庞大表情和情绪推导模型

Sachiko KITAZAKI and Takehisa ONISAWA

University of Tsukuba
Institute of Engineering Mechanics
1-1-1, Tennodai, Tsukuba, Ibaraki 305, Japan
phone:+81-298-53-6188, fax:+81-298-53-6471
E-mail:kita@fhuman.esys.tsukuba.ac.jp

Key Words: neural networks, facial expressions, emotions

Abstract: This paper mentions an inference model of facial expressions and emotion from the viewpoint that non-verbal information as well as verbal information is useful in human computer interaction, and that especially, facial expressions are useful since they reflect human emotion well. First this paper considers the inference model of human emotion from only facial expressions. Next this paper considers the inference model of human emotion from not only facial expressions but also the situation in which human is put. Neural network models are constructed based on the data obtained by the questionnaire about facial expressions, emotions, and situations. It is found by the evaluation questionnaire that obtained models infer human emotions from facial expressions and situations well.

1. Introduction

In human face to face communication, human comprehends emotional expressions, gesture, i.e., non-verbal information[1], as well as verbal information. Human employs the information usefully in order to communicate each other smoothly. Even in human computer interaction, non-verbal information as well as verbal information is useful. Especially, facial expressions are useful since they reflect human emotion well. Then understanding of human facial expressions becomes more important. There have been many approaches to construct the relationship model between facial expressions and human emotions, e.g., a neural network model approach[2][3], a fuzzy measure and fuzzy integral approach[4]. These are so called soft computing approaches. Of course, there have been approaches from the psychological point of view[5]. In many of these approaches, however, only the relationship between facial expressions and human emotion is considered. Understanding of human facial expressions is also dependent on the situation in which human is put. The situation becomes an important factor for the inference of human facial expressions and emotion.

First this paper considers the inference model of human emotion from only facial expressions. Neural network models are constructed based on the data obtained by the questionnaire. Next this paper considers an inference model of human emotion from not only facial expressions but also the situation in which human is put. Neural network models, which infer emotion from facial expressions and the situation, are also constructed by the data obtained by the questionnaire. In both models inference results of human emotion are expressed by natural language.

Finally, this paper verifies by the evaluation questionnaire that obtained models can infer human emotion not only from facial expressions but also the situation in which human is put.

2. Inference Model of Emotion from Only Facial Expressions

2.1 Model Structure

This paper employs a neural network model in order to infer emotion from facial expressions. Eight parameters of facial expressions, i.e., the size of eyes, the slant of eyes, the shape of eyes, the slant of eyebrows, the shape of eyebrows, the size of a mouth, the width of a mouth, and the shape of a mouth, are considered as inputs of the neural network model. On the other hand 6 kinds of emotions, i.e., happiness, sadness, fear, surprise, anger and disgust, are considered as outputs. The output of the model is the evaluation of one of 6 kinds of emotions or the evaluation of unnaturalness. The unnaturalness means that the subject has a feeling of unnaturalness for the presented facial expressions.

2.2 Questionnaire about Facial Expression

In the questionnaire about facial expressions, drawings of facial expressions as shown in Fig. 1 are used since they are controlled more easily than real expressions such as a face photograph. For the questionnaire 150 facial expressions are prepared. Ten subjects, who are undergraduate or graduate students, and eight male and two female students, are asked to compare the facial expressions at the right side of the arrow with the expressions at the left side in Fig.1, and to infer emotion.
with three scale evaluations, i.e., a weak feeling, a feeling, or a strong feeling of happiness, sadness, fear, surprise, anger and disgust. The left side expressions are standard expressions. If the subjects do not have any feeling of emotion, they are asked not to mark the questionnaire sheet. And if the subjects have a feeling that the facial expressions in the questionnaire sheet are unnatural, they are asked to mark unnatural in the sheet. In order to avoid the influence of the former expressions, a couple of facial expressions is drawn in one questionnaire sheet.

