



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0087162
(43) 공개일자 2024년06월19일

(51) 국제특허분류(Int. Cl.)
A61B 5/16 (2006.01) A61B 5/00 (2021.01)
A61B 5/024 (2006.01) A61B 5/0533 (2021.01)
(52) CPC특허분류
A61B 5/165 (2013.01)
A61B 5/02416 (2013.01)
(21) 출원번호 10-2022-0172584
(22) 출원일자 2022년12월12일
심사청구일자 2022년12월12일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
김덕환
서울특별시 양천구 목동로 230, 103동 1004호(목동, 아이파크아파트)
강동현
경상북도 포항시 북구 중흥로139번길 23-1, 1층(죽도동)
(74) 대리인
양성보

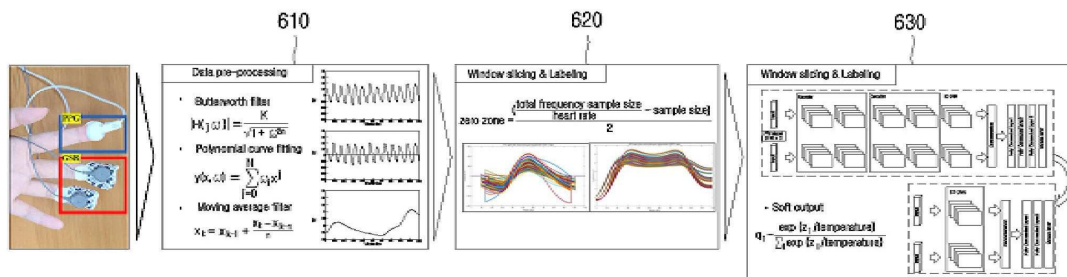
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 방법 및 시스템

(57) 요약

실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 방법 및 시스템이 제시된다. 본 발명에서 제안하는 실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 시스템은 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출하는 전처리부, 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플 사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행하는 라벨링부 및 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습하는 학습부를 포함한다.

대표도



(52) CPC특허분류

A61B 5/0533 (2021.01)

A61B 5/7207 (2013.01)

A61B 5/7267 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711160589
과제번호	2022-0-00641-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	방송통신산업기술개발 사업
연구과제명	[Ezbaro] XVoice: 멀티모달 음성 메타학습(1차년도)
기 여 율	1/2
과제수행기관명	인하대학교 산학협력단
연구기간	2022.04.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711153997
과제번호	2021R1F1A1050750
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	기본연구
연구과제명	[Ezbaro] IoT 데이터를 위한 연합 심층학습을 통한 블록체인지향 분산 엣지 컴퓨팅
솔루션에 대한 연구	
기 여 율	1/2
과제수행기관명	인하대학교
연구기간	2022.03.01 ~ 2023.02.28

명세서

청구범위

청구항 1

실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출하는 전처리부;

상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플 사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행하는 라벨링부; 및

상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습하는 학습부

를 포함하는 감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 시스템.

청구항 2

제1항에 있어서,

상기 전처리부는,

PPG 신호에 대하여 선형 주파수 영역의 필터를 적용하고, 고차다항식 적합(High order Polynomial)으로 데이터의 기저선 변동 잡음과 이동평균 필터(Moving Average Filter)을 통해 동적 잡음을 감소시키고,

GSR 신호에 대하여 PPG 신호 기준으로 분할한 후 동일한 전처리를 진행하는

감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 시스템.

청구항 3

제1항에 있어서,

상기 라벨링부는,

심장박동 수를 이용하여 중복 방지를 위해 하나의 파형 구간 외에는 제로 존(zero zone)으로 처리하고,

상기 추출된 파형 단위의 PPG 신호 및 GSR 신호의 라벨링은 상기 PPG 신호의 첨두치를 기준으로 진행하고,

추출 및 전처리된 데이터와 라벨링을 통해 복수의 세그먼트를 추출하여 라벨링 범위에 따라 복수의 클래스로 분류하는

감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 시스템.

청구항 4

제1항에 있어서,

상기 학습부는,

상기 PPG 신호 및 GSR 신호 기반의 일차원 컨볼루션 오토인코더를 실시간 프로그램 및 하드웨어 성능이 제한된 환경에서 적용하기 위해 지식 증류 기법을 적용하여, 정확도를 유지하면서 네트워크 압축, 매개변수 감소, 연산 소요 시간을 감소시키는

감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 시스템.

청구항 5

전처리부를 통해 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출하는 단계;

라벨링부를 통해 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플

사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행하는 단계; 및
 학습부를 통해 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습하는 단계
 를 포함하는 감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 방법.

청구항 6

제5항에 있어서,
 상기 전처리를 통해 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출하는 단계는,
 PPG 신호에 대하여 선형 주파수 영역의 필터를 적용하고, 고차다항식 적합(High order Polynomial)으로 데이터의 기저선 변동 잡음과 이동평균 필터(Moving Average Filter)를 통해 동적 잡음을 감소시키고,
 GSR 신호에 대하여 PPG 신호 기준으로 분할한 후 동일한 전처리를 진행하는
 감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 방법.

청구항 7

제5항에 있어서,
 상기 라벨링부를 통해 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 침두치를 기준으로 샘플 사이즈가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행하는 단계는,
 심장박동 수를 이용하여 중복 방지를 위해 하나의 파형 구간 외에는 제로 존(zero zone)으로 처리하고,
 상기 추출된 파형 단위의 PPG 신호 및 GSR 신호의 라벨링은 상기 PPG 신호의 침두치를 기준으로 진행하고,
 추출 및 전처리된 데이터와 라벨링을 통해 복수의 세그먼트를 추출하여 라벨링 범위에 따라 복수의 클래스로 분류하는
 감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 방법.

청구항 8

제5항에 있어서,
 상기 학습부를 통해 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더를 사용하여 감정분류를 위한 모델을 설정하고 학습하는 단계는,
 상기 PPG 신호 및 GSR 신호 기반의 1D 일차원 컨볼루션 오토인코더를 실시간 프로그램 및 하드웨어 성능이 제한된 환경에서 적용하기 위해 지식 증류 기법을 적용하여, 정확도를 유지하면서 네트워크 압축, 매개변수 감소, 연산 소요 시간을 감소시키는
 감정인식 및 분류를 위한 일차원 컨볼루션 오토인코더 방법.

발명의 설명

기술 분야

[0001] 본 발명은 실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 효율적인 인간-로봇의 상호작용을 위해 실시간 감정인식 및 분류는 필수적인 사항이다. 일반적으로 인간의 감정은 외부적인 영향과 복잡한 생각, 신체적인 특징에 민감하게 반응한다. 따라서 취득된 데이터는 정보 손실이 없고, 유지할 수 있는 전처리 과정, 빠르게 결과를 도출할 수 있는 작은 크기의 모델 설정이 필요로 한다. 표면적으로 나타나는 인간의 표정, 목소리 신호에 대한 실시간 감정인식 기반 연구는 활발히 진행되고 있지만, 표정과

목소리는 인간이 제어 가능한 체성 신경계 신호(Somatic Nervous System; SNS)로써 정확한 내면 감정인식으로 볼 수가 없다. 따라서 인간이 제어하지 못하는 자율 신경계 신호(Auto-nomic Nervous System; ANS)를 사용하여 정확한 내면 감정을 인식할 수 있다. 인간의 내면 감정을 이용한다면, 다양한 분야에서 인간-로봇의 상호작용 연구 및 개발, 적용이 가능하다. 복잡한 상관관계가 있는 감정은 내부와 외부의 환경적인 요소와 개인적인 특성 때문에 감정인식 및 분류의 정확도에 영향을 미친다. 따라서 생체신호(ANS)와 표정, 목소리 데이터(SNS)의 융합 기술을 사용하면, 내부와 외부의 감정인식 및 분류 성능 개선이 가능하다. 즉, 강력한 감정인식 모델을 구축할 수 있다. 예를 들어 정확한 내면 감정인식 기술은 복잡한 비행기 조종을 위한 훈련에서 조종사의 정서적인 생체 신호에 따른 개별적인 기준을 제공할 수 있으며, 환자들의 재활치료와 같은 의료 산업에 적용할 수 있다.

[0003] 다양한 생체신호를 정보 손실 없이 취득하기 위해 침습형과 비침습형 방법이 있다. 침습형은 피험자의 신체에 직접 센서와 바늘을 삽입하여, 정확한 신호를 측정할 수 있지만, 부작용과 안전성에 문제가 생긴다. 따라서 주로 사용되는 방법은 피험자의 피부에 전극과 센서를 부착하는 비침습형 방식이다.

[0005] 도 1은 종래기술에 따른 측정 위치 및 유형 기반 자율신경계의 생체신호 분류를 설명하기 위한 도면이다.

[0006] 인간의 피부에서 취득이 가능한 ANS 신호는 도 1과 같이 취득되는 신체 부위와 취득 방법에 따라 신호의 종류가 다르다. 예를 들어, 인간의 피부에서 취득이 가능한 ANS 신호는 뇌의 전기적 활동을 나타내는 뇌전도 (EEG, electroencephalogram), 심장의 전기적 활동을 측정한 심전도 (ECG, electro-cardiogram), 혈관에서 반사된 광도를 측정한 광혈류량 (PPG, photoplethysmogram), 땀 배출량과 체온 변화를 측정한 피부 전도도 (GSR, Galvanic Skin Response), 골격근에서 발생하는 전기적인 신호를 측정한 근전도 (EMG, electromyography) 등이 있다.

