

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일

2024년 2월 8일 (08.02.2024)



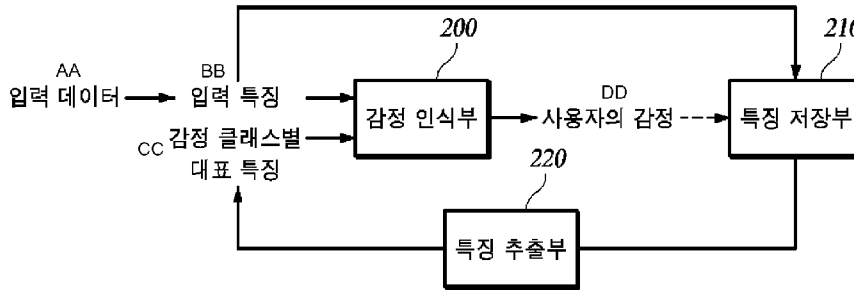
(10) 국제공개번호

WO 2024/029726 A1

- (51) 국제특허분류: *G10L 25/63* (2013.01) *G06F 40/30* (2020.01)
G10L 15/04 (2006.01)
- (21) 국제출원번호: PCT/KR2023/008107
- (22) 국제출원일: 2023년 6월 13일 (13.06.2023)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보: 10-2022-0095862 2022년 8월 2일 (02.08.2022) KR
- (71) 출원인: 에스케이텔레콤 주식회사 (SK TELECOM CO., LTD.) [KR/KR]; 04539 서울특별시 중구 을지로 65 SK T-타워, Seoul (KR).
- (72) 발명자: 정원진 (JUNG, Won Jin); 04539 서울특별시 중구 을지로 65, SKT-타워, Seoul (KR). 주민정 (JU, Min Jeong); 04539 서울특별시 중구 을지로 65, SKT-타워, Seoul (KR).
- (74) 대리인: 이철희 (LEE, Chulhee); 06229 서울특별시 강남구 도곡로33길 26, 베리타스빌딩 2-4층, Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.
- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

(54) Title: DEVICE AND METHOD FOR EMOTION RECOGNITION

(54) 발명의 명칭: 감정 인식을 위한 장치 및 방법



200 ... Emotion recognition part
210 ... Feature storage part
220 ... Feature extraction part
AA ... Input data
BB ... Input feature
CC ... Representative feature for each emotion class
DD ... Emotion of user

(57) Abstract: A device and a method for emotion recognition are disclosed. According to an aspect of the present invention, an emotion recognition method implemented by a computer is provided. The emotion recognition method includes the steps of: obtaining an input feature indicating text data, audio data, or image data of a user; extracting representative features corresponding to a plurality of emotion classes from a queue for storing at least one feature with respect to each of the plurality of emotion classes, the representative features corresponding to the plurality of emotion classes, respectively, the features stored in the queue being obtained by classifying user's previous input features into one of the plurality of emotion classes; determining an emotion of the user by using the input feature and the representative features; and storing, as a feature of an emotion class corresponding to the emotion of the user, the input feature in the queue.

[다음 쪽 계속]



WO 2024/029726 A1

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제21조(3))

(57) 요약서: 감정 인식을 위한 장치 및 방법을 개시한다. 본 발명의 일 측면에 의하면, 컴퓨터에 의해 구현되는 감정 인식 방법에 있어서, 사용자의 텍스트 데이터, 오디오 데이터 또는 이미지 데이터를 나타내는 입력 특징을 획득하는 단계; 복수의 감정 클래스들 각각에 대해 적어도 하나의 특징을 저장한 큐로부터 복수의 감정 클래스들에 대응되는 대표 특징들을 추출하는 단계 - 대표 특징들 각각은 복수의 감정 클래스들 각각에 대응되고, 큐에 저장된 특징들은 사용자의 이전 입력 특징들이 복수의 감정 클래스들 중 어느 하나로 분류된 것임; 입력 특징 및 대표 특징들을 이용하여 사용자의 감정을 판단하는 단계; 및 입력 특징을 사용자의 감정에 대응되는 감정 클래스의 특징으로서 큐에 저장하는 단계를 포함하는 감정 인식 방법을 제공한다.

명세서

발명의 명칭: 감정 인식을 위한 장치 및 방법

기술분야

- [1] 본 발명의 실시예들은 감정 인식 장치 및 방법, 자세하게는 문맥(context)에 기초하여 사용자의 감정을 판단하거나 인식하기 위한 감정 인식 장치 및 방법에 관한 것이다.

배경기술

- [2] 이하에 기술되는 내용은 단순히 본 실시예와 관련되는 배경 정보만을 제공할 뿐 종래기술을 구성하는 것이 아니다.
- [3] 최근 인공지능 기술의 발전에 따라, 인간-컴퓨터 상호작용(Human Computer Interaction; HCI) 기술분야가 급속도로 발전하고 있다.
- [4] 대표적인 HCI 시스템의 애플리케이션으로서, 사용자가 컴퓨터 장치와 텍스트 데이터 또는 오디오 데이터를 주거나 받음으로써, 대화와 같은 특정 작업을 수행할 수 있는 챗봇(chatbot)이 있다.
- [5] HCI 시스템은 사용자와의 자연스러운 대화를 위해 사용자의 감정을 이용할 수 있다. HCI 시스템은 감정 인식(Emotion Recognition; ER) 방법을 통해 사용자의 감정을 인식할 수 있다. 사용자의 감정을 인식하는 방법의 예로서, 사용자의 음성 데이터를 이용하는 음성 감정 인식(Speech Emotion Recognition; SER) 방법이 있다.
- [6] 도 1은 음성 감정 인식 방법을 설명하기 위한 도면이다.
- [7] 도 1을 참조하면, 특정 발화 데이터(utterance data, 100), 감정 인식 모델(110) 및 사용자 감정이 도시되어 있다.
- [8] 음성 감정 인식 방법은 사용자의 음성(speech)을 마이크를 통해 입력 받고, 사용자의 음성을 나타내는 음성 데이터를 복수의 시간 윈도우들(time windows)에 기초하여 복수의 발화 데이터로 나눈다. 음성 데이터는 여러 프레임들에 해당하는 복수의 발화 데이터로 구분된다. 예를 들면, 하나의 발화 데이터는 사용자가 발화한 하나의 문장에 대응될 수 있고, 하나의 시간 윈도우에 대응될 수 있다. 음성 감정 인식 방법은 복수의 발화 데이터 중 특정 발화 데이터(100)에 관련된 사용자의 감정을 판단한다. 음성 감정 인식 방법은 특정 발화 데이터(100)를 감정 인식 모델(110)에 입력하기 위한 스펙트로그램으로 변환한다. 이후, 음성 감정 인식 방법은 감정 인식 모델(110)을 이용하여 스펙트로그램으로부터 특정 발화 데이터(100)에 관한 특징(feature)들을 추출한다. 음성 감정 인식 방법은 특정 발화 데이터(100)에 관한 특징들을 감정 인식 모델(110)에 적용함으로써, 사용자 감정을 기 정의된 종류의 감정들 중 하나로 분류할 수 있다. 이때, 감정 인식 모델(110)은 합성곱 신경망(Convolution Neural Network; CNN)과 같은 신경망을 포함하는 기계 학습 모델(machine learning model)일 수 있다.

