Emotional Boundaries for Choosing Modalities according to the Intensity of Emotion in a Linear Affect-Expression Space

Jeong Woo Park, Hui Sung Lee, Member, IEEE, Su Hun Jo, Min-gyu Kim, and Myung Jin Chung, Senior Member, IEEE

Abstract—Recently, in the field of HRI, multimodal expression has been an issue. Synchronizing modalities and determining what modality to use are important aspect of multimodal expression. For example, when robots express emotional states, they may use only facial expressions or facial expressions with gestures, neck motions, sounds, etc. In this paper, emotional boundaries are proposed for multimodal expression in a three-dimensional affect space. The simultaneous expression of facial expression and gestures was demonstrated using proposed emotional boundaries on a simulator.

I. INTRODUCTION

One of the most important roles of robots as they become more widespread in our daily lives will be to convey social functions, including communicating with humans. If the robot can convey not only information but emotion to humans, as humans do, it will be able to interact more naturally and effectively. Namely, users expect robots to behave like humans, and research to enable communication with humans has been vigorously conducted.

The importance of HRI, wherein human emotion is recognized and analyzed, and the robot’s emotional state is conveyed, with respect to a robot carrying out social functions is widely known on the basis of HHI (Human Human Interaction). In HHI, humans exchange information or emotion using many methods such as language, behavior, facial expression, etc. According to Mehrabian, words account for 7% of liking, tone of voice accounts for 38%, and facial expression, etc. account for 55% [1]. Among the forms of body language, facial expressions and gestures occupy a major proportion. Hence, research on implementing facial expressions and gestures is important for expressing humans’ emotional state [2]. To date, many robots have been developed with a focus on HRI [3], [4], [5], [6], [7]. However, most studies have focused on using facial expressions, whereas humans express their emotional state or emphasizes their intended meaning by using not only facial expressions but also neck motion, gestures, etc. [1]. In these aspects, it is important that robots also have the capacity to display their emotional state through multimodality.

Determining how the various different types of modalities are displayed has to be considered in multimodal expression. Information about the position and orientation of each joint is needed to express posture and velocity is similarly needed to express gestures. In the case of sound or color (with blinking), information about time is needed. If a modality that does not need to be expressed is displayed or the modalities of expression are not synchronized, the human user or participant in the interaction may be confused.

In previous works about multimodal expression, they predefined which modalities are used and when modalities are expressed. Hence, in this paper, we attempt to address problems in multimodal expression by using emotional boundaries in a three-dimensional affect space.

Research on emotional boundaries was initiated by Etcoff [8]. In previous studies, the criterion to divide emotional boundaries was not distinct [3], [8], [9], [10]. Hence, it is difficult to apply these methods to multimodal expression. In addition, facial expressions or gestures were changed discretely, because one emotional boundary may be adjacent to another in some cases [4], [11]. It was also known that it is difficult to apply emotional boundaries to robots, since these boundaries are determined based on psychology and have only been investigated between neutrality and other emotions [12].

In the present work, we propose a systemic approach to address the problems encountered in previous researches. First, we find emotional boundaries using the linear affect-expression space model that was proposed in advance [13], [14]. Next, we propose a general method to express the robot’s emotional state using multimodality based on emotional boundaries. Finally, as an example of multimodal expression, we present experimental results of emotional expressions using a facial expression and gestures simultaneously.
II. A LINEAR DYNAMIC AFFECT-EXPRESSION MODEL

A linear dynamic affect-expression model consists of the linear affect-expression space model and the dynamic emotion model [15]. The linear affect-expression space model is proposed to express robot’s emotional state effectively with reduced system complexity [13], [14]. The dynamic emotion model is proposed to make a change of facial expressions with different property according to the kind of stimuli [15]. So we can make a robot’s facial expression more natural and similar to humans.