[0007] 위와 같이 각각의 생체신호를 측정하는 방법에 따라 사용자의 편의성, 상용화 여부, 신호가 가지는 특정한 규칙성 여부가 다르게 나타난다. 또한 다양한 생체신호를 취득하기 위해 사용되는 전극의 개수와 위치에 따라 사용자의 사용성과 효율성이 크게 변할 수 있다. 신호가 가지는 정보와 활용 방법이 다양하므로 많은 시도와 연구가 다양한 관점에서 진행 중이다. EEG 신호는 최대 64개의 채널로 많은 양의 전극을 사용하며, 복잡한 정보를 포함하고 있다. 따라서 감정 정보에 대한 특징을 추출하기 위해, Mantini Dante et al은 EEG 신호의 파워 스펙트럼 (Power Spectrum Density; PSD)를 이용하였고, Ante Topic et al은 Topography를 사용하여 감정 정보 특징들을 추출하였다. 많은 전극에서 수집되는 정보와 데이터를 기반으로 높은 감정분류 정확도를 나타내고 있지만 데이터 처리에 따른 소요 시간, 사용자 편의성 저하로 실시간 및 일상생활에서의 사용이 어렵다. 심장의 전기적인 활동을 측정한 ECG 신호는 혈압, 심박수 등의 정보를 포함하고 있다. Maciej Golgowski et al은 wavelet transformation을 이용하여 심장의 이상감지를 연구하였으며, Cai Jing et al은 이산 웨이블릿 변환(Discrete Wavelet Transform; DWT)과 TS(Tabu Search) 알고리즘을 통해 감정을 분류하였다. ECG 신호는 감정에 따른 심장 변화 기반의 민감성과 규칙성을 포함하고 있지만, 심장 부근 가슴 근처에 최대 6개의 전극을 부착해야 하는 불편함이 있다. PPG 신호는 체적 변동기록으로 감정의 변화를 혈관에서 반사된 광도를 기반으로 혈류량 변화, 규칙적인 심장박동 정보를 포함하고 있다. Raj Rakshit et al은 PPG 신호에서 심장변이도(Heart Rate Variability; HRV)를 추출하여 감정분류 하였다. Min Seop Lee et al은 짧은 길이의 PPG 신호를 간단한 전처리와 1D CNN(Convolutional Neural Networks) 모델을 이용하여 빠른 감정분류를 진행하였다. GSR 신호는 감정 변화에 따른 땀 배출량과 체온변화 정보를 포함하고 있으며, 육체적인 영역에 밀접한 상관관계가 있다. Nagarajan Ganapathy et al은 푸리에 변환을 이용하여 시간-주파수 영역의 38가지 특징 추출 후 CNN 모델에 적용하였고, Imam Yogie Susanto et al은 1D CNN과 Residual Bidirectional GRU를 사용하여 감정분류를 진행하였다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) [1] Mert, Ahmet, and Aydin Akan. "Emotion recognition from EEG signals by using multivariate empirical mode decomposition." Pattern Analysis and Applications 21.1 (2018): 81-89.

(특허문헌 0002) [2] Lee, Min Seop, et al. "Fast emotion recognition based on single pulse PPG signal with convolutional neural network." Applied Sciences 9.16 (2019): 3355.

(특허문헌 0003) [3] Chao, Hao, et al. "Emotion recognition from multiband EEG signals using CapsNet." Sensors 19.9 (2019): 2212.

(특허문헌 0004) [4] Xing, Xiaofen, et al. "SAE+ LSTM: A New framework for emotion recognition from multi-channel EEG." *Frontiers in neurorobotics* 13 (2019): 37.

(특허문헌 0005) [5] Lee, MinSeop, et al. "Emotion recognition using convolutional neural network with selected statistical photoplethysmogram features." *Applied Sciences* 10.10 (2020): 3501.

(특허문헌 0006) [6] Raheel, Aasim, Muhammad Majid, and Syed Muhammad Anwar. "Dear-mulsemmedia: Dataset for emotion analysis and recognition in response to multiple sensorial media." *Information Fusion* 65 (2021): 37-49.

(특허문헌 0007) [7] Hwang, Woo-Hyun, Dong-Hyun Kang, and Deok-Hwan Kim. "Brain lateralisation feature extraction and ant colony optimisation-bidirectional LSTM network model for emotion recognition." *IET Signal Processing* 16.1 (2022): 45-61.

발명의 내용

해결하려는 과제

[0009] 본 발명이 이루고자 하는 기술적 과제는 PPG 신호와 GSR 신호의 생체신호를 사용하여, 감정 분류의 정확성과 안정성, 사용자의 편의성 개선이 가능한 실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 방법 및 시스템을 제안한다. 또한, 일차원 컨볼루션 오토인코더에 대해 지식 증류 기법을 적용하여, 정확도를 유지하면서 네트워크 압축, 매개변수 감소, 연산 소요 시간을 감소시킴으로써 실시간 및 제한된 하드웨어 환경에서도 사용이 가능한 모델을 제안한다.

과제의 해결 수단

[0010] 일 측면에 있어서, 본 발명에서 제안하는 실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 시스템은 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출하는 전처리부, 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플 사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행하는 라벨링부 및 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습하는 학습부를 포함한다.

[0011] 상기 전처리부는 PPG 신호에 대하여 선형 주파수 영역의 필터를 적용하고, 고차다항식 적합(High order Polynomial)으로 데이터의 기저선 변동 잡음과 이동평균 필터(Moving Average Filter)을 통해 동적 잡음을 감소시키고, GSR 신호에 대하여 PPG 신호 기준으로 분할한 후 동일한 전처리를 진행한다.

[0012] 상기 라벨링부는 심장박동 수를 이용하여 중복 방지를 위해 하나의 파형 구간 외에는 제로 존(zero zone)으로 처리하고, 상기 추출된 파형 단위의 PPG 신호 및 GSR 신호의 라벨링은 상기 PPG 신호의 첨두치를 기준으로 진행하고, 추출 및 전처리된 데이터와 라벨링을 통해 복수의 세그먼트를 추출하여 라벨링 범위에 따라 복수의 클래스로 분류한다.

[0013] 상기 학습부는 상기 PPG 신호 및 GSR 신호 기반의 일차원 컨볼루션 오토인코더를 실시간 프로그램 및 하드웨어 성능이 제한된 환경에서 적용하기 위해 지식 증류 기법을 적용하여, 정확도를 유지하면서 네트워크 압축, 매개변수 감소, 연산 소요 시간을 감소시킨다.

[0014] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 방법은 전처리부를 통해 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출하는 단계, 라벨링부를 통해 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플 사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행하는 단계 및 학습부를 통해 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습하는 단계를 포함한다.

발명의 효과

[0015] 본 발명의 실시예들에 따르면 실시간 감정인식 및 분류를 위한 PPG 와 GSR 신호기반 일차원 컨볼루션 오토인코더 방법 및 시스템을 통해 PPG 신호와 GSR 신호의 생체신호를 사용하여 감정 분류의 정확성과 안정성, 사용자의 편의성을 개선할 수 있다. 또한, 일차원 컨볼루션 오토인코더에 대해 지식 증류 기법을 적용하여, 정확도를 유지하면서 네트워크 압축, 매개변수 감소, 연산 소요 시간을 감소시킴으로써 실시간 및 제한된 하드웨어 환경에서도 사용이 가능하다.

도면의 간단한 설명

[0016] 도 1은 종래기술에 따른 측정 위치 및 유형 기반 자율신경계의 생체신호 분류를 설명하기 위한 도면이다.

도 2는 종래기술에 따른 하드웨어 리소스가 제한된 임베디드 시스템에 배포 및 사용을 위한 딥러닝 모델의 매개변수를 줄이는 방법의 예시를 나타내는 도면이다.

도 3은 종래기술에 따른 2차원 구조에서의 육체적 정서적 기반 감정인식 및 분류를 설명하기 위한 도면이다.

도 4는 종래기술에 따른 DEAP 데이터셋의 라벨링 방법을 설명하기 위한 도면이다.

도 5는 종래기술에 따른 단기 생체신호 데이터를 위한 세분화된 라벨링 방법을 설명하기 위한 도면이다.

도 6은 본 발명의 일 실시예에 따른 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호기반 일차원 컨볼루션 오토인코더 시스템의 구성을 나타내는 도면이다.

도 7은 본 발명의 일 실시예에 따른 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호기반 일차원 컨볼루션 오토인코더 방법을 설명하기 위한 흐름도이다.

도 8은 본 발명의 일 실시예에 따른 PPG 신호 및 GSR 신호의 전처리 결과를 나타내는 도면이다.

도 9는 본 발명의 일 실시예에 따른 최종 윈도우 슬라이싱 결과를 나타내는 도면이다.

도 10은 본 발명의 일 실시예에 따른 멀티모달 일차원 컨볼루션 오토인코더의 구성을 나타내는 도면이다.

도 11은 본 발명의 일 실시예에 따른 학생모델(1D CNN) 구성을 나타내는 도면이다.

도 12는 본 발명의 일 실시예에 따른 PPG 신호 및 GSR 신호 기반의 감정분류 결과 그래프를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0017] 많은 양의 전극이 필요한 다른 신호와 다르게 PPG 신호와 GSR 신호는 사용자의 손가락 끝에서 측정과 취득이 가능하며, 편의성 및 추가적인 특징 추출 방법이 없어도 사용이 가능하다. 따라서, 본 발명에서는 두 개의 생체신호를 사용하여, 감정 분류의 정확성과 안정성, 사용자의 편의성 개선이 가능한 멀티모달 감정분류 방법을 제안한다. 또한 실시간 및 제한된 하드웨어 환경에서도 사용이 가능한 모델을 개발하였다. 이하, 본 발명의 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0018] 인간-로봇 상호 작용 분야에서 감정인식 및 분류 기술은 다른 연구 분야에 적용하기 위해 빠른 데이터 처리와 모델 경량화를 설정해야 한다. 본 발명은 아시안 멀티모달 데이터 MERTI-Apps의 관찰자 주석 라벨링을 이용한 PPG(photoplethysmogram), GSR(Galvanic Skin Response) 신호 라벨링 처리방법, 1차원 합성곱 신경망 모델과 오토인코더(autoencoder)를 이용한 실시간 감정 분류 방법을 제안한다. 또한, 모델의 성능을 확인하기 위해 공개 데이터 DEAP와 아시안 멀티모달 데이터 MERTI-Apps를 사용하여 결과를 비교한다. 본 발명에서의 주요 연구는 감정 분류를 위한 멀티모달 생체신호 기반의 파형 단위 데이터 전처리, 세부적인 학습데이터 설정, 감정 분류 모델 학습 및 경량화를 하는 것이다. 감정 분류를 위해 생체신호 데이터는 1-Pulse 단위로 윈도우 슬라이싱(window slicing) 되며, PPG와 GSR 신호의 특성을 유지하기 위한 라벨링을 재설정한다. 또한, 추가적인 특징추출의 핸드크래프트(handcraft) 작업 없이 간단한 데이터 전처리와 데이터의 파형 중복과 손실 방지를 진행한다. 실험결과 MERTI-Apps의 정확도는 3-class 기준 육체적(arousal) 79.18%, 정서적(valence) 74.84%를 확인하였으며, DEAP의 정확도는 2-class 기준 arousal 81.33%, valence 80.25%로 모든 데이터 셋에서 높은 성능을 확인하였다. 또한, 제안하는 방법은 MERTI-Apps를 사용한 종래기술보다 개선된 연산 시간 및 정확도를 확인하였으며, 제한된 하드웨어 환경 및 빠른 연산, 실시간 감정분류 시스템에 사용할 수 있도록 지식 증류를 적용하였다. 경량화를 진행한 모델의 정확도는 3-class 기준 arousal 77.87%, valence 73.49%이며, 경량 모델의 연산 소요 시

