

## 뇌전도 신호를 이용한 실시간 감정변화 인식 기법

최동윤, 이상혁, 송병철  
인하대학교

[pride0723@gmail.com](mailto:pride0723@gmail.com), [xxxblanky@gmail.com](mailto:xxxblanky@gmail.com), [\\*bcsong@inha.ac.kr](mailto:*bcsong@inha.ac.kr)

### Real-Time Emotional Change Recognition Technique using EEG signal

Dong Yoon Choi, Sang Hyuk Lee, \*Byung Cheol Song  
Inha university

#### 요 약

감정인식 기술은 인간과 인공지능이 감정적인 상호작용을 위하여 매우 중요한 기술이다. 얼굴영상 기반의 감정인식 연구는 가장 널리 진행되어 왔으나 우리는 표정에서 드러나지 않는 내면의 감정을 인식하기 위하여 뇌전도를 이용한 감정인식 기법을 제안한다. 먼저 2 초 구간의 뇌전도 신호에 대하여 time, frequency, time-frequency 영역에서 특징점을 추출하고 이를 3 개의 fully connected layer 로 구성되어 있는 regressor 를 이용하여 valence 정보를 추정한다. MAHNOB-HCI 데이터세트에 대한 실험결과에서 제안기법은 종래기법보다 낮은 오차를 보이며 감정의 변화를 실시간으로 인식하는 결과를 보인다.

#### 1. 서론

인공지능과 인간과 감정적인 상호작용을 위해 중요한 기술이다. 감정인식은 인공지능 분야에서 널리 연구되고 있으며 그 중 얼굴 영상을 이용한 컴퓨터 비전기반 연구가 가장 활발하다. 그러나 컴퓨터 비전 분야에서는 제시된 감정에 맞추어 표정을 연기하거나 다소 큰 감정반응에 대한 영상을 수집한 데이터세트를 널리 이용한다. 그러나 실제 인간이 자연스럽게 나타내는 감정의 경우 표정을 크게 표현하지 않는 경우가 대부분이다. 실제 피험자에게 감정을 유발하는 콘텐츠를 제공함과 동시에 느낀 감정의 정도를 확인하는 경우 피험자는 감정을 크게 느꼈지만 표정에서는 크게 드러나지 않는 경우가 많다.

따라서 본 논문에서는 표정에서 드러나지 않는 인간의 감정을 인식하기 위하여 뇌전도 신호(electroencephalography, EEG)를 이용한 감정인식을 수행한다. 뇌전도 신호는 일반적으로 12 개에서 32 개의 전극을 머리에 접촉하여 각 전극 즉 채널마다 128~256Hz 의 1 차원 신호를 취득한다. 실제 감정반응에 대한 뇌전도 신호를 취득한 데이터세트 및 연구사례가 존재한다. DEAP[1]와 MAHNOB-HCI[2] 데이터세트는 대표적인 감정인식용 뇌전도 데이터 세트로 피험자에게 감정유발 영상을 제공함과 동시에 뇌전도를 측정하였다. 두 데이터 모두 정교한 감정인식을 위해 arousal/valence 영역에서 연속적인 값을 갖는 감정 label 를 생성하였다. 한편 DEAP 데이터세트[1]는 가장 널리 사용하는 데이터세트이지만 1 분 동안 피험자가 느낀 평균적인 감정정보가 label 로 존재하기 때문에 실시간 감정인식을 수행하는데 적합하지 않다. 따라서 본 논문에서는 4Hz 단위로 valence 정보가 annotation 되어 있는 MAHNOB-HCI dataset [2]을 이용한다.



그림 1. 제안기법의 구성도

본 논문에서는 뇌전도 신호에 대하여 여러 domain 에서 특징점을 추출하고 딥러닝 기반의 알고리즘을 이용하여 실시간 감정인식을 수행하였다. 실험결과에서는 제안기법을 통해 실시간으로 감정의 변화를 유효하게 인식하는 것을 확인하였다.

