

# 효과적인 로봇 행동 생성을 위한 선형의 정서-표정 공간 내 감정 경계의 결정

비선형의 제스처 동기화를 위한 정서, 표정 공간의 영역 결정

## The Emotional Boundary Decision in a Linear Affect-Expression Space for Effective Robot Behavior Generation

조수훈(Su Hun Jo)\*, 이희승(Hui Sung Lee)\*\*, 박정우(Jeong Woo Park)\*\*\*,  
김민규(Min Gyu Kim)†, 정명진, Myung Jin Chung‡

**요약** ~ 미래에는 로봇이 사람의 감정 상태를 이해하거나 적절하게 자신의 행동을 표현하는 기능은 중요해질 것이다. 사람간의 교류에서 메시지의 93%가 행동 표현에 있으며, 바디 랭귀지는 감정의 양을 표현하므로 행동 표현은 중요한 감정 표현 수단이다. 최근의 로봇들은 얼굴, 제스처, LED, 소리 등의 복합적 방법을 이용하여 사람과 교류하고 있지만 포즈는 위치와 방위, 얼굴이나 제스처는 속도, 말이나 색 등은 시간에 대한 정보가 필요하기 때문에 하나의 모델로 통합하거나 동기화 시키기 어렵다. 한편 작은 세기의 감정에서, 얼굴은 쉽게 표현이 가능하지만 제스처는 표현이 힘들다. 또한 기존의 감정 경계는 같은 형태와 크기를 가지거나, HHI 분야에 국한되어 연구되어 왔다.

본 논문에서는 정서 공간에서 감정의 경계를 어떻게 정의할 것이며, 복합적 표현 방법을 시스템적으로 어떻게 동기화할 수 있을지를 제안한다.

**Abstract** ~ In the near future, robots should be able to understand human's emotional states and exhibit appropriate behaviors accordingly. In Human-Human Interaction, the 93% consist of the speaker's nonverbal communicative behavior. Bodily movements provide information of the quantity of emotion. Latest personal robots can interact with human using multi-modality such as facial expression, gesture, LED, sound, sensors and so on. However, a posture needs a position and an orientation only and in facial expression or gesture, movements are involved. Verbal, vocal, musical, color expressions need time information. Because synchronization among multi-modalities is a key problem, emotion expression needs a systematic approach. On the other hand, at low intensity of surprise, the face could be expressed but the gesture could not be expressed because a gesture is not linear. It is need to decide the emotional boundaries for effective robot behavior generation and synchronization with another expressible method. If it is so, how can we define emotional boundaries? And how can multi-modality be synchronized each other?

**핵심어:** *emotional boundary, emotional region, emotion, behavior synchronization, facial robot, affect space*

### 1. 서론

퍼스널 로봇(Personal Robot)의 영역이 공장 중심에서 회사, 가정으로 확대되어 감에 따라 로봇의 행동에 대한 사람들의 기대가 커지고 있다. 초기 로봇은 산업용으로 단지 주어진 목표만을 수행하기 위한 머니플레이터로 여겨졌으나, 기술이 발전함에 따라 로봇은 인간의 생활 속에서 편의를 제공하고 인간과 교류하기 위한 동반자로 여겨지고 있다. 또

한 로봇은 인간과의 정보 교류뿐만 아니라 감정의 교류를 통해 인간에게 좀 더 친숙하게 다가갈 수 있을 것이다. 로봇이 사회적 로봇(Social Robot)으로써의 기능을 수행하기 위해서 인간의 감정을 인식하고 분석하여 인간에게 감정을 표현하는 로봇과 인간간의 상호교류(HRI: Human-Robot Interaction)가 필요하다는 점은 인간간의 상호교류(HHI: Human-Human Interaction)에서 찾아볼 수 있다. 인간간의 상호교류에서는 의사 소통을 위해 말, 행동, 표정 등 다양한

본 논문은 산업자원부 지원으로 수행되는 인간기능 생활지원 지능로봇 기술개발사업의 일환으로 수행되었음.

\*주저자, 교신저자 : 한국과학기술원 전자전산학과부 석사과정; e-mail: jsh@cheonji.kaist.ac.kr

\*\*공동저자 : 한국과학기술원 전자전산학과부 박사과정; e-mail: eins@cheonji.kaist.ac.kr

\*\*\*공동저자 : 한국과학기술원 전자전산학과부 박사과정; e-mail: pjw@cheonji.kaist.ac.kr

†공동저자 : 한국과학기술원 전자전산학과부 석사과정; e-mail: mgkim@cheonji.kaist.ac.kr

‡공동저자 : 한국과학기술원 전자전산학과부 교수; e-mail: mjchung@ee.kaist.ac.kr

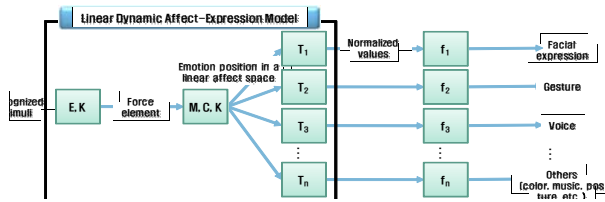


그림 1. 감정 표현의 시스템적 접근 방법

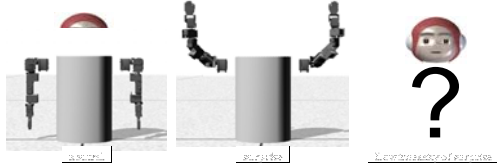


그림 2. 제스처의 비선형적 특성(놀람의 예)