간은 일차원 컨볼루션 오토인코더 모델과 비교하면 80% 이상 단축되었고, 모델의 매개변수가 98% 감소했음을 확인하였다.

[0019] 고차원 영역의 감정을 인식 및 분류하기 위해 가장 많이 사용되는 데이터는 표정과 목소리로 진행되었다. 얼굴 표정을 기반한 감정 및 의도 인식은 높은 성능과 다양한 기능 나타내고 있다. 얼굴에는 눈빛, 표정, 피부색 변화, 미간 등 인간의 상태 정보를 포함하는 데이터가 있으며, 표정 인식 및 분류하기 위한 주요한 단계는 얼굴 이미지 전처리, 특징 추출, 특징 분류가 있다. 특징 추출의 방법은 크게 기하학적 특징 추출과 얼굴 모양 특징 추출하는 두 가지 유형이 있다. 기하학적 특징에는 입, 눈, 코, 턱 등의 얼굴 구성요소를 포함하며, 얼굴 모양 특징에는 주름과 같은 얼굴의 변화를 포함하는 특정 영역을 포함한다. Timothy et al은 외모 통계 모델을 이미지와 일치시키는 방법을 제안하였다. 이 방법은 모델 파라미터의 오류와 결과에서 오류 사이의 학습된 상관관계를 활용하여 효율적인 반복 매칭 알고리즘을 구성하며, 모델 파라미터는 트레인 셋(Train set)에서 학습한 형태와 그레이-레벨(Gray-level)의 변화를 제어한다. Phavish et al은 CNN 모델을 사용하여 딥러닝 접근 방식을 통해 인간의 얼굴 표정 분류를 탐구하였다. 엣지 감지 프레임 워크를 사용하여 얼굴 구성 정보의 모양과 위치를 추출하고, 얼굴 표정 분류는 확률적 분류기를 사용하였다. 사진, 비디오 기반의 얼굴표정 데이터를 수집하기 위해 비용이 많이 들고, 개인 정보 유출에 대한 문제점들을 포함하고 있다. 음성 신호는 연령별, 지역별, 언어에 따른 기준으로 다양한 형태로 표현되기 때문에, 취득된 데이터가 복잡하게 된다. 따라서 개인적인 특성을 표준화 및 전처리를 통한 특징 추출을 진행해야만 한다. Albornoz et al은 2단계의 계층적 분류기를 사용한 음성인식 및 분류를 진행하였으며, 음성 신호의 운율적 특성과 스펙트럼 특성, 음향적 특성 기반의 음성 감정분류를 연구하였다. Eyben et al은 음성 신호 기반의 실시간으로 감정을 인식과 별도의 윈도우 슬라이싱(window slicing)이 필요 없는 LSTM-RNN 모델을 사용하였다. 또한 특징 추출과 융합을 통한 음향적 특징과 언어적 특징을 결합하였다. 목소리 기반의 감정인식 및 분류는 이미지에 비해 개인정보 유출의 위험이 줄어든다. 하지만 인간의 행동이 제한되는 도서관, 공사장, 수중 등 환경적인 요소와 발성이 어려운 장애인에는 적용하지 못한다.

[0020] 자율 신경계에서 취득된 생체신호는 근육과 뼈를 통해 전달되는 전기적인 신호로 이루어져 있기 때문에, 외부 환경적인 영향이 적다. 따라서 표정, 목소리 신호가 가지는 문제점을 해결할 수 있으며, 추가적인 생체신호를 융합한 기술은 강력한 감정인식 및 분류 성능과 기능을 나타내고 있다. 사용자가 제어 가능한 체성신경계와 다르게, 자율 신경계는 제어 불가능한 신호로 정확한 내면 감정을 인식 및 분류가 가능하다. 하지만 내면의 감정 인식을 위해 단일 생체신호만 사용한다면, 정확성과 학습데이터 한계를 가지게 된다. 향상된 성능을 위해서 다중 생체신호 기반의 멀티모달 감정인식 및 분류 방법이 제안되었다. Hui et al은 차원 감정인식의 연속성을 결합하기 위해 결정 인식 퓨전(decision recognition fusion) 모델을 제안하였다. 비침습형으로 취득된 EEG, ECG 신호를 전처리 및 특징 추출, PNN(Probabilistic Neural Network) 기반으로 감정인식을 수행하였고, 실험결과에서 다중 신호의 결과가 단일 신호의 결과보다 높은 성능을 나타내었다. Yang et al은 EEG, PPG 신호를 11개의 통계적 시간 영역, 5가지의 주파수 영역으로 추출하였고, 1D CNN을 사용하여 감정 분류를 진행하였다. 또한 ECG와 PPG 신호 사이의 PTT(Pulse Transit Time)를 계산하여 최종 분류 정확도를 개선하였다. 하지만 앞의 연구에서는 특징 추출 및 감정 인식을 위한 최소 10초에서 최대 60초의 신호 대기 시간이 필요로 한다. 10초 이상의 신호 대기 시간은 실시간 감정인식 시스템에서 적합하지 않으며, 다른 신호(영상, 이미지, 음성 등)와 융합하는 연구에서 사용이 제한된다. Ayata et al은 PPG, GSR 신호의 감정을 위해 분류, 회귀 분석에서 사용되는 앙상블 학습 방법인 랜덤 포레스트와 KNN, SVM을 사용하였다. 3초 길이의 GSR 신호와 8초 길이의 PPG 신호에서 특징을 추출하여, arousal 기준 72.06%와 valence 기준 71.05%의 정확도를 도출하였다. 또한 PPG 신호와 GSR 신호의 감정에 대한 arousal, valence 영역에서의 상관관계를 확인하였다. Lee et al은 PPG 신호에서 심장박동변이에 대한 NN(normal to normal) 시간 영역의 특징과 정규화를 통한 주파수 영역의 특징을 이용하였다. 10초 길이의 PPG 신호에서 추출된 특징 기반 arousal 기준 82.1%와 valence 기준 80.9%의 정확도를 도출하였다. 감정 인식을 위해 필요한 최소한의 시간이 기존 연구에 비해 단축되었지만, 기존 연구와 동일하게 데이터의 처리 과정에서 추가적인 특징 추출을 필요로 한다. 앞의 연구에서는 다양한 특징 추출 방법으로 생체신호 기반 감정인식 및 분류 성능이 향상되었지만, 시간-주파수 영역 특징 추출과 같은 Handcraft 작업은 추가적인 연산 시간과 낮은 활용도 문제가 여전히 존재하고 있다. 기존 문제점을 개선하고, 실시간 감정 인식을 위해서는 빠른 데이터 처리 방법과 신호 특성을 유지하는 방법이 필요하다.

[0022] 도 2는 종래기술에 따른 하드웨어 리소스가 제한된 임베디드 시스템에 배포 및 사용을 위한 딥러닝 모델의 매개변수를 줄이는 방법의 예시를 나타내는 도면이다.

[0023] 인공지능 연구의 발전은 하드웨어 및 소프트웨어 발전에 따라 크게 영향을 많이 받았다. 과거 1980년대에 개발

된 멀티레이어 퍼셉트론(Multilayer perceptron)은 기존 퍼셉트론(perceptron) 구조에 히든 레이어(hidden layer)를 추가하여, 기존 선형 분류기의 한계를 극복하였다. 하지만 다양한 산업 및 제품에는 당시 하드웨어의 성능, 매개변수 증가에 따른 학습 시간 및 요구 성능의 한계로 접근하기 어려웠다. 최근 병렬 연산과 여러 개의 코어 사용이 가능한 고성능의 GPU가 개발됨에 따라 빠른 학습 및 알고리즘을 통한 인공지능 모델 연구가 진행되고 있다. 하지만 개발된 방대한 양의 데이터 연산과정과 모델은 인간-로봇 상호작용을 위한 다양한 임베디드 장치의 자원에 적용하기 부적절하다. 하드웨어 리소스가 제한된 임베디드 시스템에 배포 및 사용을 위한 딥러닝 모델 압축, 가지치기(도 2(c)), 지식 증류(도 2(b)) 등 모델의 매개변수를 줄이는 방법이 도 2와 같이 있으며, 다양한 연구가 진행되었다. Han et al은 특정 임계점을 기준으로 낮은 값을 가지는 레이어 단위로 가지치기를 진행 후 반복적인 재학습을 진행하였다. 전체적인 파라미터는 61M 개에서 약 10% 줄어든 6.7M로 감소시켰으며, 기존 모델과 정확도 오차는 0.1%로 크게 차이는 나지 않았다. 지식 증류 기술은 앙상블 기법을 통해 학습된 큰 모델인 교사(Teacher)로부터 지식을 전달하여, 동일한 성능을 나타내는 작은 모델인 학생(Student)을 구축할 수 있다. Ho et al은 복잡한 교사 모델의 지식을 가벼운 학생 모델로 추출하거나 이러한 학생 모델을 자체 훈련하여 다중 레이블 패 질환 분류를 진행하였다. 앞서 언급된 모델 경량화 기술은 생체신호를 사용하는 제한된 하드웨어 환경, 임베디드 보드, 생체신호 취득용 모듈에서 효과적으로 사용 가능하다. 또한 기존 모델 대비 정확도 손실을 최소화, 요구되는 메모리, 연산량 등을 줄일 수 있기 때문에 영상, 음성신호를 같이 사용하는 병렬적 시스템에도 적용이 가능하다.

