

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2024-0081342
(43) 공개일자 2024년06월07일

- (51) 국제특허분류(Int. Cl.)
G06F 3/01 (2006.01) A61B 5/369 (2021.01)
A61B 5/389 (2021.01) G06F 3/00 (2006.01)
- (52) CPC특허분류
G06F 3/015 (2013.01)
A61B 5/369 (2022.01)
- (21) 출원번호 10-2023-0152497
(22) 출원일자 2023년11월07일
심사청구일자 2023년11월07일
- (30) 우선권주장
1020220163781 2022년11월30일 대한민국(KR)
- (71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
- (72) 발명자
이성환
서울특별시 강남구 언주로30길 13 (대림아크로빌, 대림아크로텔) B-2902
- 이정선
경상남도 창원시 진해구 조천북로 131 경화베스티움 102동 803호
(뒷면에 계속)
- (74) 대리인
특허법인위솔

전체 청구항 수 : 총 13 항

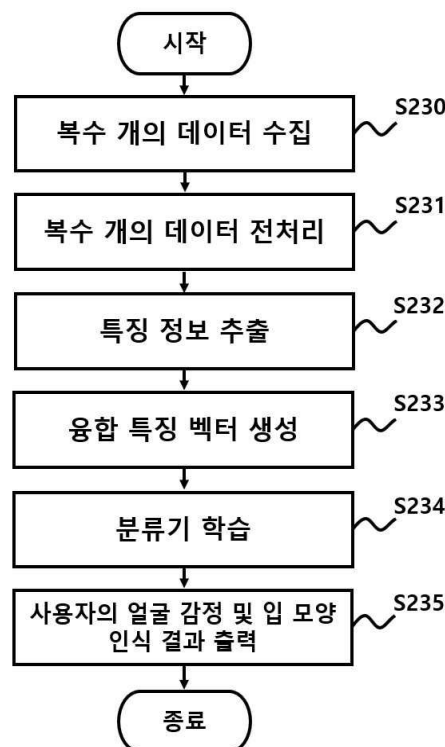
(54) 발명의 명칭 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법 및 시스템

(57) 요약

본 발명의 일 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법은 데이터 수집부가 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집하는 단계, 전처리부가 수집한 상기 복수 개의 데이터 각각을 전처리하는 단계, 특징 정보 추출부가 전처리된 상기 복수 개의 데이터로부터 특

(뒷면에 계속)

대표도 - 도2



징(Feature) 정보를 추출하는 단계, 융합 특징 벡터 생성부가 추출한 상기 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성하는 단계, 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계 및 디스플레이부가 학습한 상기 분류기를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력하는 단계를 포함하며, 상기 학습용 사용자로부터 수집하는 복수 개의 데이터는, 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터 및 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 상기 학습용 사용자의 얼굴을 촬영하여 수집하는 영상 데이터 중 적어도 하나를 포함하고, 상기 영상 데이터는, 상기 학습용 사용자의 음성 신호를 포함할 수 있다.

(52) CPC특허분류

A61B 5/389 (2022.01)

G06F 3/005 (2013.01)

G06F 2203/011 (2013.01)

(72) 발명자

이지원

서울특별시 도봉구 도봉로180길 6-23 동아에코빌아파트 106동 2203호

이서현

서울특별시 강남구 압구정로 321 한양아파트 9동 403호

고병관

강원도 동해시 평릉길 60-1 사랑드림아파트 102동 504호

이영은

경기도 성남시 분당구 중앙공원로 17 시범단지한양아파트 311동 701호

박지하

서울특별시 동대문구 무학로45길 30 (용두동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711152863
과제번호	2021-0-02068-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능 혁신 허브 연구 개발
기 여 율	1/2
과제수행기관명	고려대학교산학협력단
연구기간	2022.01.01 ~ 2022.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711093609
과제번호	2019-0-00079-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능핵심고급인재양성(R&D)
연구과제명	인공지능대학원지원(고려대학교)
기 여 율	1/2
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

데이터 수집부가 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집하는 단계;

전처리부가 수집한 상기 복수 개의 데이터 각각을 전처리하는 단계;

특징 정보 추출부가 전처리된 상기 복수 개의 데이터로부터 특징(Feature) 정보를 추출하는 단계;

융합 특징 벡터 생성부가 추출한 상기 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성하는 단계;

분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계; 및

디스플레이부가 학습한 상기 분류기를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력하는 단계;

를 포함하며,

상기 학습용 사용자로부터 수집하는 복수 개의 데이터는,

상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터 및 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 상기 학습용 사용자의 얼굴을 촬영하여 수집하는 영상 데이터 중 적어도 하나를 포함하고,

상기 영상 데이터는,

상기 학습용 사용자의 음성 신호를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 2

제1항에 있어서,

상기 데이터 수집부가 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집하는 단계는,

상기 데이터 수집부가 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 상기 학습용 사용자의 얼굴을 촬영하여 수집한 영상 데이터로부터 상기 학습용 사용자의 얼굴 이미지를 수집하는 단계;

변환부가 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터를 기 설정된 시간 구간 별로 샘플링하는 단계; 및

상기 변환부가 상기 샘플링한 생체 신호 데이터를 이미지로 변환시키는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 3

제2항에 있어서,

상기 변환부가 상기 샘플링한 생체 신호 데이터를 이미지로 변환시키는 단계는,

상기 변환부가 변환시킨 이미지를 멜 스펙트로그램(Mel Spectrogram) 이미지로 변환시키는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 4

제1항에 있어서,

상기 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계는,

상기 분류기 학습부가 상기 데이터 수집부가 수집한 상기 학습용 사용자의 얼굴 이미지 및 상기 변환부가 변환시킨 이미지를 이용하여 감정 분류 모델을 생성하는 단계; 및

상기 분류기 학습부가 생성된 감정 분류 모델을 이용하여 상기 학습용 사용자의 감정 상태를 분류하는 단계;
를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 5

제4항에 있어서,

상기 분류기 학습부가 생성된 감정 분류 모델을 이용하여 상기 학습용 사용자의 감정 상태를 분류하는 단계는,

상기 융합 특징 벡터를 이용하여 분류기가 상기 학습용 사용자의 시간의 흐름에 따른 감정 특징 변화를 분류하는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 6

제1항에 있어서,

상기 디스플레이부가 학습한 상기 분류기를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력하는 단계는

상기 디스플레이부가 상기 분류기 학습부가 분류한 상기 학습용 사용자의 감정 상태를 출력하는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 7

제1항에 있어서,

상기 생체 신호 데이터는,

뇌전도(EEG) 데이터, 근전도(EMG) 데이터 및 음성 신호 데이터 중 어느 하나 이상을 포함하며,

상기 근전도 데이터는,

상기 학습용 사용자의 얼굴 근육으로부터 측정하여 수집한 데이터인,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 8