수단을 이용하여 상호간의 정보나 감정을 교류하고, 관계를 유지할 수 있으며, 그 중 얼굴 표정은 상호 교류 시 주요 수단으로써 상호간의 지식이나 감정 상태를 예상할 수 있게 해준다. 더불어 목소리, 몸짓, 자세 등의 추가적인 정보를 이용하여 상대방의 느낌을 이해하게 되고 감정을 쉽게 교류할 수 있게 된다[1,2]. 현재 인간과의 의사소통과 감정교류를 위해 많은 로봇이 개발되고 있지만 얼굴 표정만으로는 만족할 수 있는 사람과의 상호교류가 이루어질 수 없기 때문에 복합적 방법(Multi-modality)을 통한 추가적인 정보를 제공해야 할 필요가 생겼고, 최근의 로봇들은 얼굴 표정, 제스처(gesture), LED, 소리, 음악, 센서 등 다양한 방법을 통해 인간에게 자신의 감정 상태에 대한 정보를 제공하고 있다.

하지만 복합적 방법들을 무조건 사용하는 데에는 근본적인 문제가 있다. 포즈(posture)는 각 관절(joint)의 위치(position)와 방위(orientation)에 대한 정보가 필요하고, 제스처는 속도에 대한 정보가 필요하다. 소리(sound)나 음악(music), 색(color) 같은 경우에는 시간에 대한 정보도 필요하다. 각 방법에 따라 필요한 정보가 다르기 때문에 어떤 감정에, 어떤 표현을, 어느 시간에 맞추는 식의 접근은 복잡하고, 동기를 맞추기 힘들다. 따라서 복합적 방법으로 로봇의 감정을 표현하기 위해서는 동시에 표현을 시작시킬 수 있는 시스템적인 접근 방법이 필요하다. 기존에 제시된 선형의 동적 정서-표정 모델은 이 문제를 해결할 수 있는 실마리를 제공한다. 감정이 인지된 후 연산 과정을 거치면, 정서 공간에서 동적으로 움직이는 감정이 결정되는데, 그 감정을 표현 방법에 맞추어 표정 공간으로 사상시키면 우리가 원하는 얼굴 표정이나 제스처, 음성 등의 표현 방법들을 동기화시켜서 표현할 수 있게 된다 [3][그림 1].

그러나 이런 동기화 문제를 해결하더라도 또 하나 문제가 있다. 기존의 얼굴 표정이나 LED 같은 표현은 비선형 특성을 선형화시킬 수 있어서 모델을 적용하는 데 큰 어려움이 없었지만 제스처나 음성 같은 경우에는 감정의 세기에 따라 비선형의 특성을 가지기 때문에 선형 모델을 적용하기 위해서는 적당한 동기 시점이 필요하게 된다. 예를 들어 놀란 감

정을 표현하려면, 얼굴 표정은 눈썹이 올라가고 눈꺼풀이 올라가서 눈동자 위의 흰자가 보이게 되며 입은 크게 벌리게 된다. 그리고 팔의 제스처는 팔을 높이 들어올리고 손바닥을 상대방에게 보이는 표현을 하게 된다. 하지만 약한 세기의 놀란 감정을 표현할 경우, 얼굴 표정은 선형화될 수 있기 때문에 작은 각도만 움직이면 약간 놀란 표정같이 보인다. 그러나 제스처의 경우에는 약한 세기의 감정에 대한 표현을 명확하게 정의하기 힘들다 [그림 2]. 이것은 표현하고자 하는 감정에 대한 경계를 알고 있으면 해결할 수 있다. 약한 세기의 감정에 대해서는 경계 바깥쪽에 위치하기 때문에 선형화시킬 수 있는 표현 방법만 표현하다가, 목표 감정이 어떤 감정의 경계 안쪽에 있으면 사람이 인지할 수 있는 해당 감정이 확실하기 때문에 선형화가 가능한 표현 방법과 함께 제스처를 표현하거나 다른 비선형 복합적 방법들을 표현하면 된다.

본 논문은 로봇이 인간과 원활하고 효율적으로 감정 교류를 할 수 있도록 제안된 선형의 동적 정서-표정 모델에서 감정의 경계를 결정하고, 이를 바탕으로 얼굴 표정과 함께 행동 표현의 동기화를 목적으로 한다. 표정 공간에서 80% 인식률을 가지는 감정의 경계를 결정한 다음, 정서-표정 모델을 이용하여 표정 공간의 경계를 정서 공간의 경계를 제안한다.