#### 2. 제안 기법

본 논문의 제안기법은 그림 1 과 같이 이루어 진다. 먼저 뇌전도 신호에 대하여 특징점을 추출하고 MLP 기반 regressor 를 통해 긍정/부정의 강도인 valence 정보를 출력한다.

##### 2.1 특징점 추출

MAHNOB-HCI 데이터세트 [2]의 뇌전도 신호는 1 차원 신호가 32 채널로 존재한다. 우리는 뇌전도 신호를 딥러닝 알고리즘의 입력으로 사용하기 위하여 특징점 추출과정을 진행한다. 현재 시점을 기준으로 이전의 2 초 동안의 뇌전도 신호에 대하여 time, frequency, time-frequency 영역에서의 특징점을 추출한다. Time 영역의 경우 신호 강도에 대한 최대값, 평균, 표준편차 등을 이용하여 frequency domain 의

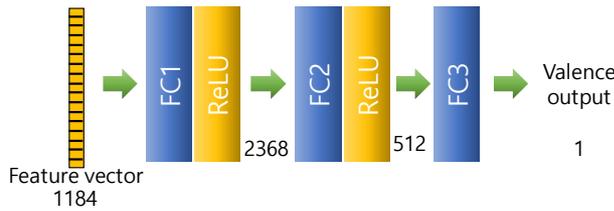


그림 2. MLP 기반 regressor

경우 주파수 대역별로 power spectral density(PSD)를 추출하여 이용한다. 마지막으로 time-frequency 영역에서는 discrete wavelet transform(DWT)를 이용하여 특징점을 추출한다. 우리는 각 채널별로 37 개의 특징점을 추출하였으며 32 채널에 대하여 총 1184 의 길이를 갖는 특징점 vector 로 구성하여 딥러닝 알고리즘의 입력으로 이용한다.

### 2.2 MLP 기반 regressor

뇌전도 신호에 대하여 추출된 특징점 vector 에 대하여 여러 개의 fully connected layer(FC)로 구성된 multi-layer perceptron(MLP) 기반의 regressor 를 통해 valence 정보를 추정한다. Regressor 에 대한 구조는 그림 2.와 같다. 초기 1184 길이의 vector 가 FC1 을 거쳐 2368 길이로 확장되며 FC2 를 통해 512 길이로 감소된 후 FC3 을 통해 최종적으로 valence 값을 출력한다. FC layer 이후에는 relu 연산자가 추종간에는 network 의 overfitting 을 방지하기 위하여 batch normalization 과 dropout 과정을 추가하였다. 또한 network 의 학습성능을 위하여 He normalization 을 통하여 network weight parameter 를 초기화 하였다. 또한 training loss 는 network 에서 출력되는 valence 값과 ground truth 의 mean squared error(MSE)를 측정하여 이용하였다.

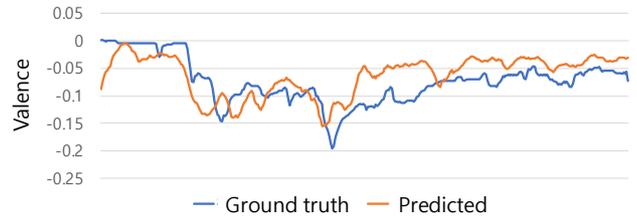
### 3. 실험 결과

본 절에서는 MAHNOB-HCI 데이터세트[2]에 대한 제안기법의 성능을 확인한다. 해당 데이터세트는 239 개의 video clip 과 뇌전도 신호 데이터가 존재한다. 각 데이터는 4Hz 단위로 valence 정보가 annotation 되어 있으며 총 75000 여개의 time step 에 대한 valence 정보가 존재한다. 해당 논문의 경우 학습데이터와 테스트데이터가 별도로 나누어져 있지 않다. 해당 데이터세트를 이용한 비교 기법의 경우 10-fold cross validation 을 수행하여 뇌전도 기반의 감정인식 성능을 확인하였다. N-fold cross validation 은 전체 데이터세트를 N 개의 서브세트로 나누고 하나의 그룹은 테스트 데이터로, 나머지는 학습데이터로 이용하여 모델의 성능을 평가한다. 이때 N 개의 각 서브세트를 테스트데이터로 이용하여 총 N 회에 대한 성능의 평균을 측정한다. 제안 기법의 MLP 기반 regressor 를 학습 시 pytorch library 를 이용하였고 batch 의 크기는 256 으로 하였다. Dropout 의 비율은 0.5 로 지정하였으며 Adam optimizer 를 이용하였다.