[0025] 도 3은 종래기술에 따른 2차원 구조에서의 육체적 정서적 기반 감정인식 및 분류를 설명하기 위한 도면이다.

[0026] 지금까지 심리학자들은 일반적인 감정과 복잡한 감정의 두 가지 관점에서 감정 연구를 하였다. 일반적인 감정은 행복, 분노, 슬픔, 두려움과 같은 범주로 나뉘어져 있고, 복잡한 감정은 다차원적 구조를 가지고 있으며 지속적으로 변한다. 생체신호 기반의 감정인식 및 분류 연구에서 많이 사용되는 구조는 도 3과 같이 Russell의 2차원 구조의 육체적(arousal), 정서적(valence) 기반으로 이루어져 있다. 감정을 2차원 구조로 표현함으로써, 복잡한 감정과 일반적인 감정을 arousal, valence의 척도로 세분화하여 나타낼 수 있다. 감정분석을 위한 대표적인 생체신호 공개데이터는 아래의 표 1과 같이 DEAP dataset, MAHNOB-HCI database로 동일한 2차원 영역 기반의 감정을 라벨링 하였다.

[0027] <표 1>

DEAP DATA SET	
Participants	32 (16 male, 16 female)
Video	40 videos (60 second)
Bio-signals	EEG, EOG, EMG, GSR, PPG
Labeling	Self-Assessment Labeling (Arousal, Valence)
MAHNOB-HCI DATA BASE	
Participants	27 (11 male, 16 female)
Video	20 videos (34 to 114 second)
Bio-signals	EEG, EOG, GSR
Labeling	Annotation Labeling (Valence)

[0028]

[0029] 실험자의 감정에 대한 정확한 훈련 데이터를 만들기 위해서는 인간의 감정 라벨이 중요하게 된다. 또한 2가지의 공개데이터는 서양인 기준의 생체신호 기반 감정인식 및 분류에 주로 사용되었다.

[0031] 도 4는 종래기술에 따른 DEAP 데이터셋의 라벨링 방법을 설명하기 위한 도면이다.

[0032] DEAP 데이터셋은 총 32명의 피험자(남성 16명, 여성 16명)에 대한 다중 생체신호 데이터 세트이다. 실험자는 감정유발을 위한 40개의 뮤직 비디오 (1분 간격)를 시청하여, 실험자의 EEG, EOG, EMG, GSR, PPG, 체온, 호흡을 취득하였다. DEAP 데이터셋이 사용하는 라벨링 방법은 다차원 감정을 정량적으로 설명하기 위해 도 4와 같이 SAM(self-assessment manikins)을 사용하였고, 피험자는 Arousa(410), Valence(420), Liking(430), 및 Dominance(440)을 자체 측정하였다. 자기평가 라벨링은 Arousal의 척도는 지루함에서 흥분까지, Valence의 척도는 슬픔에서 기쁨까지 1에서 9까지의 범위로 표현되었다. 피험자는 60초 길이의 감정유발 영상을 시청하고, 자기평가 라벨링을 표기하였다. 따라서 시청 중일 때는 감정의 변화를 표현하기 어려움이 있었으며, 주어진 영상

전체의 감정 기준으로 세분화하기에는 한계점이 존재한다.

[0034] 도 5는 종래기술에 따른 단기 생체신호 데이터를 위한 세분화된 라벨링 방법을 설명하기 위한 도면이다.

[0035] MAHNOB-HCI 데이터셋은 DEAP과 동일한 공개데이터로써, 총 27명의 피험자(남성 13명, 여성 16명)가 20개의 감정 유발 영상에 대한 감정 데이터를 수집하였다. 실험자의 EEG, EOG, GSR, 체온, 호흡을 취득하였다. MAHNOB-HCI 데이터셋에서 사용한 라벨링은 사전 교육을 받은 5명의 관찰자가 피험자의

[0036] 표정에 대한 주석을 1에서 9까지의 범위로 표현하였다. 관찰자 주석 라벨링은 교육된 관찰자가 피험자의 얼굴 정면 영상을 보며 실시간으로 감정 변화를 측정하였기 때문에, 도 5와 같이 단기 생체신호 데이터에는 세분화된 라벨링을 사용할 수 있는 것을 확인할 수 있다. 도 5(a)는 DEAP 데이터셋의 Self-Assessment, 도 5(b)는 MAHNOB-HCI 데이터셋의 Annotationdmf 나타내는 그래프이다. 결과적으로 관찰자 주석 라벨링은 실험자에게 하나의 감정이 도출될 때까지의 감정 변화를 관찰할 수 있다. 따라서 학습데이터의 라벨에 감정의 변화가 잘 드러나는 관찰자 주석 라벨링이 실시간 감정인식 및 세분화된 감정 분류에 적절하다고 볼 수 있다.

[0037] 본 발명에서는 공개데이터 중에서 DEAP 데이터셋과 표 2의 아시안 멀티모달 데이터베이스 MERTI-Apps 실험1을 사용하여, PPG와 GSR 신호 기반의 멀티모달 감정분류를 진행하였다.

[0038] <표 2>

Participants	62 (28 male, 34 female)
Video	40 videos (60 second)
Experiment 1	
Bio-signals	EOG, EMG, GSR, PPG
Emotion response to videos	5 videos (Sad 1, Happy1, Angry 2, Scared 1)
Labeling	Annotation Labeling (Arousal, Valence) Self-assessment Labeling (Arousal, Valence)
Experiment 2	
Bio-signals	EEG, EMG, EOG
Emotion response to videos	5 videos (Sad 1, Happy1, Angry 1, Scared 1, Neutral 1)
Labeling	Annotation Labeling (Arousal, Valence) Self-assessment Labeling (Arousal, Valence)
Experiment 3	
Bio-signals	EEG, EOG, PPG, GSR
Emotion response to videos	5 videos (Sad 1, Happy2, Angry 1, Scared 1)
Labeling	Self-assessment Labeling (Arousal, Valence)

[0039]

[0040] 아시안 멀티모달 데이터 MERTI-Apps는 총 62명(남성 28, 여성 34)가 15개의 감정을 유발 영상에 대한 데이터를 추출하였다. 취득된 데이터는 실험1(EMG, EOG, PPG, GSR), 실험2(EEG, EMG, EOG), 실험3(EEG, EOG, PPG, GSR)을 측정되었으며, 사전 교육을 받은 5명의 관찰자에 의한 라벨링 및 자기 평가 라벨링을 모두 진행하였다. 취득된 감정유발 영상 기반 라벨링의 척도는 -100~100까지 표현되었다.

[0042] 도 6은 본 발명의 일 실시예에 따른 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호기반 일차원 컨볼루션 서널 오토인코더 시스템의 구성을 나타내는 도면이다.

[0043] 제안하는 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호기반 일차원 컨볼루션 서널 오토인코더 시스템은 전처리부(610), 라벨링부(620) 및 학습부(630)를 포함한다.

[0044] 본 발명의 실시예에 따른 전처리부(610)는 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 과형 단위로 추출한다.

[0045] 본 발명의 실시예에 따른 전처리부(610)는 PPG 신호에 대하여 선형 주파수 영역의 필터를 적용하고, 고차다항식 적합(High order Polynomial)으로 데이터의 기저선 변동 잡음과 이동평균 필터(Moving Average Filter)을 통해

동적 잡음을 감소시킨다.

- [0046] 본 발명의 실시예에 따른 전처리부(610)는 GSR 신호에 대하여 PPG 신호 기준으로 분할한 후 동일한 전처리를 진행한다.
- [0047] 본 발명의 실시예에 따른 라벨링부(620)는 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플 사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행한다.
- [0048] 본 발명의 실시예에 따른 라벨링부(620)는 심장박동 수를 이용하여 중복 방지를 위해 하나의 파형 구간 외에는 제로 존(zero zone)으로 처리한다.
- [0049] 상기 추출된 파형 단위의 PPG 신호 및 GSR 신호의 라벨링은 상기 PPG 신호의 첨두치를 기준으로 진행하고, 추출 및 전처리된 데이터와 라벨링을 통해 복수의 세그먼트를 추출하여 라벨링 범위에 따라 복수의 클래스로 분류한다.
- [0050] 본 발명의 실시예에 따른 학습부(630)는 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습한다.
- [0051] 본 발명의 실시예에 따른 학습부(630)는 상기 PPG 신호 및 GSR 신호 기반의 일차원 컨볼루션 오토인코더를 실시간 프로그램 및 하드웨어 성능이 제한된 환경에서 적용하기 위해 지식 증류 기법을 적용하여, 정확도를 유지하면서 네트워크 압축, 매개변수 감소, 연산 소요 시간을 감소시킨다.
- [0053] 도 7은 본 발명의 일 실시예에 따른 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호기반 일차원 컨볼루션 오토인코더 방법을 설명하기 위한 흐름도이다.
- [0054] 제안하는 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호기반 일차원 컨볼루션 오토인코더 방법은 전처리부를 통해 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출하는 단계(710), 라벨링부를 통해 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플 사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행하는 단계(720) 및 학습부를 통해 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습하는 단계(730)를 포함한다.
- [0055] 단계(710)에서, 전처리부를 통해 실시간 감정인식 및 분류를 위한 PPG 신호 및 GSR 신호의 데이터 전처리를 수행하여 파형 단위로 추출한다.
- [0056] 본 발명의 실시예에 따른 전처리부는 PPG 신호에 대하여 선형 주파수 영역의 필터를 적용하고, 고차다항식 적합(High order Polynomial)으로 데이터의 기저선 변동 잡음과 이동평균 필터(Moving Average Filter)를 통해 동적 잡음을 감소시킨다.
- [0057] 본 발명의 실시예에 따른 전처리부는 GSR 신호에 대하여 PPG 신호 기준으로 분할한 후 동일한 전처리를 진행한다.
- [0058] 단계(720)에서, 라벨링부를 통해 상기 PPG 신호 및 GSR 신호를 분할하기 위해 미리 저장된 PPG 신호의 첨두치를 기준으로 샘플 사이즈(sample size)가 동일하게 분할하고, 상기 PPG 신호 및 GSR 신호의 라벨링을 수행한다.
- [0059] 본 발명의 실시예에 따른 라벨링부는 심장박동 수를 이용하여 중복 방지를 위해 하나의 파형 구간 외에는 제로 존(zero zone)으로 처리한다.
- [0060] 상기 추출된 파형 단위의 PPG 신호 및 GSR 신호의 라벨링은 상기 PPG 신호의 첨두치를 기준으로 진행하고, 추출 및 전처리된 데이터와 라벨링을 통해 복수의 세그먼트를 추출하여 라벨링 범위에 따라 복수의 클래스로 분류한다.
- [0061] 단계(730)에서, 학습부를 통해 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습한다.