## 2. 선형의 동적 정서-표정 모델

### 2.1 선형의 정서-표정 모델

선형의 정서-표정 모델을 제안하면서 만들어진 얼굴 표정 로봇 '돌도리'에서는 에크만(Ekman)이 제안한 6가지 기본 표정(화남, 놀람, 기쁨, 혐오, 슬픔, 공포)으로 이루어진 표정 공간을 정의하였다. 이를 이용하여 역으로 정서 공간을 정의한 후, 선형 사상 관계가 있음을 증명하였다 [3,4]. '돌도리'에서는 기본 6가지 표정이 정서 공간의 기저가 될 수 있다고 가정하였다. '돌도리'의 6가지 기본 컨트롤 포인트는 눈썹, 안쪽 눈꺼풀, 바깥쪽 눈꺼풀, 윗입술, 아랫입술, 입 꼬리다.

식 (1)에서  $P_1, P_2, P_3, P_4, P_5, P_6$ 는 기본 6가지 표정에 해당하는 컨트롤 포인트의 정규화된 값을 실험적으로 얻어낸 것이고, 시스템의 복잡도를 줄이기 위해 PCA, MDS

$$\begin{aligned}
 P_0 &= [P_1, P_2, P_3, P_4, P_5, P_6] \\
 &= \begin{bmatrix} P_{Anger} & P_{Surprise} & P_{Happiness} & P_{Disgust} & P_{Sadness} & P_{Fear} \end{bmatrix} \\
 &= \begin{bmatrix} P_{Eyebrows} \\ P_{InnerEyelids} \\ P_{OuterEyelids} \\ P_{UpperLip} \\ P_{LowerLip} \\ P_{LipCorners} \end{bmatrix} = \begin{bmatrix} 0.78 & -0.27 & -0.26 & 0.55 & -0.65 & -0.17 \\ -0.24 & 0.86 & 0.62 & -0.07 & 0.53 & 0.56 \\ 0.60 & 0.82 & 0.39 & 0.46 & -0.13 & 0.42 \\ 0.46 & 0.51 & -0.13 & 0.42 & 0.22 & 0.13 \\ -0.15 & -0.73 & -0.37 & 0.06 & -0.14 & -0.36 \\ -0.19 & -0.13 & 0.54 & -0.49 & -0.57 & -0.50 \end{bmatrix} \\
 E &= \begin{bmatrix} e_{Anger} & e_{Surprise} & e_{Happiness} & e_{Disgust} & e_{Sadness} & e_{Fear} \end{bmatrix} \\
 &= \begin{bmatrix} 1.00 & 0.00 & -0.83 & 1.00 & 0.00 & 0.22 \\ 0.00 & 1.00 & 0.94 & -0.21 & 0.00 & 0.33 \\ 0.00 & 0.00 & -0.90 & 0.52 & 1.00 & 0.49 \end{bmatrix} \quad (1)
 \end{aligned}$$

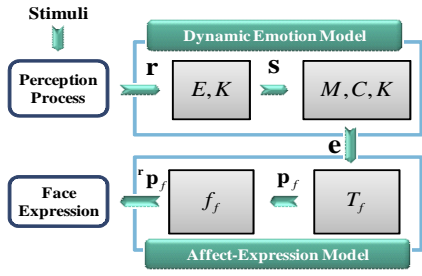


그림 3. 선형의 동적 정서-표정 모델의 개요

방법을 사용하여 전체 표정 공간의 90% 이상을 확장할 수 있는 정서 공간의 기저 개수를 3개로 선택하였다. '돌도리'의 경우에는 3차원 정서 공간의 기저로 화남, 놀람, 슬픔을 표현할 수 있는 정서가 결정되었다.

## 2.2 선형의 동적 정서-표정 모델

사람의 얼굴 표정은 얼굴 근육 움직임의 조합으로 만들어 질 수 있는데, 2차 미분 방정식으로 얼굴 근육의 움직임을 표현할 수 있기 때문에 정서 공간의 기저도 2차 미분 방정식으로 정의할 수 있다. 감정은 질량(m)을 가지는 물체로 3차원 정서 공간상에서 움직인다. 또한 자극에 대한 감정 표현 속도가 자극의 종류에 따라 다르다. 놀람 같은 경우에는 매우 빠르게 나타나는 반면, 슬픔은 천천히 나타나는 특징이 있기 때문이다. 그리고 놀람은 오버슈트(overshoot)도 존재할 수 있다. 사람 같은 경우에는 놀랐을 때 눈과 입을 크게 벌렸다가 살짝 오므리는 현상을 볼 수 있기 때문이다. 마지막으로 자극을 입력하였을 때와 제거했을 때의 반응 속도가 달라야 한다. 놀람 감정이 사라질 경우에는 오버슈트가 없고 천천히 평정을 찾아가기 때문이다.

