

신경망을 이용한 얼굴 표정인식

신영숙 이일병

연세대학교 인지과학과

서울시 서대문구 신촌동 134

E-mail: hellogen@csai.yonsei.ac.kr yblee@csai.yonsei.ac.kr

Recognizing Facial Expressions Using a Neural Network

Youngsuk Shin Yillbyung Lee

Dept. of Cognitive Science, The Graduate School Yonsei University

134, Shinchon-Dong, Seodaemun-Ku, Seoul

E-mail: hellogen@csai.yonsei.ac.kr yblee@csai.yonsei.ac.kr

요 약

기존의 표정인식 연구는 Ekman의 기본 정서모형의 틀에 의하여 표정인식이 이루어져왔다. 그러나 이러한 6가지 기본정서(행복, 놀람, 공포, 분노, 혐오, 슬픔)에 의한 표정인식은 6개 정서 중에서 선택하는 강제 선택법이 아닌 자유응답방식을 택했을 때는 훨씬 인식률이 떨어진다. 이는 표정이 기본정서이외에도 여러 가지 미묘한 마음상태를 표현하고 있기 때문이다. 본 연구는 기본정서모형의 한계성을 극복하고 섬세한 표정인식을 위한 방법으로, 차원모형을 근거로 MLP를 적용한 표정인식을 수행하였다.

차원 모형에 의한 표정은 3가지 차원으로 하나의 표정을 이룬다. 3가지 차원은 쾌-불쾌, 각성-수면과 외부지향-내부지향이다. 3가지 차원을 갖는 각각의 표정은 MLP에 의하여 쾌-불쾌차원 68%, 각성-수면차원 60%, 외부지향-내부지향차원 76%의 인식률을 보였다. 연구결과에서 차원모형에 근거한 표정인식을 통하여 기존의 기본정서모형의 한계성을 극복하고 섬세한 표정인식을 수행할 수 있었다.

1. 서 론

기계가 인간의 표정으로부터 인간의 상황을 이해할 수 있다면 감지된 상황에 따라 능동적으로 인간을 도울 수 있을 것이다. 이것은 기계와 인간이 마음에서 마음으로 통신을 실현하는 것으로 능동적인 인간 인터페이스 개발이라고 할 수 있다. 이러한 능동적인 인간 인터페이스의 개발은 인간의 얼굴표정의 연구로부터 시작된다고 할 수 있다. 인간은 얼굴표정을 인식할 수 있는 뛰어난 능력을 갖는다. 심리학 분야에서는 얼굴 분석과 인식은 수십 년간 연구되었다. Ekman[1]의 연구에 따르면 표정은 행복, 놀람, 공포, 분노, 혐오, 슬픔의 6가지로 범주화된다. 이는 Facial Action Coding System(FACS)[2]을 토대로 한 것으로서 얼굴근육의 움직임을 상세하게 기록하여 여섯가지 표정으로 분류할 수 있도록 구성되어 있다. 이러한 얼굴표정에 대한 초기의 연구는 연구자의 주관적인 판단이나 일반 사람들의 평정자료를 근거로 하여 얼굴표정으로 표현되는 경향의 연구였다. 최근에는 얼굴근육의 움직임을 측정하여 얼굴표정을 내

적상태의 표현과 연결시키는 작업들이 진행되었다. 이는 내적 상태에서부터 얼굴표정으로의 대응관계를 조사한 한재현[3]의 연구에서 나타나고 있다. 이 논문에서 내적 상태는 얼굴표현으로 발견되며 구별되는 두 내적 상태는 서로 다른 얼굴표정으로 나타난다고 소개하고 있으며, 얼굴표정으로부터 내적 상태를 추정하는 과정을 밝히기 위하여 Pilowsky 등[4]의 얼굴특징요소들의 12차원을 사용하고 있다. 특히 Pilowsky는 특정 내적 상태를 인식하는 요소로서 얼굴표정 공간의 12차원을 사용하고 있는데 이것은 얼굴표정을 수량화하기 위하여 얼굴의 눈, 코, 입간의 기하학적인 요소들을 특징으로 한 것이다. 기하학적인 특징을 이용한 표정연구는 Kobayashi 등[5]의 연구와 신영숙 등[6]의 연구에서도 밝히고 있다. 그러나 이러한 기존의 연구들은 Ekman의 기본정서모형을 근거로 한, 제한된 상황에서 비교적 용이한 한정된 정서만을 대상으로 한 연구들이었다. 따라서 섬세한 표정을 인식하는 데는 기본적인 한계가 있을 것으로 보인다. 본 연구는 이러한 기본정서모형의 한계를 극복하고, 자연스런 표정을 인식하기 위한 방법으로 차원모형을 기반으로 하여 신경망을 적용한 표정인식을 수행하였다.