- [9] 한편, HCI 시스템은 자연어 이해(Natural Language Understanding; NLU) 또는 자연어 생성(Natural Language Generating; NLG)를 포함하는 자연어 처리(Natural Language Processing; NLP) 모델을 이용함으로써, 사용자의 감정의 정확히 인식할 수 있고 사용자와 자연스러운 대화를 수행할 수 있다. 특히, 방대한 양의 훈련 데이터로 훈련된 라지-스케일(large-scale) 모델에 의해, HCI 시스템의 성능이 개선되고 있다.
- [10] 나아가, 최근의 HCI 시스템은 하나의 입력 데이터뿐만 아니라 다양한 종류의 데이터를 함께 이용하는 멀티모달(multimodal) HCI 방식을 채용하고 있다. 멀티모달 HCI 방식은 다양한 종류의 입력들에 기초하여 사용자와 상호작용하는 것을 지칭한다. 예를 들면, 멀티모달 HCI 방식은 사용자의 감정을 인식하기 위해 사용자에게 의해 입력된 텍스트 데이터, 사용자의 발화를 나타내는 오디오 데이터 및 사용자의 얼굴표정을 나타내는 이미지 데이터를 입력으로서 이용할 수 있다. 멀티모달 HCI 시스템은 사용자에게 필요한 인공지능 서비스를 제공할 수 있다.
- [11] 하지만, 종래의 감정 인식 모델을 이용하는 HCI 시스템은 현재 대화에만 기초하여 사용자의 감정을 인식할 뿐, 과거 대화에 따른 문맥(context)을 고려하지 않는다. 즉, 종래의 HCI 시스템은 현재 프레임 내 정보에만 기초하여 사용자의 감정을 인식한다. 예를 들면, 사용자의 현재 감정을 인식하기 위한 입력으로서, 종래의 HCI 시스템은 하나의 프레임에 대응하는 하나의 문장(sentence)을 이용한다. 다른 예로서, 종래의 HCI 시스템은 단일 이미지 프레임으로부터 사용자의 감정을 인식한다.
- [12] 이처럼, 현재의 입력에만 의존하는 HCI 시스템은 사용자의 감정을 정확하게 판단하는 데 한계가 있다. 즉, 사용자의 감정에 관해 HCI 시스템에 의해 판단된 결과는 신뢰도가 낮다.

발명의 상세한 설명

기술적 과제

- [13] 본 발명의 실시예들은, 사용자의 이전 입력 데이터에 따른 문맥을 고려하여 사용자의 현재 감정을 정확하게 판단하기 위한 감정 인식 장치 및 방법을 제공하는 데 주된 목적이 있다.
- [14] 본 발명의 다른 실시예들은, 특정 사용자의 감정을 정확하게 판단함으로써 특정 사용자에게 맞춤형 서비스를 제공하기 위한 감정 인식 장치 및 방법을 제공하는 데 일 목적이 있다.

과제 해결 수단

- [15] 본 발명의 일 측면에 의하면, 컴퓨터에 의해 구현되는 감정 인식 방법에 있어서, 사용자의 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 나타내는 입력 특징을 획득하는 단계; 복수의 감정 클래스들 각각에 대해 적어도 하나의 특징(feature)을 저장한 큐(queue)로부터 상기 복수의 감정 클래스들에 대응되는 대표 특징들을 추출하는 단계 - 상기 대표 특징들 각각은 상기 복수의 감

정 클래스들 각각에 대응되고, 상기 큐에 저장된 특징들은 상기 사용자의 이전 입력 특징들이 상기 복수의 감정 클래스들 중 어느 하나로 분류된 것임 -; 상기 입력 특징 및 상기 대표 특징들을 이용하여 상기 사용자의 감정을 판단하는 단계; 및 상기 입력 특징을 상기 사용자의 감정에 대응되는 감정 클래스의 특징으로서 상기 큐에 저장하는 단계를 포함하는 감정 인식 방법을 제공한다.

- [16] 본 실시예의 다른 측면에 의하면, 감정 인식 장치에 있어서, 명령어들을 저장하는 메모리; 및 적어도 하나의 프로세서를 포함하되, 상기 적어도 하나의 프로세서는 상기 명령어들을 실행함으로써, 사용자의 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 나타내는 입력 특징을 획득하고, 복수의 감정 클래스들 각각에 대해 적어도 하나의 특징(feature)을 저장한 큐(queue)로부터 상기 복수의 감정 클래스들에 대응되는 대표 특징들을 추출하고 - 상기 대표 특징들 각각은 상기 복수의 감정 클래스들 각각에 대응되고, 상기 큐에 저장된 특징들은 상기 사용자의 이전 입력 특징들이 상기 복수의 감정 클래스들 중 어느 하나로 분류된 것임 -, 상기 입력 특징 및 상기 대표 특징들을 이용하여 상기 사용자의 감정을 판단하고, 상기 입력 특징을 상기 사용자의 감정에 대응되는 감정 클래스의 특징으로서 상기 큐에 저장하는, 감정 인식 장치를 제공한다.