제1항에 있어서,

상기 전처리부가 수집한 상기 복수 개의 데이터 각각을 전처리하는 단계는,

상기 전처리부가 상기 복수 개의 데이터 각각에 대하여 기 설정된 주파수 대역대를 획득하기 위하여 상기 복수 개의 데이터 각각을 밴드 패스 필터(Band-pass filtering) 및 노치 필터(Notch filter) 중 어느 하나를 통과시키는 단계; 및

상기 전처리부가 밴드 패스 필터(Band-pass filtering) 및 노치 필터(Notch filter) 중 어느 하나를 통과시킨 상기 복수 개의 데이터 각각에 대하여 안전도(Electrooculography)를 포함하는 뇌파 신호 아티팩트를 제거하는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 9

제1항에 있어서,

상기 특징 정보 추출부가 전처리된 상기 복수 개의 데이터로부터 특징(Feature) 정보를 추출하는 단계는,

전처리된 상기 복수 개의 데이터가 전처리된 영상 데이터인 경우,

상기 특징 정보 추출부가 전처리된 상기 복수 개의 데이터에서 상기 학습용 사용자의 얼굴 영역을 검출하는 단계; 및

상기 특징 정보 추출부가 검출된 상기 학습용 사용자의 얼굴 영역에서 얼굴 특징점을 탐색하는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 10

제1항에 있어서,

상기 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계는,

상기 분류기 학습부가 복수의 학습용 사용자에게서 공통으로 발생하는 패턴을 분리하는 단계;

상기 분류기 학습부가 분리 결과를 이용하여 서포트 벡터 머신, 인공신경망(Artificial Neural Network, ANN) 및 합성곱 신경망(Convolution Neural Network, CNN) 중 어느 하나의 학습 모델을 통해 상기 분류기를 학습시키는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 11

제1항에 있어서,

상기 융합 특징 벡터 생성부가 추출한 상기 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성하는 단계는,

상기 융합 특징 벡터 생성부가 전처리된 상기 복수 개의 데이터를 인코딩하는 단계;

상기 융합 특징 벡터 생성부가 인코딩한 결과 생성된 전처리된 상기 복수 개의 데이터에 대한 임베딩 벡터를 생성하는 단계;

상기 융합 특징 벡터 생성부가 상기 임베딩 벡터를 입력 데이터로 하여 전처리된 상기 복수 개의 데이터에 대응하는 얼굴 감정 특징 데이터를 인식하는 단계; 및

상기 융합 특징 벡터 생성부가 인식한 얼굴 감정 특징 데이터에 대응하는 융합 특징 벡터를 생성하는 단계;

를 포함하는

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 12

제1항에 있어서,

상기 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계는

상기 분류기 학습부가 상기 얼굴 감정 특징 데이터에 대응하는 융합 특징 벡터를 반영하여 상기 분류기를 학습시키는 단계;

를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법.

청구항 13

출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집하는 데이터 수집부;

상기 데이터 수집부가 수집한 상기 복수 개의 데이터 각각을 전처리하는 전처리부;

전처리된 상기 복수 개의 데이터로부터 특징(Feature) 정보를 추출하는 특징 정보 추출부;

추출한 상기 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성하는 융합 특징 벡터 생성부;

상기 융합 특징 벡터 생성부가 생성한 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 분류기 학습부; 및

상기 분류기를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력하는 디스플레이부;

를 포함하며,

상기 학습용 사용자로부터 수집하는 복수 개의 데이터는,

상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터 및 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 상기 학습용 사용자의 얼굴을 촬영하여 수집하는 영상 데이터 중 적어도 하나를 포함하고,

상기 영상 데이터는,

상기 학습용 사용자의 음성 신호를 포함하는,

발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법 및 시스템에 관한 것이다. 보다 자세하게는 뇌파를 비롯한 생체 신호 데이터를 이용함으로써 사용자의 얼굴 감정 및 입 모양을 정확하고 효과적으로 인식하여 판단할 수 있는 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법 및 시스템에 관한 것이다.

배경 기술

[0002] 생체 신호의 대표적인 예시인 근전도(Electromyography, EMG)는 운동으로 인한 근육 수축 상태에서 발생하는 전류 및 안정 상태에서의 전류를 기록하여 근육의 구조적 특이정보보다는 근육의 기능적 이상 여부를 진단할 수 있는 생체 신호이다.

[0003] 이러한 근전도를 측정하는 방법으로는 인간의 피부 표면에 전극을 부착한 후 전위를 유도하는 표면 전극법과 근육에 전극을 직접적으로 꽂아 넣어 근육 내 특정한 한 지점에서 발생한 활동 전위를 유도해 운동 단위의 활동을 검출하는 바늘 전극법이 있다.

[0004] 최근에는 근전도를 이용하여 개인의 근육 상태를 실시간으로 검사하고 측정하여 측정값에 대응하는 진단 결과를 출력하는 시스템을 구현하는 등 각종 재활 의학, 의학 연구 및 스포츠 과학 등 다양한 분야에서 활용하고 있다.

[0005] 또 다른 대표적 생체 신호인 뇌전도(Electroencephalography, EEG)는 인간의 뇌의 피질에 존재하는 신경 세포 간 정보 전달이 일어날 때 나타나는 전기적 신호가 합성되어 나타나는 미세한 신호를, 두피에 부착된 전극을 이용하여 측정된 전위인 뇌파를 측정하는 것이다.