'돌도리'의 경우에는 정서 공간의 기저가 화남, 놀람, 슬픔을 표현할 수 있는 정서이기 때문에 앞의 가정에 의해서  $\omega_n$  (natural frequency)은 놀람이 가장 크고, 화남, 슬픔 순으로 결정된다.  $\zeta_n$  (damping ratio)는 슬픔이 가장 크고, 화남, 놀람의 순으로 결정할 수 있다. 한편  $c_i$  (viscosity)는 자극이 제거될 때는 오버슈트가 없기 때문에  $c_{i-F} < c_{i-B}$ 의 관계가 성립한다. 따라서 식 (2)를 실험적으로 결정하여 사람과 비슷하게 표정을 지을 수 있다.

$$\begin{aligned}
 M\ddot{\mathbf{e}} + C\dot{\mathbf{e}} + K\mathbf{e} &= \mathbf{s} \\
 \mathbf{e} &= [e_1 \ \dots \ e_{D_e}]^T, \ \mathbf{s} = [s_1 \ \dots \ s_{D_s}]^T \\
 M &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} K = \begin{bmatrix} 150 & 0 & 0 \\ 0 & 180 & 0 \\ 0 & 0 & 130 \end{bmatrix} \\
 C_F &= \begin{bmatrix} 50 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 110 \end{bmatrix} C_B = \begin{bmatrix} 150 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 200 \end{bmatrix} \quad (2)
 \end{aligned}$$

## 2.3 전체 모델 개요

$\mathbf{r}$ 은 각 감정의 세기를 요소로 가지는 벡터로 구성된다. 감정의 개수만큼의 행을 가진 벡터라고 할 수 있다. 또한  $\mathbf{r}$ 은 힘 요소(force elements)이며 식 (3)을 통해 최종 상태 오류(steady state error)를 없앤다. 다음으로 정서 공간에서 감정의 동역학에 적용될 수 있는 힘(force)  $\mathbf{s}$ 로 변환된다. 이후에는 정서 공간의 감정 벡터  $\mathbf{e}$ 가 생긴다. 그러면 선형의 정서-표정 모델에서 제안한 식 (3)을 이용해 감정 벡터를 표정 공간으로 사상시킬 수 있고, 함수  $f$ 를 통해 표정 벡터  $\mathbf{p}$ 로 실제 로봇의 얼굴을 표현할 수 있다 [그림 3].

$$\begin{aligned}
 \mathbf{r} &= [\gamma_1 \ \gamma_2 \ \dots \ \gamma_k \ \dots \ \gamma_{D_r}]^T \quad 0 \leq \gamma_k \leq 1 \\
 \mathbf{s} &= K_{(D_s \times D_e)} E_{(D_e \times D_r)} \mathbf{r} \\
 \mathbf{p}_{(D_p \times 1)} &= T_{2(D_p \times D_e)} T_{1(D_e \times D_s)} \mathbf{e}_{(D_e \times 1)}, \quad D_p \geq D_e \\
 \mathbf{r}_{(D_r \times 1)} &= f(\mathbf{p}_{(D_p \times 1)}) \\
 T1 &= \begin{bmatrix} 0.78 & -0.24 & 0.60 & 0.46 & -0.15 & -0.19 \\ -0.27 & 0.86 & 0.82 & 0.51 & -0.73 & -0.13 \\ -0.65 & 0.53 & -0.13 & 0.22 & -0.14 & -0.57 \end{bmatrix} \quad (3)
 \end{aligned}$$

## 3. 감정의 경계

### 3.1 기존 연구

감정의 경계에 대한 연구는 에트코프(Etcoff)가 처음으로 시작하였다 [5]. 사람의 표정 이미지를 통해 6개의 기본 감정이 각 감정으로 가는 모든 경우에 대해 설문 조사하여 경계를 결정하였다. 로봇 분야에서도 로봇의 감정 경계를 사용하였다. MIT의 'Kismet'은 AVS 정서 공간에서 감정이 연속적으로 변하게 되고 감정 식에 의해 9가지의 감정 중에서 결정하게 된다 [6]. 하지만, 각 감정의 경계의 형태나 크기에 대한 근거가 없기 때문에 행동 표현에 사용하기에는 어렵고, 각 감정 경계의 크기 차이가 없기 때문에 얼굴 표정과 함께 동시에 표현하기에는 어색하다. Waseda 대학의 'WE-4RII'는 PAC 정서 공간에서 감정의 동적 모델을 통해 감정이 움직이고 6개의 감정 중에서 결정된다 [7]. 하지만, 감정의 경계가 서로 근접해 있기 때문에 얼굴 표정과 제스처가 불연속적이라는 단점이 있었다. Business Design Lab의 'Ifbot'은 신경망을 이용하여 3차원의 정서 공간을 만들었다 [8,9]. 15개의 얼굴 요소를 45개의 요소로 변환한 다음 다시 3개의 요소로 줄여서 정서 공간의 기저를 결정하게 된다. 그러나 정서 공간의 각 축에 대한 정보를 명확하게 정의할 수 없기 때문에 기저의 특성을 사용할 수 없고, 다른 로봇에 적용하기가 힘들다. 한편 바트넥(Bartneck, 2005)은 컴퓨터 그래픽 캐릭터를 이용하여 중립 감정에서 5개의 감정

으로 변화할 때의 인식률을 조사하여 30%의 기하학적인 강도만 있으면 사람은 어느 정도 감정을 인지할 수 있다고 주장하였다 [10]. 하지만 중립 감정에서 다른 감정으로 변하는 부분만 조사를 했기 때문에 감정 간의 변화에 대한 고찰이 부족하며, 로봇에 적용하기에는 미흡하다는 문제가 있다.