2. 표정정보

얼굴의 특징정보는 Pilowsky의 얼굴표정 공간에 대한 12차원을 적용하였다.[그림1]. [그림1]에서 3번째 그림의 수직과 수평성분은 기하학적인 얼굴 특징 요소들의 정규화를 위하여 사용된 것으로 수평성분의 특징들은 두 눈 사이의 거리 값으로 정규화하고 수직성분의 값들은 한쪽 눈꼬리에서 콧구멍까지의 길이로 정규화하였다. 이는 Pilowsky의 수직성분의 정규화와는 차이를 보인다.

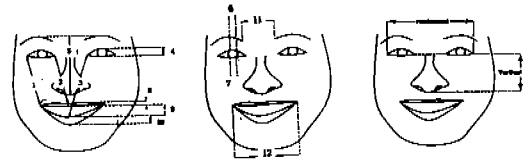


그림1. Pilowsky[4]의 얼굴표정 12차원

3. 신경망을 이용한 표정인식

3.1 차원모형

표정은 내적 상태를 기술하는 83개의 단어를 30, 40대 피험자 150명을 대상으로 실험을 통하여 김영아, 김진관, 박수경, 오경자[7]의 정서관련 어휘분석을 통한 내적 상태 차원 연구에 의하여 3가지 차원으로 구별되어진다. 각각의 표정들은 3가지 차원인 쾌-불쾌차원, 각성-수면차원, 외부지향-내부지향차원으로 하나의 표정을 표현한다. 이러한 차원모형은 기존의 표정연구가 6가지 기본정서모형을 토대로 한정된 정서만을 대상으로 이루어졌던 한계성을 극복한 것이다. 기본정서모형의 정서상태들이 서로 독립적인 데 반하여 차원모형은 정서상태들이 서로 독립적이지 않으며 매우 체계적인 방식으로 서로 관련되어 있다. 이러한 차원모형은 특정 정서를 나타내는 것이 아니라, 많은 서로 다른 정서들에 공통된 일반적 특징을 나타낸다. 또한 차원상에서 정서들이 배열되는 형태에서 범주적 속성이 나타나기도 하는데, 어떤 정서들은 다른 정서들보다 서로 보다 밀접하게 관련되어 있음을 보인다. 서로 직각을 이루는 두 개의 양극성 차원의 공간에서 이러한 몇가지 정서의 그룹들이 원모양으로 배열된다고 하여 원형모형이라고도 한다. [그림2]는 두가지 차원(쾌-불쾌와 각성수준)의 차원모형을 표현한다. 세가지 차원들은 7점 척도상에서 표정데이터를 평정하도록 나타내진다.

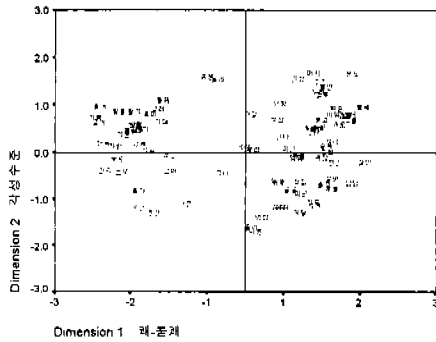


그림2. 83개 기초목록의 차원분석[7]

3.2 학습

얼굴표정과 내적 상태의 연결관계를 탐색하기 위해 인간 피험자를 대상으로 얻어진 데이터를 기반으로 MLP를 이용하여 표정인식을 위한 학습을 수행하였다. [그림3]은 표정인식에 사용된 3층 신경망 모델을 나타낸다.

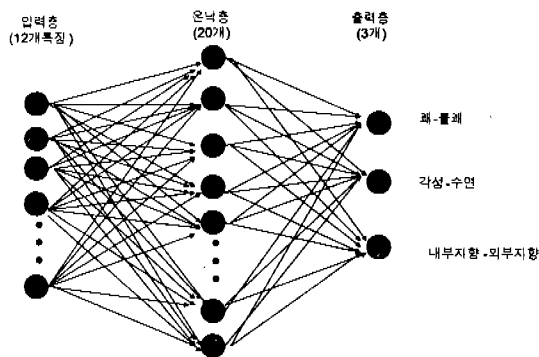


그림3. 구현된 신경망

신경망 학습방법은 BP(역전파 학습 방법) 알고리즘을 사용하였다. 이 방법의 학습은 훈련되어 출력되는 값과 목표 값과의 오류를 은닉층에 역전파하여 수행하는 것으로 이 과정에서 연결 가중치를 조정한다.