2.3 Data Analysis

Parameters of the size of eyes, the size of a mouth, the slant of eyes, the shape of eyes, the shape of eyebrows, the width of a mouth and the shape of a mouth have numerical values in [0, 1]. These values are dependent on the location of the part in facial expressions as shown in Fig. 2. With respect to emotion, natural language expressions correspond to the numerical value in [0, 1] as follows.

\[
\begin{align*}
\text{a weak feeling} & : 0.33 \\
\text{a feeling} & : 0.67 \\
\text{a strong feeling} & : 1.00
\end{align*}
\]  

(1)

If unnatural in the questionnaire sheet is marked, the numerical value about the unnaturalness is set up at 1. If it is not marked, the value is set up at 0.

Seven kinds of neural network models are obtained by the use of the backpropagation algorithm. The obtained models have an input layer, an output layer and two hidden layers. Each model has 8 input nodes and an output node about emotion or the unnaturalness of facial expressions. The numerical value in [0, 1] representing the output value of each emotion is transformed into natural language expressions as shown in Fig. 3. If three or more unpleasant emotions out of the fours (anger, disgust, sadness, fear) are felt even a little, natural language expression a feeling of unpleasant is employed in order to express unpleasant emotions together and the of unpleasantness is the largest output value grade in unpleasant emotions. However if one unpleasant emotion is much stronger than the other unpleasant emotions, the grade of each emotion is represented by natural language.
of 6 kinds of facial expressions which are used in the first questionnaire and 18 kinds of expressions which are not used in the first questionnaire. The models get the point according to their evaluation as shown in Table 1.

### Table 1 Evaluation Point

<table>
<thead>
<tr>
<th>Evaluation</th>
<th>Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>The subject does not agree with the model inference results at all.</td>
<td>-2</td>
</tr>
<tr>
<td>The subject does not agree with the model inference results.</td>
<td>-1</td>
</tr>
<tr>
<td>The subject does not conclude whether he/she agrees with the model inference results or not.</td>
<td>0</td>
</tr>
<tr>
<td>The subject agrees with the model inference results.</td>
<td>+1</td>
</tr>
<tr>
<td>The subject agrees with the model inference results completely.</td>
<td>+2</td>
</tr>
</tbody>
</table>

With respect to the individual model, the average point of evaluation among 24 facial expressions is 1.13. The best average evaluation point is 1.89. The worst average point is 0.11. On the other hand with respect to the average model, the average point of evaluation is 1.15. The best average point is 1.78 and the worst average point is 0.44. The obtained models get a high evaluation point. No subject does not agree with the inference results of models at all. The use of the expression unpleasant causes high evaluation point. There is no difference between the inference results by the individual subject model and the ones by the average model.

### 3. Inference Model of Emotion from Facial Expressions and Situation

#### 3.1 Model Structure

Inference of emotion is dependent on not only facial expressions but also the situation in which human is put. In constructing the inference model, information on the situation is as well as facial expressions is necessary.

With respect to information on facial expressions, the same parameters in the inference model of emotion from only facial expressions are employed. Moreover, 18 situations are considered. In this paper, however, information on the situation is input into the model not directly but indirectly since it is impossible to consider every situation as the input of the neural network model. That is, the information is transformed into emotion.

For example, let us consider the situation where I have passed an entrance examination unexpectedly. The subject has a strong feeling of happiness and surprise and a weak feeling of fear for the situation, the situation is transformed into a strong feeling of happiness and surprise, and a weak feeling of fear.

The model has 8 input nodes about facial expressions and 6 input nodes about emotion obtained from only the situation. The model has an output node about one of 6 emotions (happiness, surprise, anger, disgust, sadness, fear) or an output node of unnaturalness. The unnaturalness means that the subject has a feeling of unnaturalness for the combination of facial expressions and the situation.