### 3.2 설문조사

#### 3.2.1 차이 행렬을 이용한 이미지 추출

앞 장의 선형의 동적 정서-표정 모델을 '돌도리'에 적용하여, 정서 공간의 각 감정을 3차원 정서 공간에 표시하면 중립 감정은 3차원 축의 중앙에 위치하고 있고 각 위치에 6개의 감정을 표시할 수 있다. 만약 중립 위치에서 각 감정에 대한 자극을 입력한다면 감정이 변하는 경로를 표현할 수 있다. 자극이 입력되었을 때와 제거되었을 때의 경로를 모두 확인할 수 있다. 이처럼 3차원 정서 공간에서는 각 감정마다 위치가 정해져 있고, 기하학적 거리에 따라 감정의 세기를 표현할 수 있기 때문에 감정의 경계도 정할 수 있다. 선형의 정서-표정 모델을 정의하듯이 경계를 찾을 때도 표정 공간에서 정서 공간으로 변환하여 찾아올 수 있다. 따라서 표정 공간의 각 감정사이의 기하학적 거리를 알고 있기 때문에 이것을 선형으로 보간한 다음 그 표정을 이용하여 인식률을 조사할 수 있다.

#### 3.2.2 설문 조사 문항 및 참여자

표정 벡터의 차이 행렬로 추출한 292장의 이미지를 이용하여 설문 조사를 한다면 각 감정에서 다른 감정으로 변화하는 모든 경우를 생각할 수 있다. 설문 조사는 한 화면에 순차적으로 감정이 변하는 이미지 7장을 2 줄로 배열해두고, 왼쪽에서 오른쪽으로 감정이 변화할 때 어느 부분까지 왼쪽 감정 같은지에 대해 경계를 만들라고 하였다. 여기에서 하나의 가정이 필요하다. 돌도리의 얼굴 표정은 실제로는 100%의 인식률을 가지지 못한다. 하지만 각 감정의 경계를 정하기 위해서는 만들어진 표정이 100% 인식된다고 가정하지 않으면 경계를 결정하지 못하는 근본적인 문제가 생기게 된다. 따라서 가장 왼쪽에 보이는 표정은 100% 그 감정의 표정으로 인지된다고 가정한다면 설문 응답자가 표시한 영역까지는 해당 감정이라고 생각할 수 있다 [그림 4].

이 설문 조사는 총 26명에게 실시되었고, 응답자들은 20~30대 연령의 학생으로, 파일 형태로 설문지를 제시하면 결과 파일을 되돌려 주는 형식으로 진행되었다. 설문지에는 한 페이지에 이미지 14장을 배치하였고, 설문 응답자의 기본 정보에 대한 질문 페이지를 포함하여 총 22장의 페이지를 보여주었다. 설문 응답에 걸리는 시간은 평균 10분이었다.

#### 3.2.3 설문 조사 결과 및 분석

설문 조사의 문항을 살펴보면 한 변화당 이미지의 개수가 7개이고 경계를 나눌 수 있는 부분은 8개이다. 하지만 설문 조사 문항에서 첫 번째 이미지의 왼쪽으로는 경계를 줄 수

없도록 하였다. 설문 조사 하기 전에 가정된 데로 '돌도리'가 표현하는 각 얼굴 표정은 완벽한 인식률을 가지고 있다고 생각하기 때문이다. 만약 가장 왼쪽 이미지의 왼쪽으로도 경계를 줄 수 있게 되면 가정과 상충되기 때문에 설문 조사 문항에서 그 부분에 제한을 두었다. 그리고 적어도 80%의 인식률을 가지고 있어야 그 감정의 경계로 생각할 수 있기 때문에 그래프로 그려서 다시 각 점 사이를 선형 보간하면 80%의 인식률을 가진 기하학적 세기를 구할 수 있다 [그림 5]. x축은 기하학적 세기를 나타내고, y축은 인식률을 나타내었다. 점과 점 사이를 선으로 연결하고 필요한 80% 인식률의 기하학적 세기를 구할 수 있다. 한 감정에서 다른 감정으로 변할 때, 80% 인식률을 가지기 위해서는 다른 크기의 기하학적 세기가 필요하다는 것을 알 수 있다.

### 3.3 표정 공간에서의 경계 결정

기하학적 세기와 인식률의 관계 중에서, 80%의 인식률에 대한 각 감정의 기하학적 세기를 토대로 표정 공간에서 감정의 경계를 정의할 수 있다. 중립을 포함한 7개의 감정을 생각하면, 한 감정에서 다른 감정으로 변화하는 경우는 6가지가 있다. 따라서 6각형 형태의 표정 공간의 경계를 만들어 볼 수 있고, 각 감정의 경계는 각각 다른 형태와 크기를 가지고 있음을 확인할 수 있다 [그림 6].