The overview of a linear dynamic affect-expression model is as in Fig. 1. The force element vector \( r \) is generated by some perception process and the number of force elements differs depending on researchers or systems. If four kinds of emotion can be classified from external stimuli, then \( D_r \) is 4 [15]. The force vector \( s \) which makes the emotion goes to target position in affect space is computed by multiplying matrix \( E \) and \( K \) as in (2)

\[
\begin{align*}
\mathbf{r} & = \left[ r_1, r_2, \ldots, r_k, \ldots, r_{D_r} \right]^T, \quad 0 \leq r_k \leq 1 \\
\mathbf{s} & = K_{D_r \times D_e} E_{D_r \times D_e} \mathbf{r} \\
D_r & : \text{the dimension of affect space} \\
D_e & : \text{the dimension of expression space}
\end{align*}
\]

And then if the force vector \( r \) is set as input of dynamic equation, we can get the trajectory of emotion \( \mathbf{e} \) while it reaches target position by (3).

\[
M\dot{\mathbf{e}} + C\mathbf{e} + Ke = \mathbf{s}
\]

The determined emotion \( \mathbf{e} \) is transformed to expression vector \( \mathbf{p} \) by transition matrix \( T_f \) and \( T_e \) as in (4). \( T_f \) in Fig. 1 is decomposed into \( T_{f1} \) and \( T_{f2} \). These matrices are obtained by establishing affect and expression space [14].

\[
P_{(D_e \times 1)} = T_{f2(D_e \times D_f)} T_{f1(D_f \times D_e)} \mathbf{p}_{(D_e \times 1)}, \quad D_p \geq D_e
\]

As you can see in Fig. 2, the dimension of affect space is three and that of expression space is six for facial robot ‘Doldori’. This is made to propose the linear affect-expression space model. The linear affect-expression space model is completed by the following procedure: first of all, robots’ expression space is defined. The expression space has six bases with values of control points on expressions corresponding to Ekman’s six basic emotions [16], [17]. Then the robot’s three-dimensional affect space is determined by using its expression space. Finally, it is proved that the affect space is related to expression space linearly.

If we use a linear dynamic affect expression model, we can...
represent the exact position of robot’s emotion in the affect space and the trajectory between current position and target position of emotion is decided. Therefore, geometric distance between each emotion can be obtained in the affect space. So, the emotional boundaries between one emotion and the others can be found using positions of six basic emotions.

III. EMOTION BOUNDARY

First of all, we need three assumptions to decide emotional boundaries. First, gestures are non-linear according to emotional intensity. Most humans represent their emotional state with facial expression when the intensity of emotion is weak. But if the intensity of emotion is strong, humans use facial expression and gestures to express their emotional state. For instance, if the intensity of surprise is low, humans do not use gestures, but when the intensity of surprise is strong, humans usually lift up their hands. Second, we consider that the emotion with recognition rate higher than 80% is strong one. Third, it is assumed that the most left images (seven basic expressions including neutral) used on questionnaire has 100% recognition rate. Robot’s facial expression can not be recognized perfectly as humans’ facial expression can not. Thus above assumptions are necessary to decide emotional boundaries.

Emotion is linearly related to expression in a linear dynamic affect-expression model and there exists a transition matrix $T_{ij}$ as shown in Fig. 3. However, transition matrices mapping emotion into each space of expressional modalities are necessary like facial expression. Namely, as you can see in Fig. 3, $T_{ij}$, $T_{kj}$, ..., $T_{nj}$ have to be determined. If transition matrices are found, the problem of synchronizing are solved because transition matrices transform determined emotion to the expression space of each modality. But in the case of modalities, which are nonlinear according to emotion, like gestures, finding transition matrix between affect space and its expression space is not easy. So expression of gestures using emotional boundaries will be meaningful because we can decide whether to use gestures or not as to the position of target emotion which is inside of emotional boundaries or not.

If we adopt a linear dynamic affect-expression model in ‘Doldori’, we can get the position of emotions corresponding to six basic expressions in (6). Since the position of each emotion in a three-dimensional affect space and intensity of emotion can be represented according to the geometric distance between emotions, we can also find emotional boundaries which are determined by mapping emotions in the expression space to the affect space.

$$E = \begin{bmatrix} e_{\text{anger}} & e_{\text{surprise}} & e_{\text{happiness}} & e_{\text{sadness}} & e_{\text{disgust}} & e_{\text{fear}} \end{bmatrix}$$ (6)

A. Questionnaire

We can compute the geometric distance through Euclidean norm between seven expressions including neutrality using the position of each expression in expression space as presented in (7). So we can investigate recognition rate of linearly interpolated facial expressions and results of recognition results will be used to determine emotional boundaries.