- [0006] 특히, 뇌전도는 뇌에서 일어나는 활동을 실시간으로 기록하여 사용자의 발화 의도를 예측할 수 있는 중요한 수단이다.
- [0007] 인간의 두뇌 중에서도 특히 대뇌 피질은 인간의 두뇌에서 가장 발달한 부분으로 크게 두 개의 반구로 나누어 볼 수 있으며 기능적으로 운동, 감각, 그리고 통합 영역의 세 부분으로 나눌 수 있다.
- [0008] 뇌는 각 부위에 따라 각 기능이 세밀하게 나누어져 있으며 인간의 의식과 정신 활동에 따라 뇌파는 특정한 형태를 가지게 되며 실시간으로 변화한다.
- [0009] 뇌파의 측정 방법은 뇌의 어느 위치에서 뇌파를 측정하느냐에 따라 크게 세 가지 방법으로 나뉜다.
- [0010] 첫 번째, 가장 대중적으로 사용되는 방식은 두피 표면에 전극을 부착하여 비침습적으로 뇌파를 유도하는 일반적인 두피 뇌파(scalp EEG)를 측정하는 방식이다.
- [0011] 두 번째로는 대뇌 피질 표면에서 유도하는 피질 뇌파(electrocorticogram, ECoG)를 이용하는 방식이고, 마지막으로 뇌에 침습적으로 전극을 삽입하여 전위를 유도하는 심부 뇌파(deep EEG)가 있으며, 귀 착용형 뇌파 검사장비를 이용하여 귀-뇌파 기록(ear-EEG)을 측정할 수도 있다.
- [0012] 뇌-컴퓨터 인터페이스(Brain-computer interface)는 뇌의 활동으로부터 나타나는 전기적 신호를 통해 인간의 뇌와 컴퓨터와 같은 기계를 연결하고 의사소통하는 새로운 종류의 인터페이스 시스템이다.
- [0013] 뇌-컴퓨터 인터페이스 중에서도 비침습적 뇌-컴퓨터 인터페이스는 두피에서 비침습적으로 측정된 뇌파 정보를 이용해 사용자의 의도를 인지하여 외부 장치를 사용할 수 있게 하고 이를 제어할 수 있게 한다.
- [0014] 비침습적 뇌-컴퓨터 인터페이스는 수술 과정을 거칠 필요가 없으며 침습적 방법에 비하여 비교적 적은 비용과 간단한 방법으로 사용할 수 있기 때문에 다양한 방면에서 이를 적용하려는 연구가 활발하게 진행되고 있다. 이처럼 생체 신호란 뇌파부터 목소리까지 생체에서 유래하는 신호를 모두 포함하는 개념이다.
- [0015] 종래 기술로서, 대한민국 공개특허공보 제 10-2005-0063618호(2005.06.28)에서는 뇌파의 감마파 미세 변화를 이용한 긍정/부정 의사 인식방법 및 장치를 개시하고 있으나, 해당 기술은 단순히 긍정/부정 의사를 인식할 뿐, 사용자의 생체 신호 및 뇌파를 통해 사용자의 얼굴 감정 및 입 모양을 인식할 수 없으며, 주변 환경에서 잡음이 발생할 경우 사용자가 말하고 있지 않음에도 불구하고, 말을 하고 있는 것으로 결과가 두출되어 발화하는 것과 다른 결과가 나올 수 있는 문제가 있다.
- [0016] 이처럼 기존의 뇌-컴퓨터 인터페이스 기술은 뇌파를 통해 음성 및 텍스트로 출력하는 장치 및 방법에 치중되어 있으며 뇌파를 이용하여 입 모양을 예측하거나 재구성하는 방법이나 그에 대한 장치는 찾기 어려운 실정이다.

선행기술문헌

특허문헌

- [0017] (특허문헌 0001) 대한민국 공개특허공보 제 10-2005-0063618호(2005.06.28)

발명의 내용

해결하려는 과제

- [0018] 본 발명이 해결하고자 하는 기술적 과제는 사용자의 뇌파를 기반으로 입 모양을 예측하여 재구성하는 뇌-컴퓨터 인터페이스 시스템을 구현하고, 여러 잡음 신호가 섞여 들어가는 뇌파 신호의 특징을 감안하여, 여기서 더 나아가 전반적인 생체 신호를 추가적인 인공지능 모델 학습 데이터로 사용하여 기존 뇌-컴퓨터 인터페이스 장치보다 향상된 신뢰도 및 정확도를 가지며 보다 사용자 친화적인 인터페이스를 제공하는 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법 및 시스템을 제공하는 것이다.
- [0019] 본 발명의 기술적 과제들은 이상에서 언급한 기술적 과제들로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

- [0020] 상기 기술적 과제를 달성하기 위한 본 발명의 일 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법은 데이터 수집부가 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집하는 단계, 전처리부가 수집한 상기 복수 개의 데이터 각각을 전처리하는 단계, 특징 정보 추출부가 전처리된 상기 복수 개의 데이터로부터 특징(Feature) 정보를 추출하는 단계, 융합 특징 벡터 생성부가 추출한 상기 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성하는 단계, 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계 및 디스플레이부가 학습한 상기 분류기를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력하는 단계를 포함하며, 상기 학습용 사용자로부터 수집하는 복수 개의 데이터는, 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터 및 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 상기 학습용 사용자의 얼굴을 촬영하여 수집하는 영상 데이터 중 적어도 하나를 포함하고, 상기 영상 데이터는, 상기 학습용 사용자의 음성 신호를 포함할 수 있다.
- [0021] 상기 일 실시 예에 따르면, 상기 데이터 수집부가 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집하는 단계는 상기 데이터 수집부가 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 상기 학습용 사용자의 얼굴을 촬영하여 수집한 영상 데이터로부터 상기 학습용 사용자의 얼굴 이미지를 수집하는 단계, 변환부가 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터를 기 설정된 시간 구간 별로 샘플링하는 단계 및 상기 변환부가 상기 샘플링한 생체 신호 데이터를 이미지로 변환시키는 단계를 포함할 수 있다.
- [0022] 상기 일 실시 예에 따르면, 상기 변환부가 상기 샘플링한 생체 신호 데이터를 이미지로 변환시키는 단계는 상기 변환부가 변환시킨 이미지를 멜 스펙트로그램(Mel Spectrogram) 이미지로 변환시키는 단계를 포함할 수 있다.
- [0023] 상기 일 실시 예에 따르면, 상기 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계는 상기 분류기 학습부가 상기 데이터 수집부가 수집한 상기 학습용 사용자의 얼굴 이미지 및 상기 변환부가 변환시킨 이미지를 이용하여 감정 분류 모델을 생성하는 단계 및 상기 분류기 학습부가 생성된 감정 분류 모델을 이용하여 상기 학습용 사용자의 감정 상태를 분류하는 단계를 포함하고,
- [0024] 상기 일 실시 예에 따르면, 상기 분류기 학습부가 생성된 감정 분류 모델을 이용하여 상기 학습용 사용자의 감정 상태를 분류하는 단계는 상기 융합 특징 벡터를 이용하여 분류기가 상기 학습용 사용자의 시간의 흐름에 따른 감정 특징 변화를 분류하는 단계를 포함할 수 있다.
- [0025] 상기 일 실시 예에 따르면, 상기 디스플레이부가 학습한 상기 분류기를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력하는 단계는 상기 디스플레이부가 상기 분류기 학습부가 분류한 상기 학습용 사용자의 감정 상태를 출력하는 단계를 포함할 수 있다.
- [0026] 상기 일 실시 예에 따르면, 상기 생체 신호 데이터는, 뇌전도(EEG) 데이터, 근전도(EMG) 데이터 및 음성 신호 데이터 중 어느 하나 이상을 포함하며, 상기 근전도 데이터는, 상기 학습용 사용자의 얼굴 근육으로부터 측정하여 수집한 데이터일 수 있다.
- [0027] 상기 일 실시 예에 따르면, 상기 전처리부가 수집한 상기 복수 개의 데이터 각각을 전처리하는 단계는 상기 전처리부가 상기 복수 개의 데이터 각각에 대하여 기 설정된 주파수 대역대를 획득하기 위하여 상기 복수 개의 데이터 각각을 밴드 패스 필터(Band-pass filtering) 및 노치 필터(Notch filter) 중 어느 하나를 통과시키는 단계 및 상기 전처리부가 밴드 패스 필터(Band-pass filtering) 및 노치 필터(Notch filter) 중 어느 하나를 통과시킨 상기 복수 개의 데이터 각각에 대하여 안전도(Electrooculography)를 포함하는 뇌파 신호 아티팩트를 제거하는 단계를 포함할 수 있다.
- [0028] 상기 일 실시 예에 따르면, 상기 특징 정보 추출부가 전처리된 상기 복수 개의 데이터로부터 특징(Feature) 정보를 추출하는 단계는 전처리된 상기 복수 개의 데이터가 전처리된 영상 데이터인 경우, 상기 특징 정보 추출부가 전처리된 상기 복수 개의 데이터에서 상기 학습용 사용자의 얼굴 영역을 검출하는 단계 및 상기 특징 정보 추출부가 검출된 상기 학습용 사용자의 얼굴 영역에서 얼굴 특징점을 탐색하는 단계를 포함할 수 있다.
- [0029] 상기 일 실시 예에 따르면, 상기 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계는 상기 분류기 학습부가 복수의 학습용 사용자에게서 공통으로 발생하는 패턴을 분리하는 단계, 상기 분류기 학습부가 분리 결과를 이용하여 서포트 벡터 머신, 인공신경망(Artificial Neural Network, ANN) 및 합성곱 신경망(Convolution Neural Network, CNN) 중 어느 하나의 학습 모델을 통해 상기 분류기를 학습시키는 단계를 포함할 수 있다.