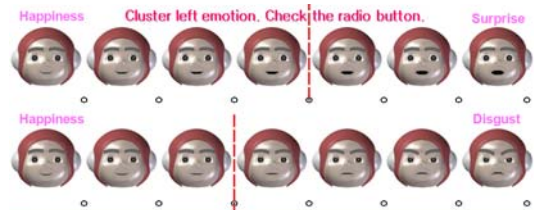


그림 4. 설문 문항 예제(기쁨에서 놀람이나 혐오로 변화 시)

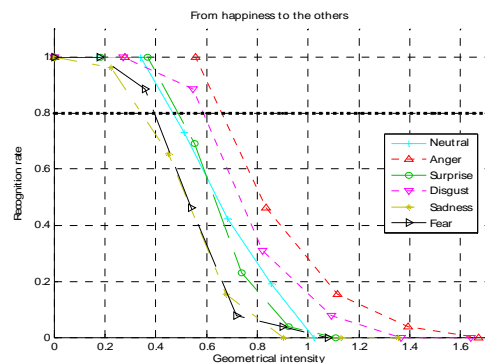


그림 5. 기쁨 감정에서 다른 감정으로 변화 시 인식률





그림 6. 표정 공간에서 기쁜 감정의 경계

하나의 그래프는 다른 6개의 감정으로 변화할 수 있고, 한 감정에서 다른 감정으로의 기하학적 거리가 다르기 때문에 짙은 색의 경계(Limit)는 다른 형태와 크기를 가지고 있다. 또한 80%의 인식률과 60%의 인식률을 표시해보면 기하학적 거리의 경계모양과 일치하지 않음을 알 수 있다. 이것은 인식률의 변화량이 감정마다 다르기 때문이다. 그리고 기존의 연구에서 행했던 '돌도리'의 인식률 조사에 대한 근거도 찾아볼 수 있다. 화남과 혐오는 기하학적 거리가 다른 감정에 비해 가깝기 때문에 조금만 감정의 세기가 낮아지면 사람들이 인식하기 힘들거나, 다른 감정과 혼동할 수 있다. 슬픔과 공포 또한 마찬가지로 설명할 수 있다.

### 3.4 정서 공간에서의 경계 결정

감정은 정서 공간에서 동적으로 움직이는 물체이기 때문에 정서 공간에서 경계를 정해야만 경계를 사용할 수 있다. 따라서 선형의 정서-표정 모델을 이용하여 표정 공간의 경계를 정서 공간으로 변환시킨 다음, 후 처리 과정을 거치면 정서 공간에서 경계를 결정할 수 있다.

#### 3.4.1 표정 공간에서 정서 공간으로의 변환

선형의 정서-표정 모델은 정서 공간을 선형 사상시켜서 표정 공간으로 바꾸고, 다시 함수를 이용해서 실제 표정 표현으로 만들어내는 모델이다. 앞의 절에서 표정 공간의 경계가 만들어졌기 때문에 식 (4)를 사용하면 정서 공간으로 변환시킬 수 있다.

$$e_i = (T_i^T T_i)^{-1} T_i^T P_{Bi}, \quad i = \text{each emotion} \quad (4)$$

#### 3.4.2 정서 공간에서 다면체 경계의 결정

6차원의 표정 공간에서 3차원 정서 공간으로 변환되면서 정서 공간에 6개의 좌표를 얻었다 [표 1][표 2]. 이것을 3차원 정서 공간에 표시를 한 다음, 서로 연결하면 다면체의 경계가 그려지게 된다 [그림 7]. 하지만 이것을 동적 모델에 바로 사용하기는 힘들다. 동적 모델에서 감정은 정서 공간의 기저의 특성을 적용 받아 움직이기 때문에 표정 공간에서 선형으로 보간하여 얻은 다면체의 경계를 통과하지 않을 수 있기 때문이다. 예를 들어 기쁜 감정에서 다른 감정으로 변화할 경우에 대한 경로를 그려볼 수 있는데, 경계를 통과하지 않고 그 주위로 이동하는 것을 확인할 수 있다 [그림 8].

따라서 동적 모델을 적용하기 위해서는 다면체 형태의 경계가 아닌 구 형태의 경계가 필요하다.

#### 3.4.3 정서공간에서 평균값 반지름을 가지는 구 형태 경계 결정

이미 알고 있는 한 감정과 80%의 인식률을 가지는 경계 지점인 6개 좌표를 유클리드 놈으로 평균하면 하나의 값을 얻을 수 있다 [표 3]. 이것을 각 감정이 가질 수 있는 구 형태 경계의 반지름으로 사용하면 3차원 정서 공간에서 각 감정의 경계를 정할 수 있다. 각 감정마다 기하학적 세기의 평균값이 다르기 때문에 반지름의 크기가 달라지게 되고 3차원 공간상에서는 다른 크기의 구 형태 경계를 만들 수 있다.