발명의 효과

- [17] 이상에서 설명한 바와 같이 본 발명의 일 실시예에 의하면, 사용자의 이전 입력 데이터에 따른 문맥을 고려하여 사용자의 현재 감정을 정확하게 판단할 수 있다.
- [18] 본 발명의 다른 실시예에 의하면, 특정 사용자의 감정을 정확하게 판단함으로써 특정 사용자에게 맞춤형된 서비스를 제공할 수 있다.

도면의 간단한 설명

- [19] 도 1은 음성 감정 인식 방법을 설명하기 위한 도면이다.
- [20] 도 2는 본 발명의 일 실시예에 따른 감정 인식 방법을 설명하기 위한 도면이다.
- [21] 도 3은 본 발명의 일 실시예에 따른 특징 저장부를 설명하기 위한 도면이다.
- [22] 도 4는 본 발명의 일 실시예에 따른 특징 추출부를 설명하기 위한 도면이다.
- [23] 도 5는 본 발명의 일 실시예에 따른 감정 인식 과정을 설명하기 위한 도면이다.
- [24] 도 6은 본 발명의 일 실시예에 따른 감정 인식 방법의 순서도이다.
- [25] 도 7은 본 발명의 일 실시예에 따른 감정 인식 장치의 구성도이다.

발명의 실시를 위한 형태

- [26] 이하, 본 개시의 일부 실시예들을 예시적인 도면을 이용해 상세하게 설명한다. 각 도면의 구성 요소들에 참조 부호를 부가함에 있어서, 동일한 구성 요소들에 대해서는 비록 다른 도면 상에 표시되더라도 가능한 한 동일한 부호를 가지도록 하고 있음에 유의해야 한다. 또한, 본 개시를 설명함에 있어, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 개시의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

- [27] 본 개시에 따른 실시예의 구성요소를 설명하는 데 있어서, 제1, 제2, i), ii), a), b) 등의 부호를 사용할 수 있다. 이러한 부호는 그 구성요소를 다른 구성 요소와 구별하기 위한 것일 뿐, 그 부호에 의해 해당 구성요소의 본질 또는 차례나 순서 등이 한정되지 않는다. 명세서에서 어떤 부분이 어떤 구성요소를 '포함' 또는 '구비'한다고 할 때, 이는 명시적으로 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [28] 본 발명에 따른 장치 또는 방법의 각 구성요소는 하드웨어 또는 소프트웨어로 구현되거나, 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 각 구성요소의 기능이 소프트웨어로 구현되고 마이크로프로세서가 각 구성요소에 대응하는 소프트웨어의 기능을 실행하도록 구현될 수도 있다.
- [29] 이하에서 설명될 실시예들은 스마트폰, 개인형 컴퓨터(personal computer) 등과 같은 다양한 전자 장치들에서 사용자의 감정을 인식하는 데 이용될 수 있다. 또한, 실시예들은 챗봇, 홈 로봇, 지능형 비서 로봇, 스마트폰, 증강현실 기기, 가상현실 기기 등에서 사용자의 감정을 인식하는 데 이용될 수 있다. 실시예들은 칩(chip) 형태로 구현되어 전술한 장치들에 탑재될 수 있다.
- [30] 본 발명의 일 실시예에 따른 감정 인식 장치는 이전에 입력된 사용자의 발화들, 얼굴 표정들 또는 문장들을 기반으로 문맥(context)을 이해하고, 문맥에 기초하여 사용자의 현재 입력 데이터로부터 사용자의 감정을 정확하게 판단할 수 있다. 즉, 감정 인식 장치는 사용자의 현재 입력 데이터뿐만 아니라 과거 입력 데이터에 포함된 특징들을 이용하여 사용자의 현재 감정 상태를 판단한다.
- [31] 도 2는 본 발명의 일 실시예에 따른 감정 인식 방법을 설명하기 위한 도면이다.
- [32] 도 2를 참조하면, 본 발명의 일 실시예에 따른 감정 인식 장치는 감정 인식부(200), 특징 저장부(210) 및 특징 추출부(220)를 포함한다. 감정 인식 장치는 입력부(미도시)를 더 포함할 수 있다.
- [33] 입력부는 사용자의 입력 데이터를 수신한다.
- [34] 사용자의 입력 데이터는 하나의 프레임 단위 또는 하나의 시간 윈도우(time window) 단위를 가질 수 있다. 예를 들면, 단일 시간 윈도우에 대응되는 텍스트 데이터는 하나의 완전한 문장일 수 있고, 단일 시간 윈도우에 대응되는 오디오 데이터는 하나의 완전한 문장에 대응되는 음성일 수 있고, 단일 시간 윈도우에 대응되는 이미지 데이터는 하나의 이미지 프레임일 수 있다.
- [35] 다른 실시예에서, 사용자의 입력 데이터는 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 포함한다. 텍스트 데이터는 키보드 또는 터치 디스플레이와 같은 인터페이스를 통해 사용자에게 의해 입력되는 단어(word), 문장(sentence), 어절, 구 또는 절 중 적어도 하나를 포함한다. 오디오 데이터는 마이크를 통해 수신되는 사용자의 음성(speech)로서, 적어도 하나의 발화(utterance)를 포함한다. 이미지 데이터는 사용자의 얼굴을 포함하는 이미지 프레임이다.
- [36] 본 발명의 일 실시예에 의하면, 입력 데이터가 텍스트 데이터일 때, 입력부는 텍스트 데이터로부터 입력 특징을 추출하기 전에 텍스트 데이터가 완전한 문장을

포함하는지 여부를 판단할 수 있다. 텍스트 데이터가 불완전한 문장인 경우, 입력부는 완전한 문장을 위한 적어도 하나의 추가 텍스트 데이터를 입력 받는다. 입력부는 텍스트 데이터와 추가 텍스트 데이터를 결합한다. 결합된 데이터가 완전한 문장을 나타낼 때, 입력부는 결합된 데이터로부터 입력 특징을 추출한다. 예를 들면, 사용자가 주어, 목적어 및 서술어를 분리하여 입력하는 경우, 입력부는 주어, 목적어 및 서술어가 모두 입력될 때까지 대기하고, 주어, 목적어 및 서술어가 결합된 문장으로부터 입력 특징을 추출한다.