- [0030] 상기 일 실시 예에 따르면, 상기 융합 특징 벡터 생성부가 추출한 상기 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성하는 단계는 상기 융합 특징 벡터 생성부가 전처리된 상기 복수 개의 데이터를 인코딩하는 단계, 상기 융합 특징 벡터 생성부가 인코딩한 결과 생성된 전처리된 상기 복수 개의 데이터에 대한 임베딩 벡터를 생성하는 단계, 상기 융합 특징 벡터 생성부가 상기 임베딩 벡터를 입력 데이터로 하여 전처리된 상기 복수 개의 데이터에 대응하는 얼굴 감정 특징 데이터를 인식하는 단계 및 상기 융합 특징 벡터 생성부가 인식한 얼굴 감정 특징 데이터에 대응하는 융합 특징 벡터를 생성하는 단계를 포함하고,
- [0031] 상기 일 실시 예에 따르면, 상기 분류기 학습부가 생성된 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 단계는 상기 분류기 학습부가 상기 얼굴 감정 특징 데이터에 대응하는 융합 특징 벡터를 반영하여 상기 분류기를 학습시키는 단계를 포함할 수 있다.
- [0032] 상기 일 실시 예에 따르면, 상기 변환부가 상기 샘플링한 생체 신호 데이터를 이미지로 변환시키는 단계는 상기 변환부가 변환시킨 이미지를 멜 스펙트로그램(Mel Spectrogram) 이미지로 변환시키는 단계를 포함할 수 있다.
- [0033] 상기 일 실시 예에 따르면, 상기 분류기 학습부가 생성된 감정 분류 모델을 이용하여 상기 학습용 사용자의 감정 상태를 분류하는 단계는 상기 융합 특징 벡터를 이용하여 분류기가 상기 학습용 사용자의 시간의 흐름에 따른 감정 특징 변화를 분류하는 단계를 포함할 수 있다.
- [0034] 상기 기술적 과제를 달성하기 위한 본 발명의 또 다른 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템은 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집하는 데이터 수집부, 상기 데이터 수집부가 수집한 상기 복수 개의 데이터 각각을 전처리하는 전처리부, 전처리된 상기 복수 개의 데이터로부터 특징(Feature) 정보를 추출하는 특징 정보 추출부, 추출한 상기 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성하는 융합 특징 벡터 생성부, 상기 융합 특징 벡터 생성부가 생성한 상기 융합 특징 벡터를 이용하여 분류기를 학습시키는 분류기 학습부 및 상기 분류기를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력하는 디스플레이부를 포함하며, 상기 학습용 사용자로부터 수집하는 복수 개의 데이터는, 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터 및 상기 학습용 사용자가 상기 출력되는 제시문을 읽는 동안 상기 학습용 사용자의 얼굴을 촬영하여 수집하는 영상 데이터 중 적어도 하나를 포함하고, 상기 영상 데이터는, 상기 학습용 사용자의 음성 신호를 포함할 수 있다.

발명의 효과

- [0035] 상기와 같은 본 발명에 따르면, 사용자의 뇌파를 기반으로 입 모양을 예측하여 재구성하는 뇌-컴퓨터 인터페이스 시스템을 구현하고, 여러 잡음 신호가 섞여 들어가는 뇌파 신호의 특징을 감안하여, 여기서 더 나아가 전반적인 생체 신호를 추가적인 인공지능 모델 학습 데이터로 사용하여 기존 뇌-컴퓨터 인터페이스 장치보다 향상된 신뢰도 및 정확도를 가지며 보다 사용자 친화적인 인터페이스를 제공하는 효과가 있다.
- [0036] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 아래의 기재로부터 통상의 기술자에게 명확하게 이해 될 수 있을 것이다.