[0062] 본 발명의 실시예에 따른 학습부는 상기 PPG 신호 및 GSR 신호 기반의 특징 추출, 감정 분류를 진행하기 위해 일차원 컨볼루션 오토인코더(1D convolutional autoencoder)를 사용하여 감정분류를 위한 모델을 설정하고 학습한다.

[0063] 본 발명의 실시예에 따른 학습부는 상기 PPG 신호 및 GSR 신호 기반의 일차원 컨볼루션 오토인코더를 실시간 프로그램 및 하드웨어 성능이 제한된 환경에서 적용하기 위해 지식 증류 기법을 적용하여, 정확도를 유지하면서 네트워크 압축, 매개변수 감소, 연산 소요 시간을 감소시킨다.

[0065] 앞서 설명된 바와 같이 기존 생체신호 기반의 감정인식 및 분류 방법은 일반적으로 데이터 전처리와 신호 특징을 시간-주파수 영역에서 추출하는 과정을 가진다. 이러한 추가적인 Handcraft 과정은 많은 시간과 실시간 시스템에 적용하기 어렵다. 본 발명에서는 사용자의 편의성과 실시간 감정 분류에 용이한 두 가지 PPG, GSR 신호를 사용하였으며, 생체신호의 특성을 유지하기 위한 라벨링 재설정하였다. 제안하는 방법은 생체신호 기반의 데이터 전처리, 학습데이터 라벨 설정, 감정분류 모델 학습 및 경량화로 크게 3가지로 나뉘어진다. 추가적으로 전처리 및 윈도우 슬라이싱이 완료된 PPG, GSR 신호를 추가적인 Handcraft 과정이 없는 멀티모달 감정분류 모델 개발, 제한된 환경에서 사용을 위한 실험을 수행한다. 감정분류에서의 기본적인 모델은 동일한 구조의 입력(encoder)과 출력(decoder)으로 구성된 1차원의 컨볼루션 오토인코더 모델을 사용하였다.

[0066] 원시 PPG 신호 데이터는 ECG 신호보다 편의성과 사용성이 높지만, 전원 노이즈와 같은 고주파 성분 및 모세혈관에 의한 저주파 성분의 노이즈를 제거해야만 한다. 본 발명에서는 단순한 선형 주파수 영역의 필터인 버터워스 필터(Butterworth Filter)를 적용하고, 고차다항식 적합(High order Polynomial)으로 데이터의 기저선 변동 잡음과 이동평균 필터(Moving Average Filter)를 통해 동적잡음을 줄였다. GSR 신호는 PPG 신호 기준으로 분할한 후에 동일한 전처리를 진행하였다.

[0067] 버터워스 필터는 통과대역의 진폭 스펙트럼이 평평하며, 천이대역(Transition Band)이 완만한 주파수 필터이다. 차수 N 이 높을수록 천이대역이 작아지고, 천이가 급격해진다. 간단한 선형 주파수 필터로써, 아래와 같이 표현된다.

[0068] 필터 근사를 위한 기초 공식은 다음과 같다:

[0069]
$$|H(j\omega)| = \frac{K}{\sqrt{1+\varepsilon^2 f(\omega^2)}} \quad (1)$$

[0070] 버터워스 필터 근사 함수 ($f(\omega^2) = \omega^{2n}$):

[0071]
$$|H(j\omega)| = \frac{K}{\sqrt{1+\varepsilon^2 \omega^{2n}}} \quad (2)$$

[0072] 정규화된 버터워스 필터 근사 함수 ($K = 1, \varepsilon = 1, \omega_c = 1$):

[0073]
$$|H(j\omega)| = \frac{K}{\sqrt{1+\omega^{2n}}} \quad (3)$$

[0074] 생체신호 취득 후 데이터가 가지는 이상치, 오차를 해결하기 위해 일반적으로 사용되는 방법으로 고차다항식 적합을 사용한다. 입력변수와 출력변수의 상관관계를 분석 및 예측을 통해 오차율을 줄여, 가장 적절한 데이터를 가지게 된다. 고차다항식 적합은 아래와 같이 표현된다.

[0075] N 차 다항식 방정식:

[0076]
$$y(x, \omega) = \sum_{j=0}^N \omega_j x^j \quad (4)$$

[0077] 일반적인 이동 필터는 현재 시점까지 측정된 데이터의 총합의 평균이다. 측정되는 값이 시간에 따라 지속적으로 변하는 값이라면, 이동 필터는 적절하지 않다. 따라서 데이터의 잡음 제거와 동적 데이터의 평균을 구하기 위해서 사용되는 방법으로 이동평균 필터를 사용한다. 즉, 모든 데이터의 평균이 아닌 N 개의 최신 값 평균을 사용한다. 이동평균 필터는 아래와 같이 표현된다.

[0078] N 개의 데이터에 대한 이동 필터:

$$\bar{x}_k = \frac{x_{k-n+1} + x_{k-n+2} + \dots + x_k}{n} = \bar{x}_{k-1} + \frac{x_k - x_{k-n}}{n} \quad (5)$$

도 8은 본 발명의 일 실시예에 따른 PPG 신호 및 GSR 신호의 전처리 결과를 나타내는 도면이다.

데이터 전처리 후 정확한 윈도우 슬라이싱을 위해 10초 길이의 PPG 신호 기준 첨두치 및 평균 심장박동 수 정보를 수집 및 저장한다. PPG, GSR 신호의 전처리 결과는 도 8과 같다. 10초 길이의 PPG, GSR 신호 전처리 결과를 나타내고, 도 8(a)는 전처리된 PPG 신호; 도 8(b)는 PPG 신호의 첨두치를 나타내고, 도 8(c)는 PPG 신호의 첨두치와 동일한 구간의 GSR 신호를 나타낸다.

PPG, GSR 신호를 정확하게 분할하기 위해 사전에 저장된 PPG 신호의 첨두치 기준으로 샘플 사이즈(sample size)가 동일하게 분할하였다. 결과적으로 1kHz로 측정된 PPG, GSR 신호의 샘플 사이즈는 1100개로 유지하였다. 또한 심장박동 수를 이용하여, 중복 방지를 위해 하나의 파형 구간 외에는 제로 존(zero zone)으로 처리하였다.

중복 방지를 위한 제로 존 설정:

$$zero\ zone = \frac{\left(\frac{total\ frequency\ sample\ size}{heart\ rate} - sample\ size \right)}{2} \quad (6)$$

도 9는 본 발명의 일 실시예에 따른 최종 윈도우 슬라이싱 결과를 나타내는 도면이다.

추출된 파형 단위 PPG, GSR 신호의 라벨링은 MERTI-Apps의 관찰자 주석 라벨을 사용하여 PPG 신호의 첨두치 기준으로 진행하였다. 5명의 해설자가 실험자의 얼굴 표정을 0.25초 간격으로 Arousal, Valence를 표현하였으며, 실험자가 자체 감정 평가 결과와 일치하지 않는 데이터는 사용하지 않았다. 62명의 실험자에게 5개의 1~5분 길이의 감정 유발영상을 통해 추출 및 전처리된 데이터와 라벨링을 통해 총 32,000개의 세그먼트를 추출하였다. 라벨의 범위는 -100~100에서 긍정적(20~100), 중립적(20~-20), 부정적(-20~-100)으로 세 가지 클래스로 나누었으며, 파형 구간 내부의 라벨의 평균을 사용하였다. 최종 윈도우 슬라이싱 결과와 분포도의 다양성은 표3, 도 9와 같이 확인되었다. 최종 PPG, GSR 신호의 결과물(sample size 1100)에 대하여 도 9(a)는 PPG 신호의 결과를 나타내고, 도 9(b)는 GSR 신호의 결과를 나타낸다.

<표 3>

		Annotation labeling	Self-assessment labeling
Number of data extracted		32000 segments	
Label (arousal)	high	14,489	2
	neutral	11,190	2
	low	6,321	1
Label (valence)	positive	6,716	1
	neutral	5,289	1
	negative	14,227	3

도 10은 본 발명의 일 실시예에 따른 멀티모달 일차원 컨볼루션 오토인코더의 구성을 나타내는 도면이다.