- [37] 입력부는 사용자의 입력 데이터로부터 입력 특징을 추출한다. 여기서, 입력 특징은 입력 데이터로부터 추출된 특징값, 특징벡터, 특징행렬, 또는 특징텐서(tensor) 중 어느 하나의 데이터 포맷을 가질 수 있다. 다른 실시예에서 입력 특징은 원시 데이터(raw data)로서, 문장 텍스트, 스펙트로그램 또는 이미지 프레임 중 어느 하나의 포맷을 가질 수 있다. 이 경우, 감정 인식부(200)에서 입력 특징을 특징벡터 등으로 변환하는 과정이 수행된다.
- [38] 입력부는 훈련된 제1 신경망을 이용하여 입력 특징을 얻을 수 있다. 예를 들면, 입력부는 복수의 레이어들을 포함하는 제1 신경망을 이용하여 입력 데이터로부터 입력 특징을 추출할 수 있다. 제1 신경망은 텍스트 입력으로부터 특징들을 추출하도록 훈련된 신경망, 오디오 입력으로부터 특징들을 추출하도록 훈련된 신경망, 또는 이미지 데이터로부터 특징들을 추출하도록 훈련된 신경망 중 어느 하나일 수 있다. 제1 신경망은 합성곱 신경망, 순환 신경망 등 다양한 아키텍처로 구현될 수 있다. 제1 신경망은 인공지능 기술분야에서 널리 알려진 방법으로 훈련된 오토인코더(auto-encoder)의 인코더 부분일 수 있다.
- [39] 한편, 특징 저장부(210)는 기 설정된 개수의 감정 클래스들마다 적어도 하나의 특징을 저장한다. 특징 저장부(210)에 저장된 초기 특징들은 모두 제로 값을 가지거나 랜덤 값을 가질 수 있다. 사용자의 감정이 판단된 후, 특징 저장부(210)는 사용자의 감정에 대응되는 감정 클래스로서 입력 특징을 저장한다. 시간이 흐름에 따라, 특징 저장부(210)는 사용자가 특정 감정을 가질 때 사용자의 행동에 관한 특징들을 축적할 수 있다.
- [40] 특징 추출부(220)는 특징 저장부(210)로부터 각 감정 클래스의 대표 특징을 추출한다. 예를 들면, 감정 인식 장치는 특징 저장부(210)로부터 중립적인 감정의 대표 특징, 화난 감정의 대표 특징, 행복한 감정의 대표 특징 및 슬픈 감정의 대표 특징을 추출할 수 있다.
- [41] 감정 인식부(200)는 입력 특징 및 감정 클래스별 대표 특징을 이용하여 사용자의 감정을 판단한다.
- [42] 본 발명의 일 실시예에 의하면, 감정 인식부(200)는 입력 특징 및 대표 특징들 간 상관관계들에 기초하여 사용자의 감정을 판단할 수 있다. 예를 들면, 감정 인식부(200)는 입력 특징과 대표 특징들 간 유사도 값들을 계산하고, 계산된 유사도 값들 중 가장 높은 유사도 값에 대응되는 대표 특징을 식별하고, 식별된 대표 특징에 대응되는 감정 클래스의 감정을 사용자의 감정으로 판단할 수 있다.

- [43] 본 발명의 다른 실시예에 의하면, 감정 인식부(200)는 감정 인식 모델을 이용하여 입력 특징 및 대표 특징들로부터 복수의 감정 클래스들에 관한 확률 분포 정보를 획득하고, 확률 분포 정보에 따라 사용자의 감정을 판단할 수 있다. 여기서, 감정 인식 모델을 훈련된 제2 신경망일 수 있다. 감정 인식 장치는 전처리부(미도시)를 더 포함할 수 있다. 구체적으로, 전처리부는 입력 특징 및 대표 특징들을 처리하여 결합 특징(concatenated feature)을 생성한다. 감정 인식부(200)는 결합 특징을 감정 인식 모델에 입력한다. 감정 인식 모델은 결합 특징의 입력에 응답하여 복수의 감정 클래스 각각에 대응되는 확률 정보를 출력한다. 감정 인식부(200)는 확률 정보에 기초하여 사용자의 감정을 판단한다.
- [44] 상기 실시예에서, 감정 인식 모델은 트랜스포머 네트워크(transformer network), 컨볼루션 뉴럴 네트워크(Convolution Neural Network; CNN), 또는 장단기 메모리(Long Short-Term Memory; LSTM) 네트워크 중 적어도 하나를 포함할 수 있다.
- [45] 상기 실시예에서, 감정 인식 모델은 훈련용 결합 특징들을 포함하는 훈련 데이터의 입력에 응답하여 복수의 감정 클래스들 각각에 대응되는 확률 정보를 출력하도록 훈련된 모델이다.
- [46] 감정 인식 장치는 사용자가 특정 감정을 가질 때 수신되는 발화 특징들, 행동 특징들, 또는 입력 텍스트의 특징들을 저장하고, 저장된 특징들에 기초하여 사용자의 현재 입력 특징으로부터 사용자의 감정을 판단할 수 있다. 예를 들면, 감정 인식 장치는 사용자가 화났을 때 주로 사용하는 단어들의 특징들에 기초하여, 사용자의 현재 발화로부터 사용자가 어떤 감정을 가지는지 판단할 수 있다.
- [47] 감정 인식 장치는 사용자와의 과거 대화의 문맥을 고려하여 사용자의 감정을 판단하므로, 현재 대화 내용에만 의존하여 사용자의 감정을 판단하는 것에 비해 높은 감정 인식 성능과 높은 신뢰도를 갖는다.
- [48] 도 3은 본 발명의 일 실시예에 따른 특징 저장부를 설명하기 위한 도면이다.
- [49] 도 3에서, 감정 클래스들은 4개인 것으로 설명한다. 제1 감정 클래스는 중립적인 감정을 나타내고, 제2 감정 클래스는 화난 감정을 나타내고, 제3 감정 클래스는 행복한 감정을 나타내고, 제4 감정 클래스는 슬픈 감정을 나타낸다. 다른 실시예에서, 감정 클래스들은 위 네 가지 감정뿐만 아니라 공포 감정, 놀란 감정, 편안한 감정, 불안한 감정 등 다양한 감정에 관한 감정 클래스들을 더 포함할 수 있다.
- [50] 도 3을 참조하면, 특징 저장부(210)는 큐(queue, 212)를 포함할 수 있다.
- [51] 큐(212)는 초기 특징들이 저장되고, 추가적으로 사용자의 입력 데이터로부터 추출된 입력 특징이 저장되는 적어도 하나의 메모리이다.
- [52] 큐(212)는 기 설정된 개수의 감정 클래스들마다 적어도 하나의 특징을 저장한다. 큐(212)는 감정 클래스들의 개수에 해당하는 부분 큐들을 포함할 수 있고, 각 부분 큐는 각 감정 클래스의 특징들을 저장한다.
- [53] 큐(212)는 초기 특징들 i_1, i_2, i_3 , 및 i_4 를 초기값으로서 저장한다. 예를 들어, 큐(212)는 제1 감정 클래스의 특징으로서 제1 초기 특징 i_1 를 저장하고, 제2 감정 클래스의 특징으로서 제2 초기 특징 i_2 를 저장한다.