Linearly interpolated facial expression is obtained by (8) and for instance, the results of facial expression from happiness to surprise are shown in Fig. 5. After all, five new images are needed between two emotions. Likewise, we used totally 217 images because we obtained the interpolated images from one emotion to other six emotions for each emotion.

$$\Delta \equiv \left[ \delta \right] = \left[ P_i - P_{j} \right]_{i=1, \ldots, 6}$$ (7)

$$P_i(i, j) = \frac{1}{6} \sum_{k=0}^{5} (P_{k} - P_{l}) + P_{l}$$ (8)

(Neutral, Anger, Surprise, Happiness, Disgust, Sadness, Fear)
B. Procedures and Results of Questionnaire

We conducted a questionnaire survey of 26 people in their 20 ~ 30's males and females. Questionnaire is composed of seven stepwise images corresponding to changing from one emotion to another. For example shown in Fig. 6, we place the images which are changed from happiness to surprise. Participants were asked to cluster images of happiness from left. Here we used second and third assumption mentioned in the beginning of this chapter. So we forced participants not to draw a vertical line in the left of the most left image, which is recognized with 100% rate by third assumption. Then we can consider the images located left of vertical line to be same category to the most left image.

TABLE I represents recognition rate of facial expression when it is changed from anger to the others. Since there are totally seven emotions (six basic emotions and neutrality), we can obtain 7 tables like TABLE I. If the geometric distance equals to 1, vertical line is placed on the most left in questionnaire. However, this situation can not be occurred because of the third assumption. If the distance equals to 8, it is for the most right. In TABLE I, we can see that all participants regarded first three images as the same emotional category of the most left image for anger.

If we draw a graph using TABLE I, we can obtain a variation of recognition rate and find the point which has 80% recognition rate for 7 emotions when one emotion changes to the others. We showed the point of 80% recognition rate in Fig. 7 because of the third assumption. As you can see in Fig. 7, distances which have 80% recognition rate between anger and the others are different. Also the distance found above is different in direction. The two graphs’ meeting point is not a center between two emotions when the emotion of happiness moves to that of surprise and vice versa. So the geometrical intensity which has 80% recognition rate is different in two cases. When the emotion of surprise moves to that of happiness, same recognition rate is obtained at longer distance (Fig. 8). Therefore it is expected that emotional boundaries have different shape for different kind of emotions.

C. Emotional boundaries in the expression space

We can define emotional boundaries in expression space based on relationship between recognition rate and geometrical intensity of emotion. Emotional boundaries are shown in Fig. 9 for anger and surprise. We can find that the shape and size are different as to each emotion and the
shape of boundaries with 80% recognition rate and those with 60% are also different. The reason is that the change of recognition rate is different between emotions. It is free addition that the anger and disgust are difficult to be distinguished because their geometric distance is closer than other emotion pairs.

D. Emotional boundaries in the affect space

Emotional boundaries in the affect space can be defined by (9). \( \mathbf{P}_{i,j} \) matrix consists of six vectors in emotional boundaries. The vector \( \mathbf{p}_{ij} \) consists of each control point value in boundary between \( p_i \) expression and another expression. There are seven \( \mathbf{P}_{i,j} \) matrices because of seven emotions.

\[
P_{i,j} \triangleq \begin{bmatrix} p_{0i} & p_{0j} & p_{1i} & p_{1j} & p_{2i} & p_{2j} \end{bmatrix}
\]  

(9)

If emotional boundaries are defined in the affect space since the emotion moves dynamically in affect space, we can apply them to gestures or other modalities. Therefore boundaries in the expression space have to be transferred in the affect space using transition matrix \( T \) which is determined in the linear affect-expression space model as presented in (10).

\[
\mathbf{E}_{i,j} = (T_i^T T_j)^{-1} T_i^T \mathbf{P}_{i,j}
\]  

(10)