본 발명의 실시예에 따른 멀티모달 일차원 컨볼루션 오토인코더를 통해 파형 단위 PPG와 GSR 신호 기반의 특징 추출, 감정 분류를 진행한다. 모델은 2개의 컨볼루션 레이어(convolution layer)와 2개의 풀링 레이어(pooling layer)로 동일한 구조의 인코더(encoder), 디코더(decoder)로 구성되어 있다. 추출된 데이터는 컨볼루션 레이어(concatenation layer)를 통해 완전 연결 레이어(fully connected layer)로 전달된다. 각 레이어에 대한 정보는 도 10과 같으며, 최적화기(optimizer)는 Adam(learning rate:0.0008)을 사용하였다. 배치 사이즈(Batch size)를 16으로 설정하여, 반복 학습 횟수를 100번 실행하였다. 전체 데이터의 64%를 학습데이터, 16%를 검증데이터, 20%를 성능 검증데이터로 학습하였다.

[0095] 도 11은 본 발명의 일 실시예에 따른 학생모델(1D CNN) 구성을 나타내는 도면이다.

[0096] 본 발명의 실시예에 따른 2가지 생체신호 기반의 일차원 컨볼루션 신경 네트워크 모델(교사)을 실시간 프로그램 및 하드웨어 성능이 제한된 환경에서 적용하기 위해, 지식 증류 기법을 적용하였다. 앞서 설명한 내용과 같이, 기존 교사모델의 지식을 간단한 구조의 1D CNN 모델(학생)에 전달하였다. 사용된 학생 모델은 도 11과 같이 단순한 구조이다. 단순히 컨볼루션 레이어와 풀링 레이어의 구조로 이루어져 있으며, 최적화기는 Adam(learning rate:0.0005)을 사용하였다. 배치 사이즈를 16으로 설정하여, 반복 학습 횟수를 200번 실행하였다. 최종 학생 모델은 정확도를 유지하면서, 네트워크를 압축 및 매개변수를 줄여, 연산 소요 시간을 감소시켰다.

[0097] 본 발명의 실시예에 따른 2가지 PPG, GSR 신호 기반의 멀티모달 일차원 컨볼루션 신경 네트워크 모델을 통해 감정 분류 방법을 제안하였다. 표 4와 같이 다양한 생체신호 기반의 감정인식 및 분류에 대한 연구가 진행되었다.

[0098] <표 4>

	Bio-signals	window size	Method	class	Accuracy	
					Arousal	Valence
Mert, 2018	EEG	60s	Multivariate extension; ANN; DEAP database	2	75.00%	72.87%
Lee, 2019	PPG	1.1s	1D CNN; DEAP database	2	76.20%	75.30%
Chao, 2019	EEG	3s	Multiband feature matrix; CapsNet; DEAP database	2	66.73%	68.28%
Xing, 2019	EEG	60s	SAE+LSTM; DEAP database	2	81.10%	74.38%
Lee, 2020	PPG	10s	1D CNN; Selected Statistical Feature; DEAP database	2	80.90%	82.10%
Raheel, 2021	EEG, GSR, PPG	60s	Time, Frequency domain; DEAR-MULSEMEDIA database	2	85.18%	76.54%
Hwang, 2022	EEG, EOG, GSR, PPG	60s	ACO-bidirectional LSTM; MERTI-Apps database	-	78.00%	81.00%
			1D Convolutional Autoencoder; DEAP database	2	81.33%	80.25%
			1D Convolutional Autoencoder (Teacher model); Annotation Labeling; MERTI-Apps database	2	80.89%	81.25%
			Annotation Labeling; MERTI-Apps database	3	79.18%	74.84%
			1D CNN (student model); Annotation Labeling; Knowledge distillation; MERTI-Apps database	3	77.87%	73.49%
Our	PPG, GSR	1.1s	LSTM; Annotation Labeling; MERTI-Apps database	3	70.40%	63.21%
			1D CNN; Annotation Labeling; MERTI-Apps database	3	71.18%	67.15%

[0099]

[0100] 생체신호 기반의 감정 분류 연구에서 적절한 입력 데이터의 윈도우 사이즈가 존재하며, 취득된 신호의 특징 추출 및 규칙성을 파악하기 위한 최소한 대기시간이 존재한다. 따라서 인식을 위한 필수적인 대기 시간으로 인한 실시간 특성 감소 및 단일 생체신호 사용으로 인한 정확도가 저하되었다. Xing et al은 많은 채널을 통해 취득되는 EEG 신호에서 상관관계와 특징을 추출하고, SAE 기반 선형 EEG 혼합 모델을 통해 연산시간을 줄였다. 하지만 감정인식을 위한 윈도우 사이즈가 60초로 실시간 시스템에 적용하기 어려운 점이 있다. Lee et al은 단일 PPG 신호를 1.1초의 짧은 과행단위의 데이터를 1D CNN을 이용하여 신호 특성을 자동 추출하기 때문에 추가적인 시간 손실을 방지하였고, 2개의 분류 결과에서 단기 감정분류의 가능성을 보였다. 하지만 단일 신호와 일체화된 라벨링은 세부적인 감정표현으로 감정인식 및 분류의 클래스 증가에 따른 정확도, 안정성 한계가 존재하게 된다. 본 발명에서는 신호의 특징을 추출하기 위한 추가적인 handcraft 과정 없이 높은 수준의 감정분류 가능성을 확인하였으며, 제안하는 방법은 사용자의 심장변화에 대한 정보를 포함하는 PPG 신호, 육체적인 변화에 민감한 GSR 신호를 모두 사용함으로써 개선된 감정분류 성능을 확인하였다. 실험 결과, DEAP에서는 arousal/valence 기준 81.33%/80.25%로 제안하는 방법이 빠르고 높은 성능을 나타낸다. 또한 세분화된 라벨링으로 MERTI-Apps에서는 기존 선례 연구보다 2가지의 클래스에서는 arousal/valence 기준 80.89%/81.25%와 3가지의 클래스에서는 79.18%/74.84%로 추가적인 클래스 증가에도 높은 성능지표를 확인하였다. 실시간 시스템 및 추가적인 신호(이미지, 영상 등)와 융합하기 위해 모델 경량화 작업을 진행하였다. 실제 차량 및 인간 상호 작용을 위한 로봇에서 많이 사용하는 임베디드 보드에서 모델의 연산 소요시간을 확인하였으며, 모델의 매개변수와 크기를 줄임으로써 모든 환경에서 기존 모델보다 절반 이하의 연산 소요시간이 확인되었다. 마지막으로 본 발명은 정확한 인간의 내면감정 인식을 위한 자율 신경계 신호 기반의 연구와 제한된 하드웨어 자원에 대응하기 위한 연구를 같이 진행하였다. 기존 연구들이 2가지의 클래스 기준으로 감정인식 및 분류를 진행하였지만, arousal과 valence 영역에서의 추가적인 클래스 기준으로 감정분류에 대한 결과를 도출하였다. 본 발명의 결과를 통해 10초~60초 사이의 감정인식 및 분류는 자기 평가 라벨링이 적절하지만, 1~5초 사이의 짧은 단위 생체신호 감정인식 및 분류는 관찰자 주석 라벨링 설정이 다양한 학습데이터를 제공할 수 있음을 확인하였다.

[0102] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴

퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0103] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

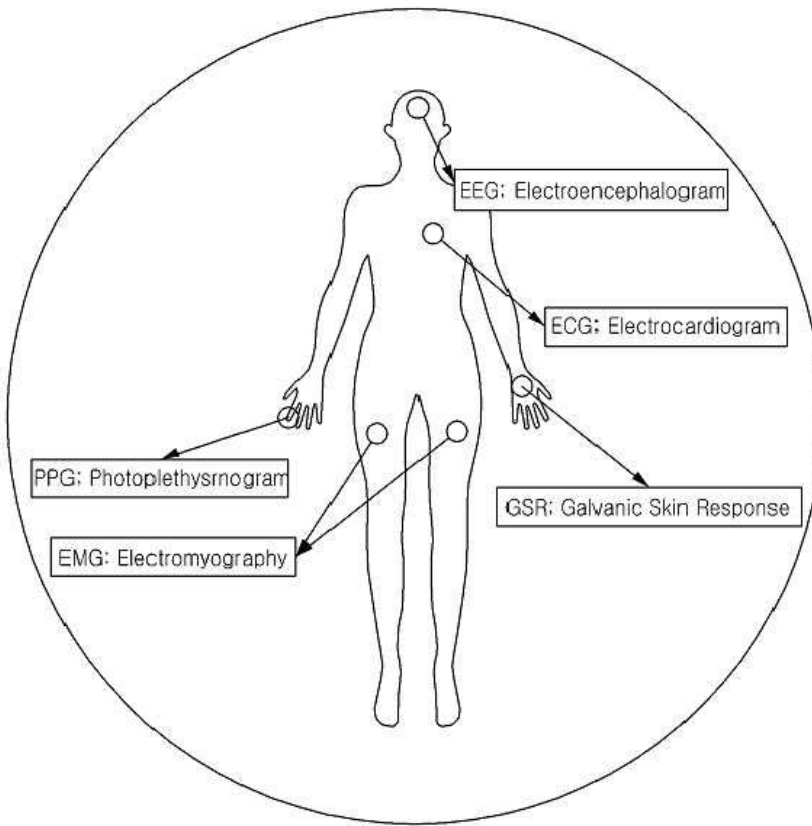
[0104] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