영상 데이터에서 Pilowsky의 12차원의 특징정보들이 입력층의 노드를 이룬다. 은닉노드의 수는 20개이며 출력노드의 수는 3가지 표정차원을 나타내기 위하여 3개의 노드로 이루어진다. 신경망 훈련을 위하여 100개의 얼굴표정이 사용되었으며, 훈련에 사용되지 않은 38개의 얼굴표

정이 인식을 위하여 사용되었다. 우선, 신경망의 객관성을 검증하기 위하여 신경망의 훈련에 사용되었던 100개의 표정 데이터를 인식 실험 데이터로 사용하였다.

3.3 인식

얼굴표정에 의한 내적 상태 추정연구[3]에서 얻어진 사람의 평정결과와 신경망의 실험 결과의 오차 0.00001 수준에서 100%에 가까운 인식률을 보였다. 3가지 표정 차원에서 사람의 평정과 신경망 실험결과를 아래의 [그림4], [그림5], [그림6]에서 각 차원별로 비교되어 표현된다.

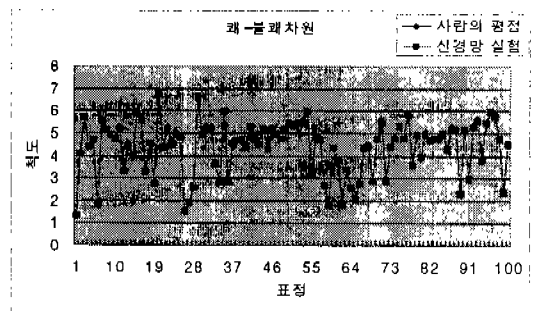


그림4.

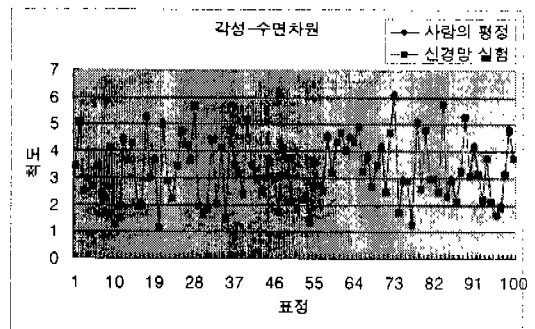


그림5

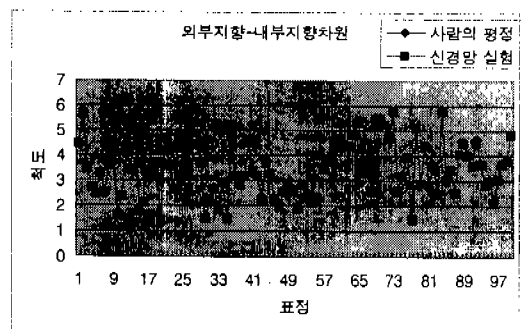


그림6

학습데이터를 훈련한 후 학습에 사용되지 않은 표정 데이터 38개를 사용하여 인식 실험을 수행하였다. 인식실험 결과 쾌-불쾌차원 68%(26/38), 각성-수면차원 60%(23/38), 외부지향-내부지향차원 76%(29/38)을 보였다. 인식수준의 임계치는 사람의 평정결과의 80%~120%를 인식의 정도로 설정하여 인식률을 측정하였다. 아래의 [그림7], [그림8], [그림9]은 표정차원별로 사람의 평정과 신경망 실험 결과를 보인다.