하지만 표정 공간의 경계에서도 알 수 있었듯이 화남과 공포, 혐오의 기하학적 거리가 가깝기 때문에 이것을 변환한 정서 공간에서도 거리가 가까울 것이라는 것을 예상할 수 있고, 실제로 정서 공간의 감정 경계는 겹치고 있음을 확인할 수 있다. 구 형태의 경계의 반지름을 평균하게 되면 기하학적 거리가 가까웠던 화남, 공포, 혐오 같은 경우에는 실제 거리보다 멀어진 점에서 80% 인식률의 경계를 가지게 되기 때문이다. 따라서 평균값을 이용하기 보다는 최소값을 이용한 경계가 필요하다.

#### 3.4.3 정서공간에서 최소값 반지름을 가지는 구 형태 경계 결정

다시 감정과 다면체의 경계 간의 유클리드 놈 중 최소값을 구한 다음 정서 공간에서 경계를 만들어 확인해 보면 겹치지 않은 경계를 확인할 수 있다 [표 4][그림 9]. 이 경계는 최소값을 기준으로 생성되었기 때문에 경계 안쪽은 반드시 80%의 인식률을 가지는 해당 감정이라고 볼 수 있다. 이제는 동적 모델을 사용하더라도 동적으로 움직이는 감정이 경계를 통과할 수 있다.

### 3.4 경계를 이용하여 행동 표현을 하는 방법

한 감정에 대한 행동 표현은 사람마다 다르고, 그 수는 매우 다양하다. 하지만 로봇은 구조적인 문제 때문에 아직 그렇게 많은 행동을 정해놓고 표현할 수 없다. 각 감정에 몇 가지 행동 표현이 있다고 한다면 어느 상황에 맞춰서 그 행동을 표현해야 하는지 정할 수가 없기 때문이다. 따라서 각 감정에 하나의 주요한 행동 표현이 있다고 가정하면서 문제

표 1. 80% 인식률을 가지는 기쁜 표정의 경계

Action Unit	Neutral	Anger	Surprise	Disgust	Sadness	Fear
Eyebrows	-0.143	0.151	-0.264	0.028	-0.358	-0.227
Inner Eyelids	0.340	0.280	0.725	0.374	0.597	0.598
Outer Eyelids	0.214	0.473	0.578	0.415	0.259	0.401
Upper Lip	-0.071	0.103	0.149	0.066	-0.042	-0.035
Lower Lip	-0.203	-0.283	-0.527	-0.217	-0.312	-0.366
Lip Corners	0.296	0.252	0.248	0.174	0.260	0.161

표 2. 80% 인식률을 가지는 기쁜 정서의 경계

Affect Value	Neutral	Anger	Surprise	Disgust	Sadness	Fear
affect1(Anger)	-0.455	-0.106	-0.467	-0.166	-0.620	-0.447
affect2(Surprise)	0.516	0.568	0.966	0.531	0.703	0.719
affect3(Sadness)	-0.494	-0.544	-0.507	-0.395	-0.421	-0.391

표 3. 정서 공간의 평균값 경계의 크기

Neutral	Anger	Surprise	Happiness	Disgust	Sadness	Fear
0.26	0.42	0.50	0.50	0.33	0.48	0.23

표 4. 정서 공간의 최소값 경계의 크기

Neutral	Anger	Surprise	Happiness	Disgust	Sadness	Fear
0.18	0.20	0.31	0.35	0.20	0.32	0.16