- [54] 하나의 입력 특징이 수신되고 입력 특징에 따른 사용자의 감정이 판단될 때, 큐(212)는 사용자의 판단된 감정에 대응되는 감정 클래스의 특징 데이터로서, 입력 특징을 저장한다. 예를 들어, $t=0$ 에서 입력 특징 $x(0)$ 로부터 사용자의 감정이 중립적인 감정으로 판단될 때, 큐(212)는 입력 특징 $x(0)$ 를 제1 감정 클래스의 특징으로 저장한다. $t=1$ 에서 입력 특징 $x(1)$ 로부터 사용자의 감정이 중립적인 감정으로 판단될 때, 큐(212)는 입력 특징 $x(1)$ 을 제1 감정 클래스의 특징으로 저장한다. 반면, $t=3$ 에서 입력 특징 $x(3)$ 로부터 사용자의 감정이 화난 감정으로 판단되므로, 큐(212)는 입력 특징 $x(3)$ 를 제2 감정 클래스의 특징으로 저장한다.
- [55] 제1 감정 클래스의 특징들, 제2 감정 클래스의 특징들, 제3 감정 클래스의 특징들 및 제4 감정 클래스의 특징들 각각은 제1 특징들, 제2 특징들, 제3 특징들 및 제4 특징들로 지칭될 수 있다. 예를 들면, 제1 특징들은 i_1 , $x(0)$, $x(1)$, 및 $x(2)$ 등을 포함한다. 제2 특징들은 i_2 , $x(3)$, $x(4)$, $x(9)$, 및 $x(10)$ 등을 포함한다.
- [56] 큐(212)는 특정 감정 상태의 사용자가 사용하는 언어 특징, 행동 특징, 표정 특징 등을 저장한다. 예를 들면, 큐(212)는 특정 감정 상태의 사용자가 사용하는 문장 형식, 문장 내 단어, 단어의 순서, 언어 습관 등을 저장할 수 있다. 이처럼, 큐(212)는 사용자의 감정과 동작 패턴을 대응시켜 저장한다.
- [57] 사용자로부터 입력되는 입력 특징들이 많을수록, 큐(212)에 저장되는 사용자의 감정별 행동 패턴도 많아진다.
- [58] 큐(212)에 저장된 감정 클래스별 특징들은 문맥 정보가 된다. 특히, 상대적으로 최근에 저장된 특징들은 사용자와의 현재 대화에서 문맥을 나타낸다. 문맥 정보는 사용자의 현재 입력 특징으로부터 사용자가 어떤 감정을 가지는지 판단하는데 유용한 정보를 제공한다.
- [59] 나아가, 감정 인식 장치가 사용자의 단말에 대응하고, 큐(212)가 특정 사용자에게 해당하는 특징들을 저장하는 경우, 저장된 특징들은 특정 사용자에게 관한 정보이므로 특정 사용자에게 대한 맞춤형(customizing) 서비스를 제공하는 데 이용될 수 있다.
- [60] 도 4는 본 발명의 일 실시예에 따른 특징 추출부를 설명하기 위한 도면이다.
- [61] 도 4를 참조하면, 특징 추출부(230)가 도시되어 있다.
- [62] 특징 추출부(230)는 복수의 감정 클래스들 각각에 대해 적어도 하나의 특징이 저장된 큐로부터 복수의 감정 클래스들에 대응되는 대표 특징들을 추출한다.
- [63] 도 3 및 도 4를 참조하면, 특징 추출부(230)는 제1 특징들로부터 제1 대표 특징 $f_{Neutral}$ 을 추출하고, 제2 특징들로부터 제2 대표 특징 f_{Angry} 를 추출하고, 제3 특징들로부터 제3 대표 특징 f_{Happy} 를 추출하고, 제4 특징들로부터 제4 대표 특징 f_{Sad} 를 추출한다. 이처럼, 특징 추출부(230)는 감정 클래스별 대표 특징을 추출한다.
- [64] 이하에서, 각 감정 클래스마다 초기 특징만 저장된 상태에서 대표 특징이 추출될 수 있지만, 각 감정 클래스마다 복수의 특징들이 저장된 것으로 가정하여 설명한다.