In the case of ‘Doldori’, six coordinates are determined in the affect space for each emotion. After these coordinates are represented in the three-dimensional affect space, we can make polyhedron-shaped boundaries by connecting each other. However it is difficult that these boundaries are applied to multimodal expression because the dynamically moving emotion may not go through these boundaries which are made by linearly interpolating expressions. Thus not polyhedron shaped but spherical shaped boundaries are needed and if radius of sphere is determined as minimum distance, boundaries are sufficient to have at least 80% recognition rate for all emotions. If we use maximum or mean of distance, the boundary of one emotion may be overlapped with that of neighboring emotions or there exist some emotion that recognition rate does not reach 80%. So the minimum distance of each boundary that obtained by (10) are shown in TABLE II.

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>MINIMUM DISTANCE FROM AN EMOTION TO IT’S BOUNDARY FOR EACH EMOTION IN THE AFFECT SPACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neu</td>
<td>Ang</td>
</tr>
<tr>
<td>0.18</td>
<td>0.20</td>
</tr>
</tbody>
</table>

TABLE III | BOUNDARY WITH RECOGNITION RATE 80% FOR HAPPINESS IN THE EXPRESSION SPACE |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AU</td>
<td>Neu</td>
</tr>
<tr>
<td>1</td>
<td>-0.143</td>
</tr>
<tr>
<td>2</td>
<td>0.340</td>
</tr>
<tr>
<td>3</td>
<td>0.214</td>
</tr>
<tr>
<td>4</td>
<td>-0.071</td>
</tr>
<tr>
<td>5</td>
<td>-0.203</td>
</tr>
<tr>
<td>6</td>
<td>0.296</td>
</tr>
</tbody>
</table>

TABLE IV | MINIMUM DISTANCE FROM AN EMOTION TO IT’S BOUNDARY FOR EACH EMOTION IN THE AFFECT SPACE |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Neu</td>
<td>Ang</td>
</tr>
<tr>
<td>0.18</td>
<td>0.20</td>
</tr>
</tbody>
</table>

IV. APPLICATION OF EMOTIONAL BOUNDARIES

Gestures may be expressed when the intensity of emotion goes over the threshold. For example, humans express their emotional state through facial expression for low intensity of sadness but shed tears with lifting up their hands toward eyes for strong intensity of that. Thus gestures are a good example to be expressed using emotional boundaries since the feasibility of expressing gestures is determined by intensity of emotions. So if the target position of emotion is inside of its emotional boundary for the pattern of recognized stimuli, the stimuli may be recognized as strong that the robot expresses its emotion state as facial expression and gestures. On the other hand, if the target position of emotion is outside of its emotional boundary, all one have to do is expressing using just facial expression because we can consider stimuli to be weak.

V. EXPERIMENTAL RESULTS

Emotional boundaries are determined in the expression and affect space through above procedure for total seven emotions including six basic emotions and neutrality. For happiness, boundaries determined in the expression space were shown in TABLE III and Fig. 10. In the affect space,
they are shown in TABLE IV. And data in TABLE III are transformed to data in TABLE IV by (10). Fig. 11 shows boundaries with minimum distance for seven emotions. Theses boundaries are used to express robot’s emotional state using facial expression and gestures. And the result is shown in Fig. 12.

VI. CONCLUSION

This paper proposed emotional boundaries for multimodal expression. First we found boundaries in expression space and then mapped them on the affect space. It was then verified that emotional boundaries have different properties according to direction. Spherical type boundaries with at least an 80% recognition rate were chosen so as to avoid overlapping between neighboring emotions when one emotion approaches others. Finally it was demonstrated that the proposed emotional boundaries can be applied to modalities with a nonlinear property for emotion such as gestures using simulator with a dynamic engine. It should be noted that emotional boundaries derived in this paper are appropriate for ‘Dodori’. However, emotional boundaries for other robots can be found based on the procedure proposed in this paper. And we experimented to express gesture using arms, however, the gestures include not only movements of arms but also movements of shoulder, wrist, neck, etc. So if movements of upper body and lower body including pose of legs are considered, we can make expression of gestures more abundantly.

REFERENCES