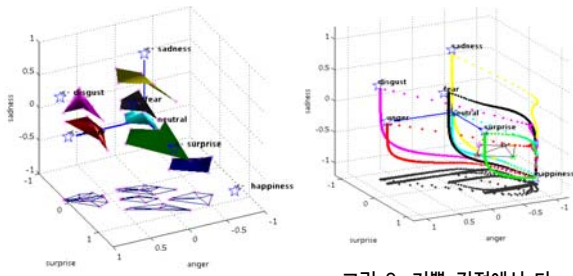


그림 7. 정서공간의 다각형경계

그림 8. 기본 감정에서 다른 감정으로 변화 시

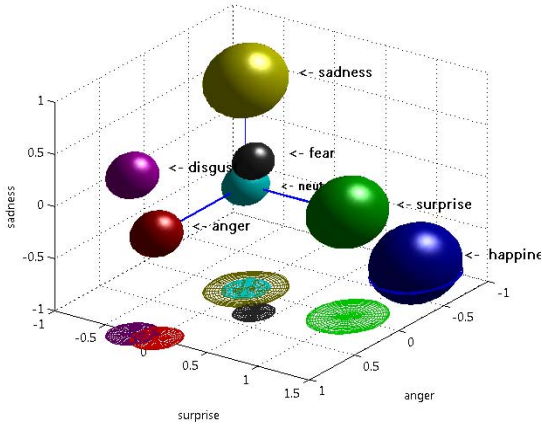


그림 9. 정서 공간에서 최소값을 이용한 구 형태의 경계

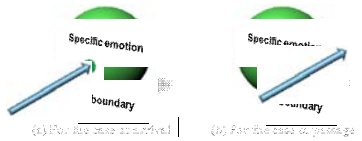


그림 10. 감정이 경계가 있는 정서 공간에서 변하는 경우

를 단순화 시킨 다음, 접근하는 방법이 타당하다. 각 감정에 하나의 주요 행동 표현이 있다면 제안한 감정의 경계를 사용하는 데 문제가 없다. 예를 들어 행동 표현 중에서 얼굴 표정과 함께 가장 풍부한 감정 표현을 할 수 있는 제스처 표현의 경우 결정된 경계를 사용해 볼 수 있다. 동적 감정 모델은 최종 목표 값을 이미 알고 있고, 목표 값까지 가는 동안의 변화를 수식으로 정의한 것이다. 따라서 어떤 감정이 인지된 후에는 어떤 감정이라고 정해지게 되고, 어느 정도의 세기인지, 어떤 감정과 섞여 있는지를 알 수 있다. 그리고 3차원 정서 공간에서 어디에 위치해 있는지도 알 수 있다. 따라서 목표 값이 경계 안쪽이라면 감정의 강도가 센 것이고, 해당 감정과의 기하학적 세기가 큰 것이기 때문에 얼굴 표정과 함께 행동 표현을 시작하더라도 어색하지 않을 것이다. 반면에 목표 값이 경계 바깥쪽이라면 감정의 강도가 약하거나, 해당 감정과의 기하학적 세기가 작고, 어떤 감정인지 판단하기 힘들 수가 있다. 이럴 경우에는 복합 감정을 표현할 수 있는 얼굴 표정만으로 감정을 표현함으로써 추가 정보를 제공하지 않으면 된다 [그림 10].

#### 4. 결론

본 논문에서는 선형의 정서-표정 모델을 비선형의 행동 표현에 적용시킬 수 있도록 감정의 경계를 제안한다. 얼굴 표정 로봇 '돌도리'를 바탕으로 6차원의 표정 공간에서 다각형의 감정 경계를 제안하였고, 3차원의 정서 공간에서 80%의 인식률을 가지는 점까지의 유클리드 놈 최소값을 이용한 구 형태의 경계를 제안하였다. 제안된 감정의 경계를 이용하여 선형화가 가능한 감정 표현과 비선형의 특성을 가진 행동 표현을 동기화할 수 있는 시스템적 방법을 제안하였다.

아직은 선형의 정서-표정 모델이 '돌도리'에만 적용되기 때문에 감정의 경계 또한 '돌도리'에만 적용해 보았고, 경계를 결정할 때 설문조사 인원이 적어 오차 범위가 크지만, 제안된 경계를 통해서 행동 표현의 동기화 실마리를 찾을 수 있을 것이다.

추후 연구 과제로 비선형적 행동 표현을 실제로 적용시켜 경계를 검증해 보는 문제와 다른 로봇에 적용해 보는 문제가 남아있다.

#### 참고문헌

- [1] Mehrabian, A., "Communication without words", *Psychology Today*, 2 (9), 52-55, 1968.
- [2] Ekman P., Friesen W. V. "Detecting deception from the body or face", *Journal of Personality and Social Psychology*, 29, pp288-298, 1974.
- [3] Hui Sung Lee, Jeong Woo Park, Su Hun Jo and Myung Jin Chung, "A Linear Dynamic Affect-Expression Model: Facial Expressions According to Perceived Emotions in Mascot-Type Facial Robots", in *Proceedings 16th IEEE International Conference on Robot & Human Interactive Communication*, Aug 26, 2007.
- [4] Hui Sung Lee, Jeong Woo Park and Myung Jin Chung, "A Linear Affect-Expression Space Model and Control Points for Mascot-Type Facial Robots", *Robotics, IEEE Transactions*, Vol. 23, No. 5, October 2007.
- [5] Etcoff, Nancy L., Magee, John J., "Categorical perception of facial expressions", *Cognition*, Sep Vol 44(3) 227-240, 1992.
- [6] C.Breazeal, "Designing Sociable Robots", Cambridge, MA: The MIT Press, Cambridge, 2002.
- [7] H. Miwa, T. Okuchi, K. Itoh, H. Takanobu, and A. Takanishi, "A New Mental Model for Humanoid Robots for Human Friendly Communication", *Proceedings of the 2003 IEEE International Conference on Robotics & Automation Taipei, Taiwan*, September 14-19, 2003.
- [8] K. Kanoh, S. Kato, and H. Itho, "Facial expressions using emotional space in sensitivity communication

- robot 'Ifbot', in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., Sendai, Japan, 2004, pp. 1568–1591.
- [9] F. Kawakami, M. Okura, H. Yamada, H. Harashima, S. Morishima, "An evaluation of 3-D emotion space", in Proceedings 4th IEEE International Conference on Robot & Human Interactive Communication, Jul 5, 1995.
- [10] C. Bartneck and J. Reichenbach, "Subtle emotional expressions of synthetic characters," Int. J. Human-Comput. Stud., vol. 62, pp. 179–193, 2005.