- [65] 특징 추출부(230)는 각 감정 클래스마다 저장된 특징들로부터 대표 특징을 추출하는 방식으로서, 생성 방식 또는 선택 방식 중 하나를 이용할 수 있다.
- [66] 본 발명의 일 실시예에 의하면, 특징 추출부(230)는 각 감정 클래스마다 저장된 특징들을 인코딩함으로써, 대표 특징들을 추출할 수 있다. 예를 들면, 특징 추출부(230)는 제1 특징들을 인코딩함으로써 제1 대표 특징을 생성할 수 있다.
- [67] 특징 추출부(230)는 표현 학습(representation learning)된 모델을 이용할 수 있다. 표현 학습은 문맥을 임베딩하는 것을 지칭한다. 일 예로서, 특징 추출부(230)는 훈련이 완료된 오토인코더의 인코더를 이용할 수 있다. 여기서, 오토인코더는 입력으로부터 특징들을 추출하는 인코더와, 추출된 특징들로부터 상기 입력에 대응되는 출력 데이터를 생성하는 디코더로 구성된 모델이다. 오토인코더는 입력과 동일한 출력을 생성하도록 자기 지도 학습 방법으로 훈련될 수 있다. 훈련이 완료된 오토인코더의 인코더는 입력으로부터 잠재 공간(latent space)을 특징으로서 추출한다. 이처럼, 특징 추출부(230)는 각 감정 클래스마다 저장된 특징들로부터 하나의 대표 특징을 추출할 수 있다.
- [68] 한편, 본 발명의 다른 실시예에 의하면, 특징 추출부(230)는 각 감정 클래스마다 복수의 특징들이 저장된 경우, 각 감정 클래스마다 저장된 복수의 특징들 중 하나를 선택함으로써, 대표 특징들을 추출할 수 있다.
- [69] 특징 추출부(230)는 원형 학습(prototypical learning)된 모델을 이용할 수 있다. 특징 추출부(230)가 원형 학습된 모델을 이용할 때, 특징 추출부(230)는 각 감정 클래스마다 원형 특징을 생성하고, 각 원형 특징과 가장 유사한 특징을 각 감정 클래스의 대표 특징으로 선택한다. 예를 들면, 제1 감정 클래스의 제1 특징들의 평균을 제1 원형 특징으로 생성한다. 제1 특징들 중 제1 원형 특징과 가장 유사한 제1 특징을 대표 특징으로 선택한다. 이후에 입력 특징이 제1 감정 클래스로서 저장되면, 제1 특징들 및 입력 특징의 평균에 따라 제1 원형 특징을 갱신한다.
- [70] 특징 추출부(230)는 각 감정 클래스의 특징들 중 어느 하나의 특징을 임의로 선택함으로써 대표 특징들을 추출할 수 있다. 예를 들면, 특징 추출부(230)는 제1 특징들 중에서 임의로 선택된 하나를 제1 대표 특징으로 결정할 수 있다.
- [71] 다른 실시예에서는, 전술한 대표 특징 추출 방법 외에, 각 감정 클래스별 대표 특징을 추출하는 다양한 방법이 이용될 수 있다.
- [72] 추출된 대표 특징은 입력 특징의 포맷과 동일한 포맷을 가지는 것이 바람직하다. 예를 들면, 입력 특징의 포맷이 벡터일 때, 대표 특징의 포맷도 벡터일 수 있다.
- [73] 도 5는 본 발명의 일 실시예에 따른 감정 인식 과정을 설명하기 위한 도면이다.
- [74] 도 5를 참조하면, 감정 인식 장치는 감정 인식부(200) 및 전처리부(500)를 포함한다.
- [75] 전처리부(500)는 복수의 감정 클래스들에 대응하는 복수의 대표 특징들과 입력 특징을 처리한다.

- [76] 본 발명의 일 실시예에 의하면, 감정 인식부(200)가 딥러닝 기술을 이용할 수 있도록, 전처리부(500)는 복수의 대표 특징들과 입력 특징을 결합(concatenate)함으로써 결합 특징을 생성한다. 대표 특징들과 입력 특징이 벡터 포맷을 가질 때, 전처리부(500)는 대표 특징 벡터들을 서로 결합하고, 연결된 대표 특징 벡터들에 입력 특징 벡터를 결합한다.
- [77] 도 5에서 입력 특징은 대표 특징들의 오른쪽에 결합되는 것으로 도시되어 있으나, 다른 실시예에서 입력 특징과 대표 특징들이 결합되는 순서는 다양하게 설정될 수 있다.
- [78] 입력 특징과 대표 특징들이 결합된 경우, 감정 인식부(200)는 사용자의 감정을 판단하기 위해 결합 특징을 감정 인식 모델의 입력으로 이용할 수 있다.
- [79] 여기서, 감정 인식 모델은 지도 학습을 통해 미리 훈련될 수 있다. 일 예로서, 감정 인식 모델은 훈련용 결합 특징들을 포함하는 훈련 데이터의 입력에 응답하여 복수의 감정 클래스들 각각에 대응되는 확률 정보를 출력하도록 훈련된다. 훈련 데이터에 포함된 결합 특징들 각각은 정답 확률 정보로 레이블된 것이다.
- [80] 감정 인식부(200)는 결합 특징을 훈련이 완료된 감정 인식 모델에 입력한다. 감정 인식 모델은 결합 특징의 입력에 따라 각 감정 클래스별 확률 값들을 포함하는 확률 정보를 출력한다. 감정 인식부(200)는 확률 정보에 기초하여 사용자의 감정을 판단한다. 일 예로서, 감정 인식부(200)는 가장 큰 확률값을 가지는 감정 클래스를 사용자의 감정으로 판단할 수 있다.
- [81] 감정 인식 모델은 적어도 하나의 레이어를 포함하는 신경망 구조를 가질 수 있다. 감정 인식 모델은 다양한 아키텍처를 가질 수 있다.
- [82] 본 발명의 일 실시예에 의하면, 감정 인식 모델은 어텐션 메커니즘(attention mechanism)에 기초하여 훈련된 것일 수 있다. 구체적으로, 감정 인식 모델은 적어도 하나의 인코더와 적어도 하나의 디코더를 포함한다. 적어도 하나의 인코더는 결합 특징을 입력 받는다. 도 5에서 결합 특징은 5개의 특징들로 나뉘어 순차적으로 입력된다. 적어도 하나의 인코더에는 결합 특징의 입력에 따른 은닉 상태들(hidden states)이 저장된다. 인코더의 각 은닉 상태는 감정 클래스들 각각에 대응될 수 있다. 적어도 하나의 디코더에는 인코더의 마지막 은닉 상태가 입력되고, 디코더의 은닉 상태가 출력된다. 디코더에는 결합 특징에 대한 레이블이 함께 입력될 수 있다. 적어도 하나의 디코더에는 입력마다 은닉 상태가 저장된다. 마지막 디코딩 단계에서, 디코더에 입력 특징과 직전 은닉 상태가 입력되고, 디코더는 최종 은닉 상태를 출력한다. 디코더의 최종 은닉 상태가 인코더에 저장된 각 은닉 상태들과 연산되는 것에 기초하여 어텐션 스코어들(attention scores)이 도출된다. 어텐션 스코어들은 인코더의 은닉 상태들에 대응된다. 각 어텐션 스코어는 대응되는 인코더의 은닉 상태의 가중치로서 적용된다. 즉, 어텐션 스코어들은 인코더의 은닉 상태들과 가중합된다. 가중합 결과는 하나의 벡터이며, 문맥 벡터(context vector)가 된다. 감정 인식 모델은 문맥 벡터와 디코더의 최종 은닉 상태 간 결합에 가중치들을 적용하고, 소프트맥스 함수를 통해 각 감정 클래스별 확률