도면의 간단한 설명

- [0037] 도 1은 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템이 포함하는 전체 구성을 예시적으로 도시한 도면이다.
- 도 2는 본 발명의 제2 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법의 대표적인 단계를 나타낸 순서도이다.
- 도 3은 멜 스펙트로그램에 대한 예시를 나타낸 도면이다.
- 도 4는 주파수 필터링에 대한 설명을 위한 도면이다.
- 도 5는 학습용 사용자의 얼굴 영역에서 탐색되는 얼굴 특징점을 나타난 도면이다.
- 도 6은 본 발명의 제3 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템의 개략적인 구성을 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0038] 본 발명의 목적과 기술적 구성 및 그에 따른 작용 효과에 관한 자세한 사항은 본 발명의 명세서에 첨부된 도면에 의거한 이하의 상세한 설명에 의해 보다 명확하게 이해될 것이다. 첨부된 도면을 참조하여 본 발명에 따른 실시 예를 상세하게 설명한다.
- [0039] 본 명세서에서 개시되는 실시 예들은 본 발명의 범위를 한정하는 것으로 해석되거나 이용되지 않아야 할 것이다. 이 분야의 통상의 기술자에게 본 명세서의 실시 예를 포함한 설명은 다양한 응용을 갖는다는 것이 당연하다. 따라서, 본 발명의 상세한 설명에 기재된 임의의 실시 예들은 본 발명을 보다 잘 설명하기 위한 예시적인 것이며 본 발명의 범위가 실시 예들로 한정되는 것을 의도하지 않는다.
- [0040] 도면에 표시되고 아래에 설명되는 기능 블록들은 가능한 구현의 예들일 뿐이다. 다른 구현들에서는 상세한 설명의 사상 및 범위를 벗어나지 않는 범위에서 다른 기능 블록들이 사용될 수 있다. 또한, 본 발명의 하나 이상의 기능 블록이 개별 블록들로 표시되지만, 본 발명의 기능 블록들 중 하나 이상은 동일 기능을 실행하는 다양한 하드웨어 및 소프트웨어 구성들의 조합일 수 있다.
- [0041] 또한, 어떤 구성요소들을 포함한다는 표현은 "개방형"의 표현으로서 해당 구성요소들이 존재하는 것을 단순히 지칭할 뿐이며, 추가적인 구성요소들을 배제하는 것으로 이해되어서는 안 된다.
- [0042] 나아가 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급될 때에는, 그 다른 구성요소에 직접적으로 연결 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 한다.
- [0043] 이하에서는 도면들을 참조하여 본 발명의 세부적인 실시 예들에 대해 살펴보도록 한다.
- [0044] 도 1은 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템이 포함하는 전체 구성을 예시적으로 도시한 도면이다.
- [0045] 그러나 이는 본 발명의 목적을 달성하기 위한 바람직한 실시 예일 뿐이며, 필요에 따라 일부 구성이 추가되거나 삭제될 수 있고, 어느 한 구성이 수행하는 역할을 다른 구성이 함께 수행할 수도 있음은 물론이다.
- [0046] 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(100)은 프로세서(10), 네트워크 인터페이스(20), 메모리(30), 스토리지(40) 및 이들을 연결하는 데이터 버스(50)를 포함할 수 있으며, 기타 본 발명의 목적을 달성함에 있어 요구되는 부가적인 구성들을 더 포함할 수 있음은 물론이라 할 것이다.
- [0047] 프로세서(10)는 각 구성의 전반적인 동작을 제어한다. 프로세서(10)는 CPU(Central Processing Unit), MPU(Micro Processor Unit), MCU(Micro Controller Unit) 또는 본 발명이 속하는 기술 분야에서 널리 알려져 있는 형태의 인공지능 프로세서 중 어느 하나일 수 있다. 아울러, 프로세서(10)는 본 발명의 제2 실시 예에 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법을 수행하기 위한 적어도 하나의 애플리케이션 또는 프로그램에 대한 연산을 수행할 수 있는바, 이를 위해 프로세서(10)에는 소정의 구조를 나타내는 딥러닝 모델을 포함할 수 있으나, 이에 대해서는 후술하도록 한다.
- [0048] 네트워크 인터페이스(20)는 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(100)의 유무선 인터넷 통신을 지원하며, 그 밖의 공지의 통신 방식을 지원할 수도 있다. 따라서 네트워크 인터페이스(20)는 그에 따른 통신 모듈을 포함하여 구성될 수 있다.
- [0049] 메모리(30)는 각종 정보, 명령 및/또는 정보를 저장하며, 본 발명의 제2 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법을 수행하기 위해 스토리지(40)로부터 하나 이상의 컴퓨터 프로그램(41)을 로드할 수 있다. 도 1에서는 메모리(30)의 하나로 RAM을 도시하였으나 이와 더불어 다양한 저장 매체를 메모리(30)로 이용할 수 있음은 물론이다.
- [0050] 스토리지(40)는 하나 이상의 컴퓨터 프로그램(41) 및 대용량 네트워크 정보(42)를 비임시적으로 저장할 수 있다. 이러한 스토리지(40)는 ROM(Read Only Memory), EPROM(Erasable Programmable ROM), EEPROM(Electrically Erasable Programmable ROM), 플래시 메모리 등과 같은 비휘발성 메모리, 하드 디스크(HDD), 보조 저장 매체(SSD), 착탈형 디스크, 또는 본 발명이 속하는 기술 분야에서 널리 알려져 있는 임의의 형태의 컴퓨터로 읽을 수 있는 기록 매체 중 어느 하나일 수 있다.
- [0051] 데이터 버스(50)는 이상 설명한 프로세서(10), 네트워크 인터페이스(20), 메모리(30) 및 스토리지(40) 사이의 명령 및/또는 정보의 이동 경로가 된다.
- [0052] 이상 간단하게 설명한 본 발명의 제1 실시 예에 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(10

0)은 독립된 디바이스의 형태, 예를 들어 전자 기기나 서버(클라우드 포함)의 형태일 수 있으며, 여기서 전자 기기는 한 장소에 고정 설치되어 사용하는 데스크톱 PC, 서버 디바이스 등과 같은 기기 뿐만 아니라, 스마트폰, 태블릿 PC, 노트북 PC, PDA, PMP 등과 같이 휴대가 용이한 포터블 기기 등이라도 무방한바, 프로세서(10)에 해당하는 CPU 등이 설치되고 네트워크 기능만 보유하고 있는 전자 기기라면 어떠한 것이라도 무방하다 할 것이다.

[0053] 이하, 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(100)이 독립된 디바이스인 전자 기기 형태 중, "서버"의 형태임을 전제로 발화시 사용자의 얼굴 감정 및 입 모양을 인식하고자 하는 사용자의 사용자 단말(미도시)에 설치된 전용 어플리케이션을 통해 본 발명의 제2 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법을 제공하는 과정에 대하여 도 2를 참조하여 설명하도록 한다.

[0054] 도 2는 본 발명의 제2 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 방법의 대표적인 단계를 나타낸 순서도이다.

[0055] 그러나 이는 본 발명의 목적을 달성함에 있어서 바람직한 실시 예일 뿐이며, 필요에 따라 일부 단계가 추가 또는 삭제될 수 있음은 물론이고, 어느 한 단계가 다른 단계에 포함되어 수행될 수도 있다.

[0056] 한편, 각 단계는 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(100)을 통해 이루어지는 것을 전제로 하며, 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(100)이 "서버"의 형태임을 전제로 하였기 때문에 사용자 단말(미도시)에 설치된 전용 어플리케이션을 본 발명의 제1 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(100)과 동일한 의미로 볼 것이고, 설명의 편의를 위해 이들 모두를 "시스템(100)"로 명명하도록 한다.

[0057] 우선, 시스템(100)의 경우, 데이터 수집부(1100)가 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집할 수 있다.(S230)

[0058] 여기서 학습용 사용자로부터 수집하는 복수 개의 데이터는, 학습용 사용자가 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터 및 학습용 사용자가 출력되는 제시문을 읽는 동안 학습용 사용자의 얼굴을 촬영하여 수집하는 영상 데이터 중 적어도 하나를 포함할 수 있다.