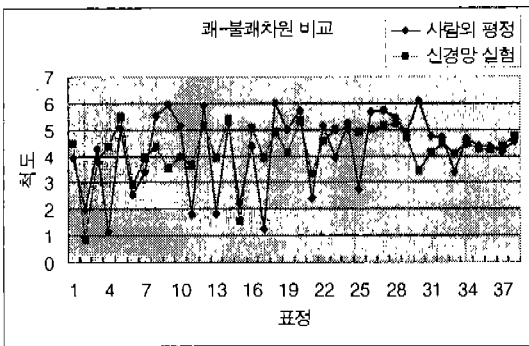


그림7

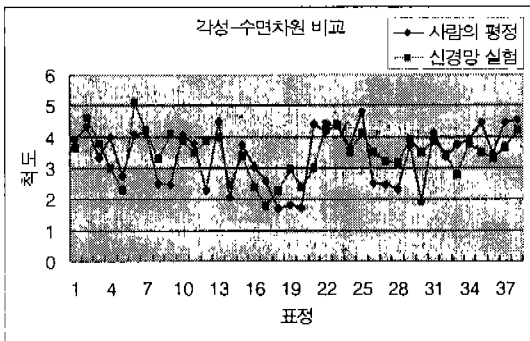


그림8

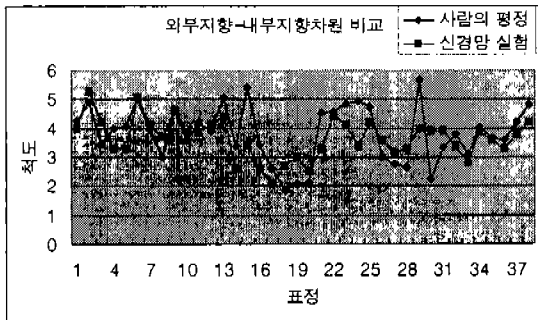


그림9

인식된 결과는 다음과 같이 내적 상태 차원 공간상에 표정이 어느 공간상에 있는지 나타내진다. [그림11]에서 별표(*)는 [그림10]을 두 가지차원(쾌-불쾌, 각성-수면)의 표정공간상에 신경망에 의하여 나타낸 인식결과이다. 위의 38개 표정 데이터 인식결과는 오차 0.01309에서 측정하였다. 학습데이터의 오차수준에서의 인식률은 0.01309 오차수준보다 낮은 인식률을 보였다.

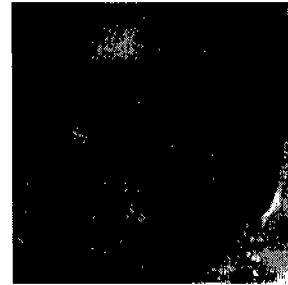


그림10

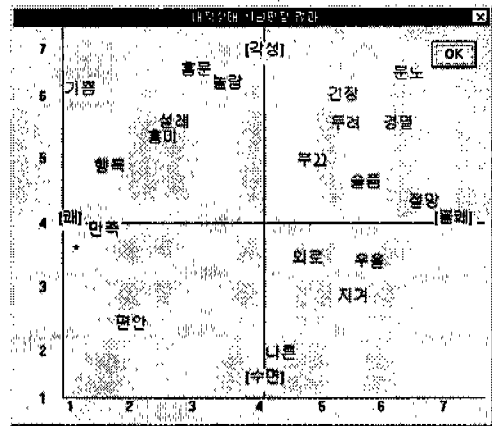


그림11

이러한 결과는 학습에 사용된 표정데이터들이 국소적인 경향을 띤 경우로 판단되며, 보다 다양한 얼굴 데이터를 기반으로 한 학습의 필요성을 시사하고 있다.

4. 결론

본 연구는 MLP를 사용하여 표정에 대한 내적상태 차원모형에서 표정공간으로의 표정인식을 수행하였다. 내적상태 차원공간상에 나타내지는 표정을 통하여 Ekman

의 6가지 기본정서모형에 의한 한정된 표정 인식을 극복하고 섬세한 표정인식을 위한 토대를 마련하였다. 향후 다양한 얼굴의 학습을 통하여 현재의 표정 인식에서 보다 일반화된 표정인식을 지향할 계획이며, 표정 특징정보들의 자동화된 추출은 계속 연구되어야 할 것이다.

pp.381-386(1992).

- [6] 신영숙, 이일병. “기하학적인 특징 추출을 이용한 얼굴표정 인식”. '97 한국감성과학회 연차학술대회 논문집,1997.
- [7] 김영아, 김진관, 박수경, 오경자. “내적 정서 관련 어휘분석을 통한 내적상태의 차원 연구”. '97 한국감성과학회 연차학술대회 논문집,1997.

참고문헌

- [1] Ekman, P.(Ed). (1982.) Emotion in the human face(2nd ed.). Cambridge, England:University of Cambridge Press.
- [2] P. Ekman and W. Friesen. "Facial Action Coding System". Consulting Psychologists, Palo Alto, CA, 1977.
- [3] 한재현. “얼굴표정에 대한 내적상태 추정”. 석사학위 논문, 1997.
- [4] Pilowsky, I., Thornton, M., & Stokes, B. B.(1996). Towards the quantification of facial expressions with the use of a mathematic model of the face. In H.D. Ellis, M. A. Jeeves FRSE,F. Newcombe & A. Young(Eds.), Aspects of face processing. Dordrecht: Martinus Nijhoff Publishers.
- [5] H.Kobayashi and F.Hara, "Recognition of six basic facial expressions and their strength by neutral network", Proc. of IEEE International Workshop on Robot and Human Communication,