[0105] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

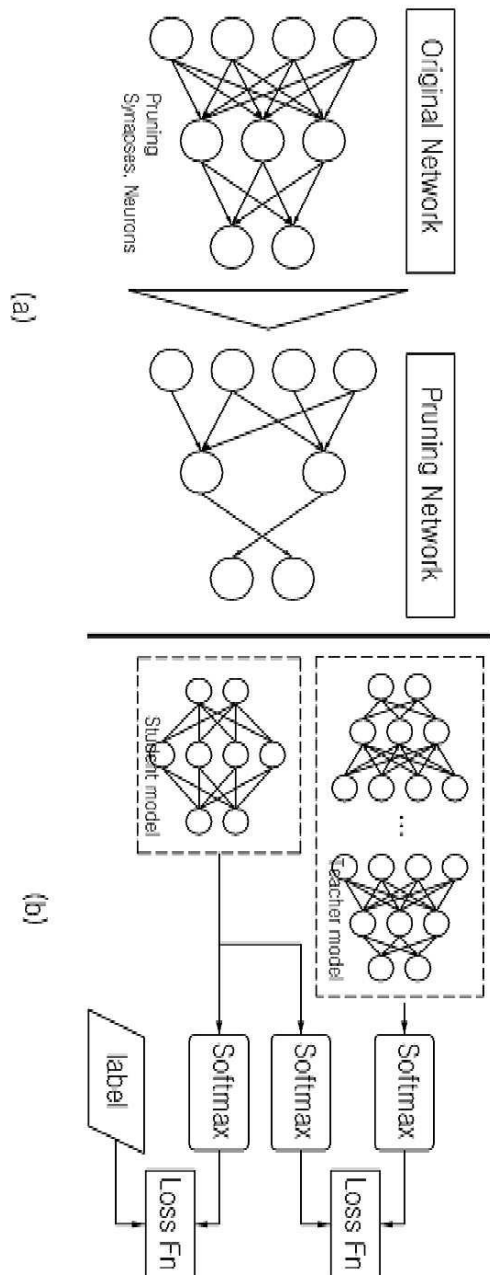
[0106] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