[0059] 보다 구체적으로, 데이터 수집부(1100)는 사용자의 두피와 얼굴, 팔 등 여러 상체의 부위 곳곳에 부착한 전극으로부터 발생한 뇌전도 데이터 및 근전도 데이터를 수집할 수 있다.

[0060] 그리고 데이터 수집부(1100)는 사용자의 발화 기반 뇌파와 더불어 사용자에게 제시한 해당 음절 소리를 낼 때의 입모양과 안면에 나타나는 얼굴 감정을 녹화하여 디지털 신호로 변환시켜 음성 신호 데이터 및 영상 데이터를 수집할 수 있다.

[0061] 또한, 저장부(1900)는 데이터 수집부(1100)가 수집한 복수 개의 데이터를 저장할 수 있다.

[0062] 그리고 디스플레이부(1600)는 학습용 사용자가 읽을 수 있도록 제시문을 출력할 수 있다.

[0063] 보다 구체적으로 디스플레이부(1600)는 사용자의 발화 기반 뇌파를 유발하고 함께 발생하는 생체 신호를 수집하기 위하여 학습용 사용자가 입모양에 변화를 주도록 학습용 사용자에게 음절, 단어 및 문장 등 다양한 제시문을 출력할 수 있다.

[0064] 또한, 영상 데이터는 학습용 사용자의 음성 신호를 포함할 수 있다.

[0065] 그리고 생체 신호 데이터는 뇌전도(EEG) 데이터, 근전도(EMG) 데이터 및 음성 신호 데이터 중 어느 하나 이상을 포함할 수 있다.

[0066] 또한, 근전도 데이터는 학습용 사용자의 얼굴 근육으로부터 측정하여 수집한 데이터를 의미할 수 있다.

[0067] 그리고 데이터 수집부(1100)는 학습용 사용자가 출력되는 제시문을 읽는 동안 학습용 사용자의 얼굴을 촬영하여 수집한 영상 데이터로부터 학습용 사용자의 얼굴 이미지를 수집할 수 있다.

[0068] 또한, 변환부(1700)는 학습용 사용자가 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터를 기 설정된 시간 구간 별로 샘플링할 수 있다.

[0069] 그리고 변환부(1700)는 샘플링한 생체 신호 데이터를 이미지로 변환시킬 수 있다.

[0070] 또한, 변환부(1700)는 변환시킨 이미지를 멜 스펙트로그램(Mel Spectrogram) 이미지로 변환시킬 수 있다.

[0071] 변환부(1700)가 이미지를 멜 스펙트로그램(Mel Spectrogram) 이미지로 변환시킴으로써, 동일한 생체 신호 데이

터베이스를 기반으로 하여 생체 신호 간 유사성과 각 신호의 두드러지는 특징점을 기준으로 분리하여 수치화할 수 있는 여러 가지 특징 추출 기법들을 이용하여 각 생체 신호의 특징점을 기준으로 데이터를 분리하여 패턴화할 수 있을 것이다.

- [0072] 또한, 도 3은 멜 스펙트로그램에 대한 예시를 나타낸 도면이며, 도 3을 참고할 경우, 멜 스펙트로그램에 대하여 보다 구체적으로 확인할 수 있을 것이다.
- [0073] 그리고 전처리부(1200)는 수집한 복수 개의 데이터 각각을 전처리할 수 있다.(S231)
- [0074] 또한, 전처리부(1200)는 복수 개의 데이터 각각에 대하여 기 설정된 주파수 대역대를 획득하기 위하여 복수 개의 데이터 각각을 밴드 패스 필터(Band-pass filtering) 및 노치 필터(Notch filter) 중 어느 하나를 통과시킬 수 있다.
- [0075] 여기서, 전처리부(1200)가 복수 개의 데이터 각각을 밴드 패스 필터(Band-pass filtering)에 통과시킬 경우, 복수 개의 데이터 각각에 대하여 주파수 대역을 제한하여 원하는 주파수 범위 내의 신호를 강조하고 다른 주파수 대역의 잡음을 제거할 수 있다.
- [0076] 이를 통하여 생체 신호 데이터를 포함하는 복수 개의 데이터의 유용한 정보를 강조하고 불필요한 주파수 대역을 제거함으로써 데이터의 정확도와 정밀도를 올릴 수 있다.
- [0077] 또한, 전처리부(1200)가 복수 개의 데이터 각각을 노치 필터(Notch filter)에 통과시킬 경우, 특정 주파수 대역의 잡음을 제거할 수 있다.
- [0078] 즉, 노치 필터 통과시킴으로써, 전처리부(1200)는 복수 개의 데이터 각각에 대하여 주파수 스펙트럼에서 특정 주파수를 선택하고 그 주파수 주위의 데이터를 제거할 수 있다. 이는 주변 환경에서 발생하는 전력 라인 주파수의 잡음과 같이 특정 주파수의 성분이 원치 않는 신호로 간주될 때 유용하게 이용될 수 있을 것이다.
- [0079] 도 4는 주파수 필터링에 대한 설명을 위한 도면이며, 상술한 밴드 패스 필터(Band-pass filtering) 및 노치 필터(Notch filter)에 대한 것은 도 4를 참고할 경우 보다 정확하게 판단할 수 있을 것이다.
- [0080] 그리고 전처리부(1200)는 밴드 패스 필터(Band-pass filtering) 및 노치 필터(Notch filter) 중 어느 하나를 통과시킨 복수 개의 데이터 각각에 대하여 안전도(Electrooculography)를 포함하는 뇌파 신호 아티팩트를 제거할 수 있다.
- [0081] 또한, 특징 정보 추출부(1300)는 전처리된 복수 개의 데이터로부터 특징(Feature) 정보를 추출할 수 있다.(S232)
- [0082] 보다 구체적으로 특징 정보 추출부(1300)는 공통 공간 패턴 분석법(CSP), 선형 판별 분석법(LDA) 등의 다양한 특징 추출 기법 등을 통하여 동일한 제시문에 대한 발화로부터 나타나는 학습용 사용자의 영상 데이터 및 발화 기반 생체 신호 데이터 간의 관련성 특징을 추출할 수 있다.
- [0083] 그리고 특징 추출부는 획득한 발화 시의 시간 동안 생체 신호 데이터 및 학습용 사용자의 영상 데이터의 특징적인 얼굴 감정 및 입모양 변화 값 간의 관련성 분석을 수행할 수 있다.
- [0084] 그리고 분류기 학습부(1500)는 상술한 관련성 분석 수행 결과를 반영하여 사용자의 생체 신호 데이터를 분리하는 인공지능 모델을 분류기(1800)에 학습시킬 수 있다.
- [0085] 그리고 전처리된 복수 개의 데이터가 전처리된 영상 데이터인 경우, 특징 정보 추출부(1300)가 전처리된 복수 개의 데이터에서 학습용 사용자의 얼굴 영역을 검출할 수 있다.
- [0086] 도 5는 학습용 사용자의 얼굴 영역에서 탐색되는 얼굴 특징점을 나타낸 도면이다.
- [0087] 도 5을 참고하면, 학습용 사용자의 얼굴 영역 및 얼굴 특징점에 대하여 확인할 수 있을 것이다.
- [0088] 또한, 특징 정보 추출부(1300)가 검출된 학습용 사용자의 얼굴 영역에서 얼굴 특징점을 탐색할 수 있다.
- [0089] 보다 구체적으로, 특징 정보 추출부(1300)는 ASM (Active shape model), AAM (Active appearance model), ESM(Explicit Shape Model) 및 SDM(Supervised Descebt Model) 등의 점 분포 모델 알고리즘을 사용하여 전처리된 복수 개의 데이터에서 학습용 사용자의 얼굴 영역을 검출할 수 있고, 검출된 학습용 사용자의 얼굴 영역에서 얼굴 특징점을 탐색할 수 있다.
- [0090] 그리고 융합 특징 벡터 생성부(1400)는 특징 정보 추출부(1300)가 추출한 복수 개의 데이터에 대한 복수 개의

