emotion manager. And then it is passed to face-expression interface as EmojiML. EmojiML describes emotion and other emotional actions. The example of a tag embedded EmojiML to make system to say "Thank you" with pleasure is as below.

```
<f-exp type="pleasure">Hello</f-exp>
```

Examples of face expression generation is figure4.

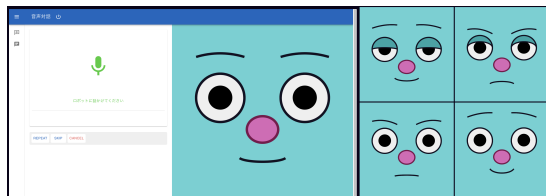


Figure 4: Display of dialogue system (left) and examples of face ganarated by the system (right)

## 5. exploratory experiment

Is the face-expression natural?			Did you feel affinity with the system?		
Face expression	not available	available	Face expression	not available	available
very natural	1	2	Strongly agree	2	3
natural	4	3	agree	1	2
Neither	0	0	Neither	2	0
unnatural	0	0	disagree	0	0
very unnatural	0	0	Strongly disagree	0	0

Table 1

Number of responses for each option in the survey

An experiment was conducted to determine whether the addition of facial expressions to the interaction with the system could create a sense of affinity. Two faces with and without facial expressions are displayed on the left and right sides of the screen, and several users are asked to talk with the system. After the dialogue, the users were asked whether their facial expressions were natural and they felt a sense of affinity with the system. The results in table1 show that both were improved with facial expressions. However, none of them received a low rating. This may be because of not strongly emotional scenario, so the presence or absence of facial expressions did not make much unnaturalness.

## 6. Issues of face-expression and emotion estimation

The results suggest that showing facial expressions gives users a positive impression. The consideration of correspondence between facial expressions and emotions is insufficient. In order to make natural face, find suitable method of generating face asking users opinion.

## References

- [1] Yuki M et al. [ Interactive Expression Expression of Sensory Robot Using Simple Recurrent Network] Simple Recurrent Network wo mochiita kansei robotto no intarakuthibu hyoujou hyousyutsu (in Japanese). Nihon robot gakkai. 2010; 28(3): 503- 512.
- [2] De Carolis, Berardina, et al. "APML, a markup language for believable behavior generation." Life-like characters: tools, affective functions, and applications (2004): 65-85. .

## Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP20H04289 "Functional Independence Measurement System based on ADL Ontology for Aged Person"