정보를 출력한다. 정리하면, 감정 인식 모델은 감정 클래스별 대표 특징과 입력 특징 간 관계에 기초하여 사용자의 감정 클래스들에 관한 확률 정보를 출력하도록 훈련된다. 이때, 감정 인식 모델은 입력 특징과 연관이 있거나 유사한 값을 가지는 대표 특징을 식별하고, 해당 대표 특징에 대응하는 감정 클래스의 확률값을 높게 출력하도록 훈련될 수 있다.

- [83] 본 발명의 다른 실시예에 의하면, 감정 인식 모델은 어텐션 메커니즘을 이용하는 트랜스포머 네트워크일 수 있다.
- [84] 본 발명의 다른 실시예에 의하면, 감정 인식 모델은 합성곱 신경망일 수 있다. 감정 인식 모델은 적어도 하나의 합성곱 레이어를 포함한다.
- [85] 본 발명의 다른 실시예에 의하면, 감정 인식 모델은 장단기 메모리 네트워크일 수 있다. 감정 인식 모델은 순환 신경망 구조를 가질 수 있고, 적어도 하나의 레이어를 포함할 수 있다.
- [86] 한편, 본 발명의 다른 실시예에 의하면, 감정 인식부(200)는 입력 특징 및 대표 특징들을 결합 없이 사용자의 감정을 판단할 수 있다. 감정 인식부(200)는 입력 특징 및 대표 특징들 간 상관관계들에 기초하여 사용자의 감정을 판단할 수 있다. 여기서, 입력 특징과 대표 특징 간 상관관계는 유사도(similarity)를 지칭한다. 일 예로서, 감정 인식부(200)는 대표 특징들 중 입력 특징과 유사도가 높은 하나의 대표 특징을 검출하고, 검출된 대표 특징에 대응하는 감정 클래스를 사용자의 감정으로 판단한다.
- [87] 도 6은 본 발명의 일 실시예에 따른 감정 인식 방법의 순서도이다.
- [88] 도 6을 참조하면, 감정 인식 장치는 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 포함하는 입력 데이터를 수신한다.
- [89] 감정 인식 장치는 사용자의 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 나타내는 입력 특징을 획득한다(S500).
- [90] 감정 인식 장치는 특징 추출기를 이용하여 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나로부터 입력 특징을 추출할 수 있다.
- [91] 본 발명의 일 실시예에 의하면, 감정 인식 장치는 단일 문장에 대응되는 입력 특징만을 추출하기 위해, 텍스트 데이터를 확인할 수 있다. 구체적으로, 감정 인식 장치는 텍스트 데이터가 불완전한 문장인지 여부를 판단한다. 여기서, 완전한 문장의 판단기준은 주어, 목적어, 서술어 등을 기반으로 미리 설정될 수 있다. 감정 인식 장치는 텍스트 데이터가 불완전한 문장인 것으로 판단한 경우, 적어도 하나의 추가 텍스트 데이터를 입력 받는다. 텍스트 데이터와 추가 텍스트 데이터의 결합이 하나의 문장을 나타내는 경우, 감정 인식 장치는 상기 결합으로부터 입력 특징을 추출한다.
- [92] 감정 인식 장치는 복수의 감정 클래스들 각각에 대해 적어도 하나의 특징을 저장한 큐로부터 복수의 감정 클래스들에 대응되는 대표 특징들을 추출한다(S510).