특징 정보를 융합하여 융합 특징 벡터를 생성할 수 있다.(S233)

- [0091] 또한, 융합 특징 벡터 생성부(1400)는 전처리된 복수 개의 데이터를 인코딩할 수 있다.
- [0092] 그리고 융합 특징 벡터 생성부(1400)는 인코딩한 결과 생성된 전처리된 복수 개의 데이터에 대한 임베딩 벡터를 생성할 수 있다.
- [0093] 또한, 융합 특징 벡터 생성부(1400)는 임베딩 벡터를 입력 데이터로 하여 전처리된 복수 개의 데이터에 대응하는 얼굴 감정 특징 데이터를 인식할 수 있다.
- [0094] 여기서, 얼굴 감정 특징 데이터는 얼굴 감정을 확인할 수 있는 특징 데이터를 의미하는 것으로 얼굴 감정은 얼굴에서 느껴지는 감정을 의미하며 실제 사용자의 기쁨, 슬픔 등과 같은 감정을 의미하는 것은 아니다.
- [0095] 그리고 융합 특징 벡터 생성부(1400)는 인식한 얼굴 감정 특징 데이터에 대응하는 융합 특징 벡터를 생성할 수 있다.
- [0096] 보다 구체적으로, 융합 특징 벡터 생성부(1400)는 생체 신호 데이터를 변환시켜 도출된 멜 스펙트로그램 이미지 및 얼굴 감정 특징 데이터를 융합하여 융합 특징 벡터를 생성할 수 있다.
- [0097] 또한, 분류기 학습부(1500)는 생성된 융합 특징 벡터를 이용하여 분류기(1800)를 학습시킬 수 있다.(S234)
- [0098] 그리고 분류기 학습부(1500)는 데이터 수집부(1100)가 수집한 학습용 사용자의 얼굴 이미지 및 변환부(1700)가 변환시킨 이미지를 이용하여 감정 분류 모델을 생성할 수 있다.
- [0099] 또한, 분류기 학습부(1500)는 생성된 감정 분류 모델을 이용하여 학습용 사용자의 감정 상태를 분류할 수 있다.
- [0100] 그리고 융합 특징 벡터를 이용하여 분류기(1800)가 학습용 사용자의 시간의 흐름에 따른 감정 특징 변화를 분류할 수 있다.
- [0101] 그리고 분류기 학습부(1500)는 복수의 학습용 사용자에게서 공통으로 발생하는 패턴을 분리할 수 있다.
- [0102] 그리고 분류기 학습부(1500)는 융합 특징 벡터 생성부(1400)가 생체 신호 데이터를 변환시켜 도출된 멜 스펙트로그램 이미지 및 얼굴 감정 특징 데이터를 융합하여 생성한 융합 특징 벡터를 이용하여 분류기(1800)를 학습시킬 수 있다.
- [0103] 또한, 분류기 학습부(1500)는 분리 결과를 이용하여 서포트 벡터 머신, 인공신경망(Artificial Neural Network, ANN) 및 합성곱 신경망(Convolution Neural Network, CNN) 중 어느 하나의 학습 모델을 통해 분류기(1800)를 학습시킬 수 있다.
- [0104] 그리고 분류기 학습부(1500)는 얼굴 감정 특징 데이터에 대응하는 융합 특징 벡터를 반영하여 분류기(1800)를 학습시킬 수 있다.
- [0105] 그리고 디스플레이부(1600)는 분류기 학습부(1500)가 학습한 분류기(1800)를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력할 수 있다.(S235)
- [0106] 또한, 디스플레이부(1600)는 분류기 학습부(1500)가 분류한 학습용 사용자의 감정 상태를 출력할 수 있다.
- [0107] 도 6은 본 발명의 제3 실시 예에 따른 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템의 개략적인 구성을 나타낸 도면이다.
- [0108] 도 6을 참고하면, 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템(1000)은 데이터 수집부(1100), 전처리부(1200), 특징 정보 추출부(1300), 융합 특징 벡터 생성부(1400), 분류기 학습부(1500), 디스플레이부(1600), 변환부(1700), 분류부(1800) 및 저장부(1900)를 포함할 수 있다.
- [0109] 보다 구체적으로 데이터 수집부(1100)는 출력되는 제시문을 읽고 그대로 발화하는 학습용 사용자로부터 복수 개의 데이터를 수집할 수 있다.
- [0110] 그리고 전처리부(1200)는 데이터 수집부(1100)가 수집한 복수 개의 데이터 각각을 전처리할 수 있다.
- [0111] 또한, 특징 정보 추출부(1300)는 전처리된 복수 개의 데이터로부터 특징(Feature) 정보를 추출할 수 있다.
- [0112] 그리고 융합 특징 벡터 생성부(1400)는 추출한 복수 개의 데이터에 대한 복수 개의 특징 정보를 융합하여 융합 특징 벡터를 생성할 수 있다.