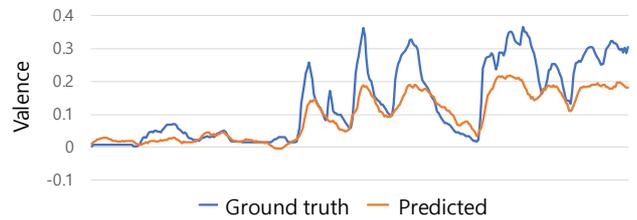
먼저 제안기법을 통한 감정정보의 정확도를 확인하기 위해 root mean squared error(RMSE)를 통해 제안기법의 정확도를 측정하였다. 10-fold validation 을 수행하였기 때문에 10 회 실험 결과에 대한 평균과 표준편차를 측정하였다. 종래

표 1. RMSE 측정 결과 비교

|      | [2]           | 제안기법          |
|------|---------------|---------------|
| RMSE | 0.0530±0.0290 | 0.0491±0.0042 |



(a)



(b)

그림 3. 시간에 따른 valence 추정 결과 (a) 부정적인 감정 발생 상황 예시 (b) 긍정적인 감정 발생 상황 예시

연구인 Soleymani 의 기법[2]과 비교하면 0.0039 만큼 낮은 RMSE 를 보이며 보다 정확한 결과를 확인할 수 있다.

제안기법을 통해 추정된 감정정보를 ground truth 와 함께 그래프로 나타낸 그림은 그림 3.과 같다. 제안기법은 부정적인 감정과 긍정적인 감정 모두에서 감정의 변화를 잘 나타내는 것을 확인할 수 있다. 그림 3(b)에서 감정의 변화 경향은 ground truth 와 유사하나 감정의 강도에 오차가 있는 것을 확인할 수 있다. 이는 ground truth 의 label 값이 인간이 표정을 보고 강도를 결정하였기 때문에 실제 EEG 로 인한 감정의 강도와 일치하지 않기 때문이다.

### 2. 결론

본 논문에서는 뇌전도 신호를 이용하여 감정의 변화를 실시간으로 인식하는 기법을 제안한다. 제안기법은 종래기법 대비 낮은 평균오차를 보이고 긍정적인 감정과 부정적인 감정 모든 상황에서 유효한 성능을 확인할 수 있다. 향후 연구로는 뇌전도 신호와 얼굴 영상정보를 동시에 이용하여 단일 신호 대비 감정인식 성능의 시너지가 발생하도록 하는 연구를 진행할 계획이다.

### 감사의 글

본 논문은 산업통상자원부의 산업기술혁신사업 [10073154, 인간 내면상태의 인식 및 이를 이용한 인간친화형 인간-로봇 상호작용 기술 개발] 및 산업용 지능융합 부품기술 개발사업 [과제명: 4K30p 급 Deep Learning 기반 Edge Computing IP 카메라용 시스템반도체 개발]의 연구비

지원으로 이루어짐.

### 참고문헌

- [1] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, and I. Patras, "Deap: A database for emotion analysis using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3 no. 1, pp. 18–31, Jan. 2012
- [2] M. Soleymani, S. Asghari-Esfeden, Y. Fu, and M. Pantic. "Analysis of EEG signals and facial expressions for continuous emotion detection," *IEEE Transactions on Affective Computing*, vol. 7 no. 1, pp. 17–28, Mar. 2016.
- [3] He, K., Zhang, X., Ren, S., & Sun, J. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," In *Proceedings of the IEEE international conference on computer vision*, pp. 1026–1034. 2015.