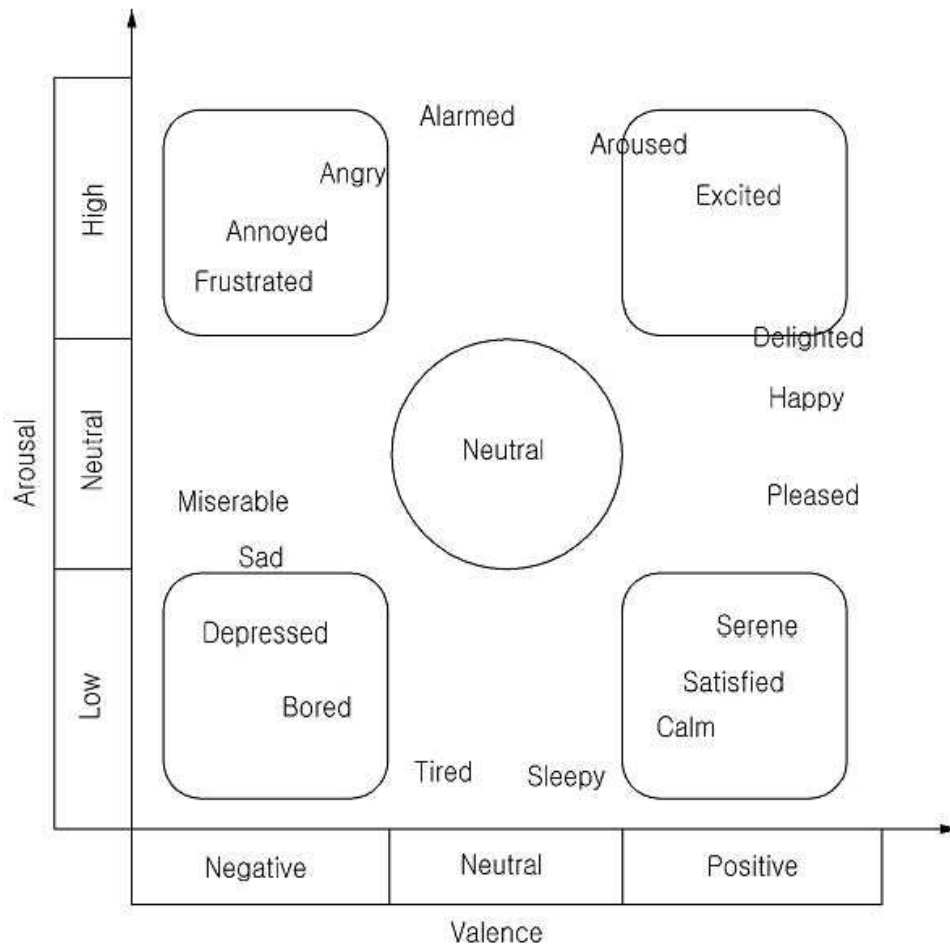
도면1



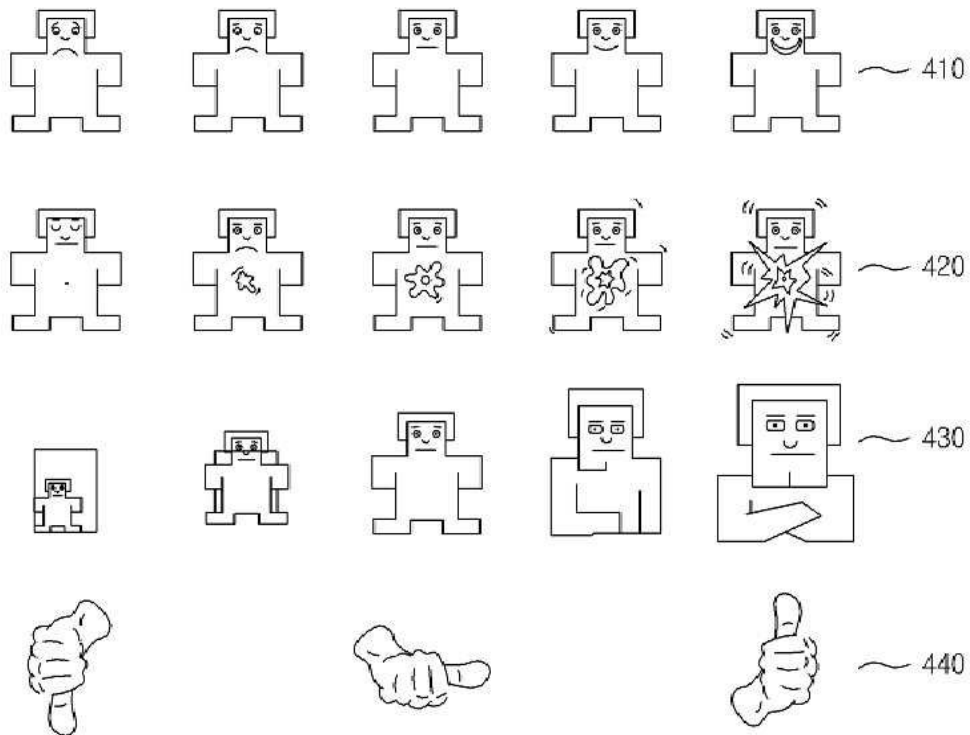
도면2



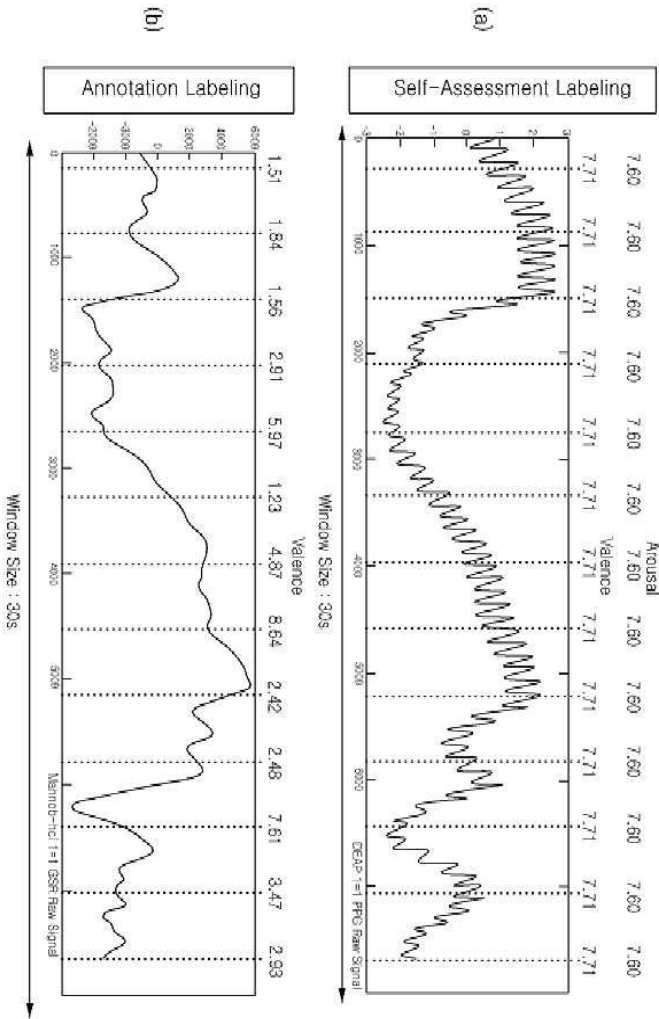
도면3



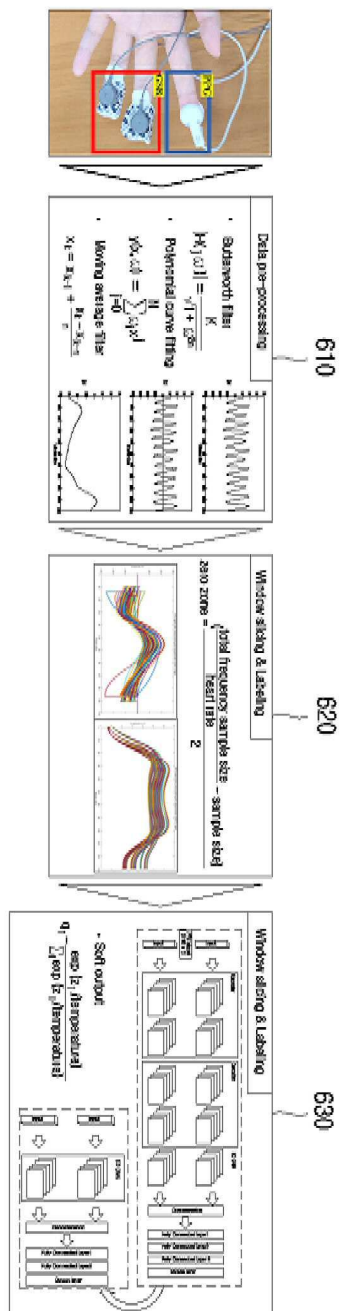
도면4



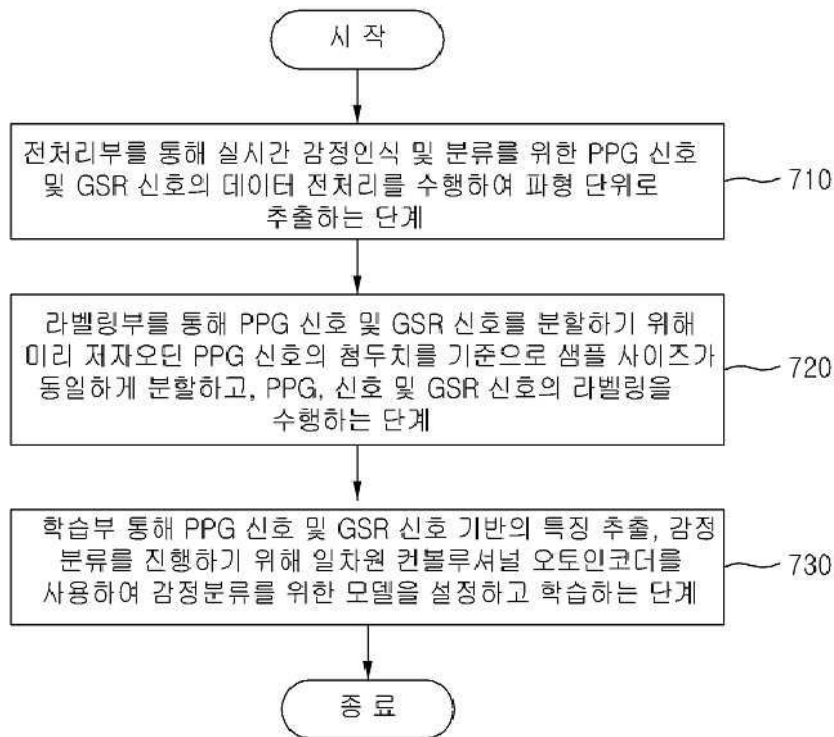
도면5



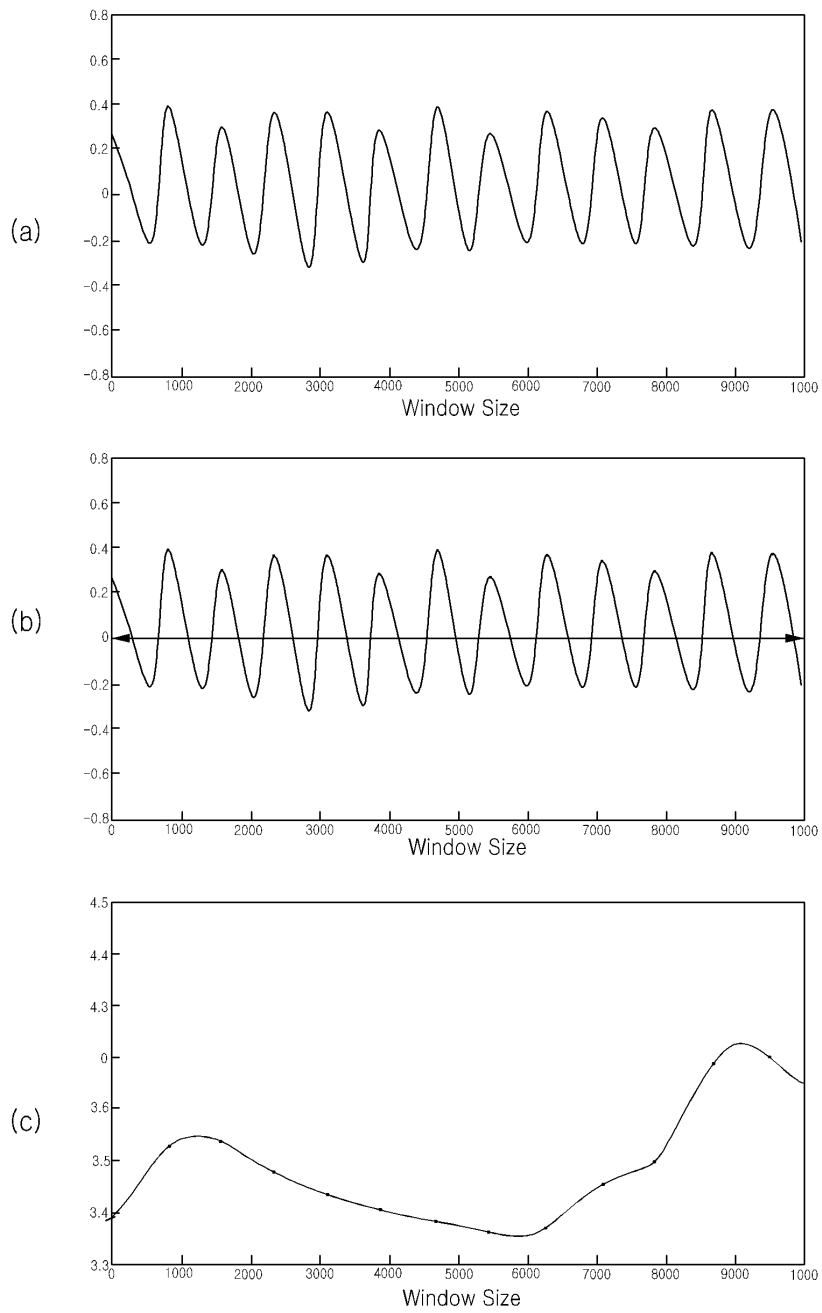
도면6



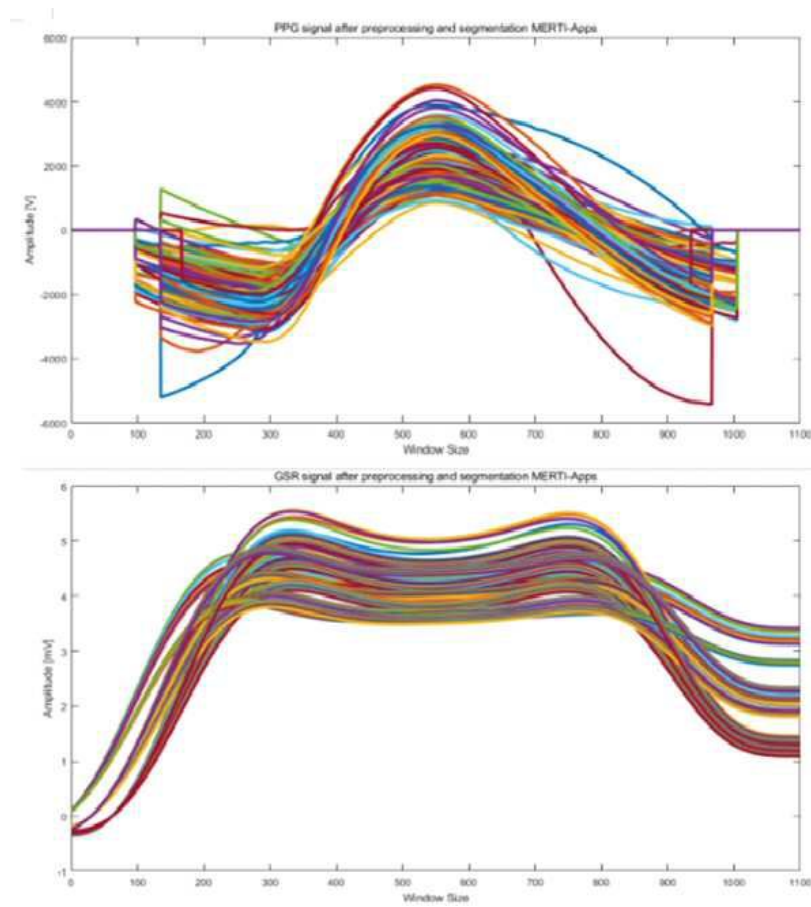
도면7



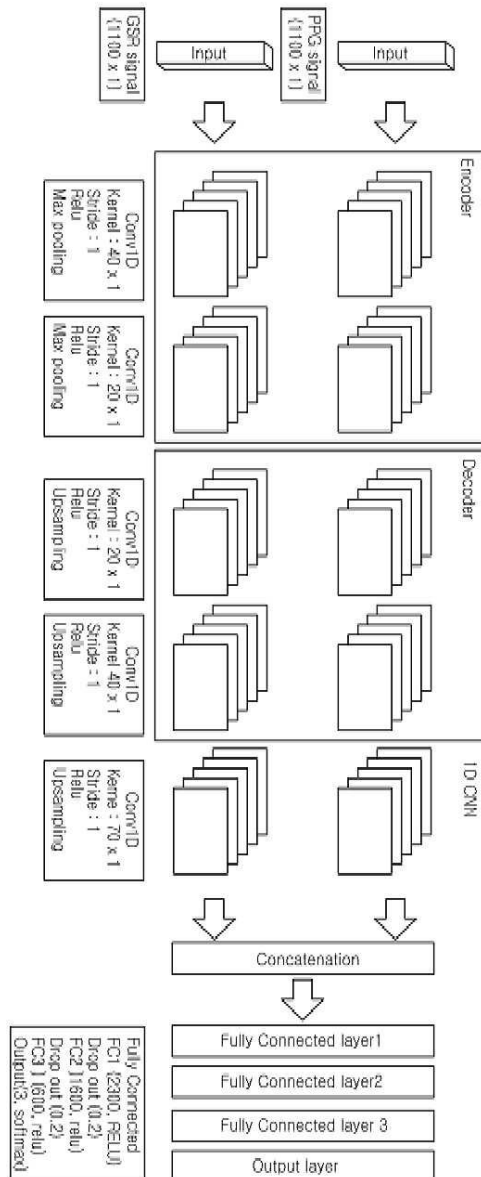
도면8



도면9



도면10



도면11