#### 3.2 Questionnaire about Situations

In order to make the relationship between the situation and the emotion which is felt under the presented situation, the questionnaire about situations is performed to the subjects. The subjects are asked to mark the questionnaire sheet according to the extent of the feeling of 6 kinds of emotions under the presented situation.

#### 3.3 Questionnaire about Facial Expressions and Situations

In this questionnaire, drawings of facial expressions are employed from the same reason as mentioned before. The subjects are asked to compare facial expressions at the right side of the arrow with the left side expressions, and to infer emotion by considering the presented situation with three scale evaluations, i.e., a weak feeling, a feeling or a strong feeling of happiness, sadness, fear, surprise, anger and disgust as shown in Fig.5. If the subjects evaluate that the combination of the situation and facial expressions is unnatural, they are asked to mark unnatural in the sheet. In the questionnaire 24 facial expressions and 7 situations are combined, where 24 facial expressions are selected out of 150 facial expressions used in Section 2, and 7 situations are selected out of 18 situations considered in 3.2. In order to avoid the influence of the former combination in the questionnaire, one combination of facial expressions and the situation is represented in one questionnaire sheet.

![Fig.5 Questionnaire about Facial Expressions and Situation](image-url)
### 3.4 Data Analysis

Eight parameters about facial expressions have numerical values in \([0, 1]\). With respect to emotion, natural language expressions also correspond to the numerical value in \([0, 1]\) as shown in Eq. (1). If unnatural in the questionnaire sheet is marked, the numerical value about the unnaturalness is set up at 1. If it is not marked, the value is set up at 0. Seven kinds of neural network models are obtained by the backpropagation algorithm. Each model has an input layer and an output layer, and two hidden layers. The model has 20 nodes in each hidden layer. The output of the neural network model is expressed by natural language in the same way as the inference model from only facial expressions.

### 3.5 Results

Only inference model of individual subject emotion is considered. Fig. 6 shows an example of inference results when the facial expressions and the situation are given as shown in Fig.6.

<table>
<thead>
<tr>
<th>facial expressions</th>
<th>Inference results</th>
</tr>
</thead>
<tbody>
<tr>
<td>He has a weak feeling of unpleasantness.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>facial expressions</th>
<th>Inference results</th>
</tr>
</thead>
<tbody>
<tr>
<td>He got bad marks in the exam.</td>
<td></td>
</tr>
<tr>
<td>He has a feeling of unpleasantness.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>facial expressions</th>
<th>Inference results</th>
</tr>
</thead>
<tbody>
<tr>
<td>He is going to see the pass announce.</td>
<td></td>
</tr>
<tr>
<td>He has a feeling of fear.</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 6 Example of Inference Results

In order to verify that the obtained models can infer human emotion from facial expressions and the presented situation satisfactorily, the ten subjects are also asked to compare model inference results of emotion with their inference results by the use of 25 combinations of facial expressions and situations. Of which 2 facial expressions and 2 situations are used in the questionnaire and 3 facial expressions and 3 situations are not used in the questionnaire. The models get the point according to subjects evaluation as mentioned before. The average point of evaluation among 25 combinations is 1.14. The best average evaluation point is 1.71 and the worst average point is 0.14.

The obtained models have as a high evaluation point as the inference model from only facial expressions. It is confirmed that the model can infer emotion well according to the situation even if the same facial expressions are given.

### 4. Conclusions

This paper considers an inference model of human emotion from not only facial expressions but also the situation in which human is put from the viewpoint that inference of human emotion is dependent on the situation as well as facial expressions. Based on the data obtained by the questionnaire about human emotions, facial expressions and situations, neural network models are obtained. It is found by the evaluation questionnaire that the obtained models can infer human emotion well from not only facial expressions but also the situation in which human is put.

This study is useful for the construction of human interface which uses non-verbal information as well as verbal information. In a future the application of the obtained model to the construction of the human interface is considered as an open problem.

### References