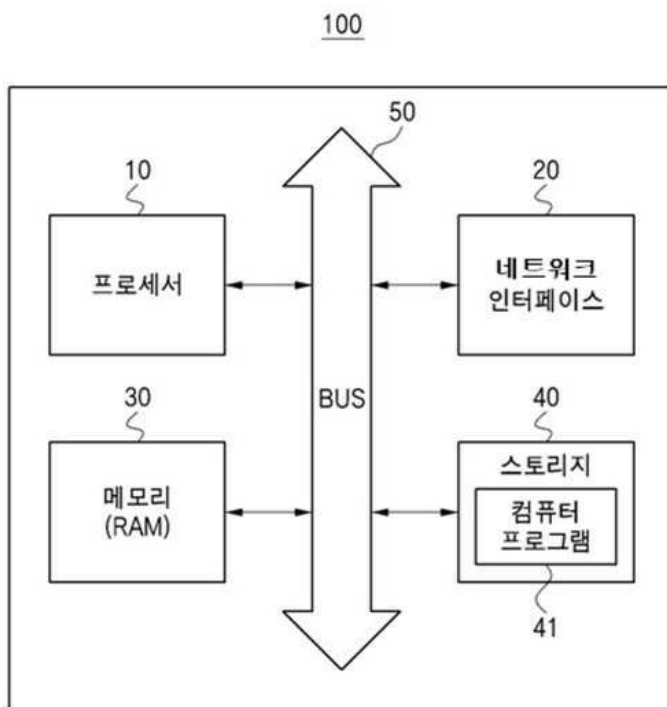
- [0113] 또한, 분류기 학습부(1500)는 융합 특징 벡터 생성부(1400)가 생성한 융합 특징 벡터를 이용하여 분류기(1800)를 학습시킬 수 있다.
- [0114] 그리고 디스플레이부(1600)는 분류기(1800)를 이용하여 임의의 사용자의 발화시 해당 사용자의 얼굴 감정 및 입 모양 인식 결과를 출력할 수 있다.
- [0115] 여기서 학습용 사용자로부터 수집하는 복수 개의 데이터는, 학습용 사용자가 출력되는 제시문을 읽는 동안 측정하여 수집하는 생체 신호 데이터 및 학습용 사용자가 출력되는 제시문을 읽는 동안 학습용 사용자의 얼굴을 촬영하여 수집하는 영상 데이터 중 적어도 하나를 포함할 수 있다.
- [0116] 그리고 영상 데이터는, 학습용 사용자의 음성 신호를 포함할 수 있다.
- [0117] 이상 첨부된 도면을 참조하여 본 발명의 실시 예들을 설명하였지만, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 그 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 실시될 수 있다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시 예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다.

부호의 설명

- [0118] 10: 프로세서
- 20: 네트워크 인터페이스
- 30: 메모리
- 40: 스토리지
- 41: 컴퓨터 프로그램
- 50: 정보 버스
- 100, 1000: 발화시 사용자의 얼굴 감정 및 입 모양을 인식하는 시스템
- 1100: 데이터 수집부
- 1200: 전처리부
- 1300: 특징 정보 추출부
- 1400: 융합 특징 벡터 생성부
- 1500: 분류기 학습부
- 1600: 디스플레이부
- 1700: 변환부
- 1800: 분류기
- 1900: 저장부

도면

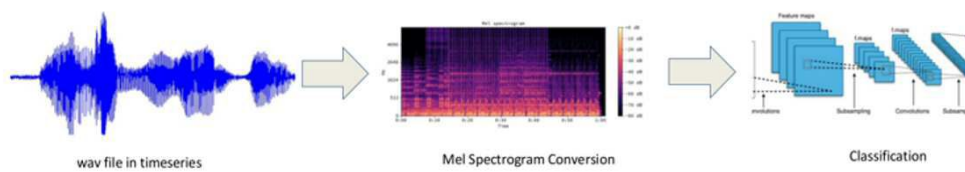
도면1



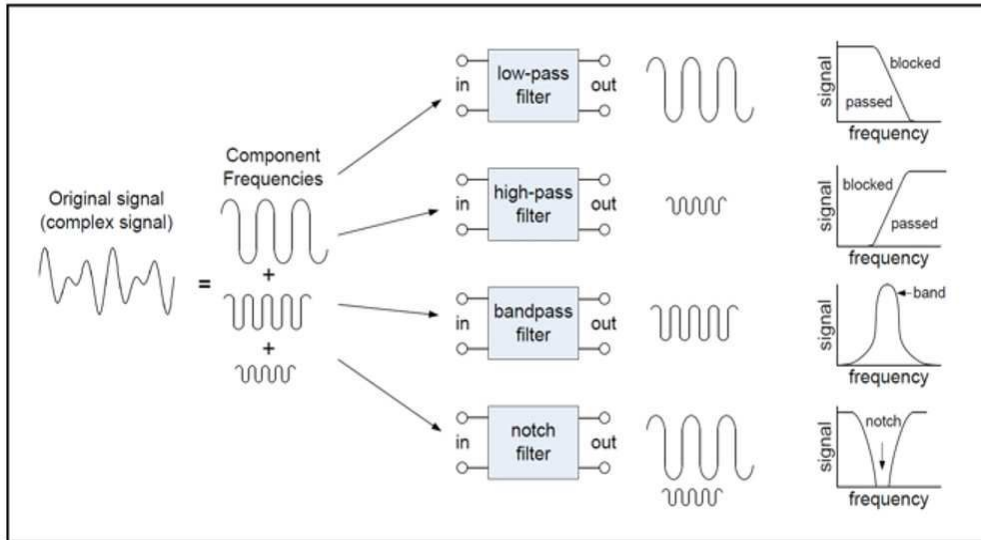
도면2



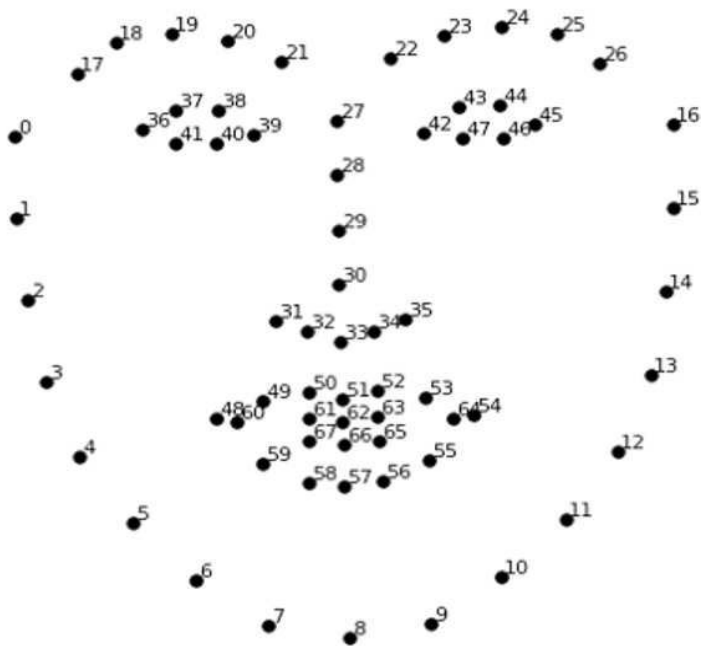
도면3



도면4



도면5



도면6