- [93] 여기서, 대표 특징들 각각은 복수의 감정 클래스들 각각에 대응된다. 큐에 저장된 특징들은 사용자의 이전 입력 특징들이 복수의 감정 클래스들 중 어느 하나로 미리 분류된 것이다. 즉, 큐에 저장된 특징들은 사용자와의 과거 대화를 통해 분석한 문맥 정보이다.
- [94] 본 발명의 일 실시예에 의하면, 감정 인식 장치는 각 감정 클래스마다 적어도 하나의 특징을 인코딩함으로써, 대표 특징들을 추출한다.
- [95] 본 발명의 일 실시예에 의하면, 감정 인식 장치는 각 감정 클래스마다 복수의 특징들이 저장된 경우, 각 감정 클래스마다 저장된 복수의 특징들 중 하나를 선택함으로써, 대표 특징들을 추출한다.
- [96] 감정 인식 장치는 입력 특징 및 대표 특징들을 이용하여 사용자의 감정을 판단한다(S520).
- [97] 본 발명의 일 실시예에 의하면, 감정 인식 장치는 입력 특징 및 대표 특징들 간 상관관계들에 기초하여 사용자의 감정을 판단할 수 있다.
- [98] 본 발명의 일 실시예에 의하면, 감정 인식 장치는 훈련된 모델을 이용하여 사용자의 감정을 판단할 수 있다. 구체적으로, 감정 인식 장치는 입력 특징 및 대표 특징들을 처리하여 결합 특징을 생성한다. 감정 인식 장치는 감정 인식 모델을 이용하여 결합 특징으로부터 감정 클래스들에 관한 확률 정보를 획득한다. 여기서, 감정 인식 모델은 훈련용 결합 특징들을 포함하는 훈련 데이터의 입력에 응답하여 복수의 감정 클래스들 각각에 대응되는 확률 정보를 출력하도록 훈련된 것이다. 감정 인식 장치는 상기 확률 정보에 기초하여 사용자의 감정을 판단한다.
- [99] 본 발명의 일 실시예에 의하면, 감정 인식 모델은 입력 특징과 대표 특징들 간 어텐션에 기초하여 확률 정보를 출력하도록 훈련된 것일 수 있다.
- [100] 사용자의 감정 판단 후, 감정 인식 장치는 입력 특징을 사용자의 감정에 대응되는 감정 클래스의 특징으로서 큐에 저장한다(S522).
- [101] 사용자의 감정 판단에 이용된 입력 특징은 문맥 정보로서 저장된다. 감정 인식 장치가 다음 입력 특징에 따른 사용자의 감정을 판단할 때, 큐에 저장된 입력 특징은 문맥 정보로서 이용된다.
- [102] 도 7은 본 발명의 일 실시예에 따른 감정 인식 장치의 구성도이다.
- [103] 도 7을 참조하면, 감정 인식 장치(700)는 메모리(710), 프로세서(720), 스토리지(730), 입출력 인터페이스(740) 및 통신 인터페이스(750) 중 일부 또는 전부를 포함할 수 있다. 감정 인식 장치(700)는 훈련 데이터 생성 장치를 나타낸다.
- [104] 감정 인식 장치(700)는 데스크탑 컴퓨터, 서버, AI 가속기 등과 같은 고정형(stationary) 컴퓨팅 장치뿐만 아니라, 랩탑 컴퓨터, 스마트폰 등과 같은 휴대용(mobile) 컴퓨팅 장치일 수도 있다.
- [105] 메모리(710)는 프로세서(720)로 하여금 본 발명의 일 실시예에 따른 데이터 생성 방법을 수행하도록 하는 프로그램을 저장할 수 있다. 예를 들면, 프로그램은 프로세서(720)에 의해서 실행 가능한(executable) 복수의 명령어들을 포함할 수

- 있고, 복수의 명령어들이 프로세서(720)에 의해서 실행됨으로써 데이터 생성 방법이 수행될 수 있다.
- [106] 메모리(710)는 단일 메모리 또는 복수의 메모리들일 수 있다. 이 경우, 데이터 생성에 필요한 정보는 단일 메모리에 저장되거나 복수의 메모리들에 나뉘어 저장될 수 있다. 메모리(710)가 복수의 메모리들로 구성된 경우, 복수의 메모리들은 물리적으로 분리될 수 있다.
- [107] 메모리(710)는 휘발성 메모리 및 비휘발성 메모리 중 적어도 하나를 포함할 수 있다. 휘발성 메모리는 SRAM(Static Random Access Memory) 또는 DRAM(Dynamic Random Access Memory) 등을 포함하고, 비휘발성 메모리는 플래시 메모리(flash memory) 등을 포함한다.
- [108] 프로세서(720)는 적어도 하나의 명령어들을 실행할 수 있는 적어도 하나의 코어를 포함할 수 있다. 프로세서(720)는 메모리(710)에 저장된 명령어들을 실행할 수 있다. 프로세서(720)는 단일 프로세서 또는 복수의 프로세서들일 수 있다.
- [109] 스토리지(730)는 감정 인식 장치(700)에 공급되는 전력이 차단되더라도 저장된 데이터를 유지한다. 예를 들면, 스토리지(730)는 비휘발성 메모리를 포함할 수도 있고, 자기 테이프, 광학 디스크, 자기 디스크와 같은 저장 매체를 포함할 수도 있다.
- [110] 본 발명의 일 실시예에 의하면, 스토리지(730)는 훈련 데이터를 구성하는 프로그램을 저장할 수 있다. 스토리지(730)에 저장된 프로그램은 프로세서(720)에 의해서 실행되기 이전에 메모리(710)로 로딩될 수 있다. 스토리지(730)는 프로그램 언어로 작성된 파일을 저장할 수 있고, 파일로부터 컴파일러 등에 의해서 생성된 프로그램은 메모리(710)로 로딩될 수 있다.
- [111] 스토리지(730)는 프로세서(720)에 의해서 처리될 데이터 및 프로세서(720)에 의해서 처리된 데이터를 저장할 수 있다.
- [112] 입출력 인터페이스(740)는 키보드, 마우스 등과 같은 입력 장치를 포함할 수 있고, 디스플레이 장치, 프린터 등과 같은 출력 장치를 포함할 수 있다. 사용자는 입출력 인터페이스(740)를 통해 프로세서(720)에 의한 프로그램의 실행을 트리거할 수도 있다.
- [113] 통신 인터페이스(750)는 외부 네트워크에 대한 액세스를 제공한다. 예를 들면, 감정 인식 장치(700)는 통신 인터페이스(750)를 통해 다른 장치들과 통신할 수 있다.
- [114] 본 명세서에 설명되는 시스템들 및 기법들의 다양한 구현예들은, 디지털 전자 회로, 집적회로, FPGA(field programmable gate array), ASIC(application specific integrated circuit), 컴퓨터 하드웨어, 펌웨어, 소프트웨어, 및/또는 이들의 조합으로 실현될 수 있다. 이러한 다양한 구현예들은 프로그래밍가능 시스템 상에서 실행 가능한 하나 이상의 컴퓨터 프로그램들로 구현되는 것을 포함할 수 있다. 프로그래밍가능 시스템은, 저장 시스템, 적어도 하나의 입력 디바이스, 그리고 적어도 하나의 출력 디바이스로부터 데이터 및 명령들을 수신하고 이들에게 데이

터 및 명령들을 전송하도록 결합되는 적어도 하나의 프로그래밍가능 프로세서 (이것은 특수 목적 프로세서일 수 있거나 혹은 범용 프로세서일 수 있음)를 포함한다. 컴퓨터 프로그램들(이것은 또한 프로그램들, 소프트웨어, 소프트웨어 애플리케이션들 혹은 코드로서 알려져 있음)은 프로그래밍가능 프로세서에 대한 명령어들을 포함하며 "컴퓨터가 읽을 수 있는 기록매체"에 저장된다.

- [115] 컴퓨터가 읽을 수 있는 기록매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 이러한 컴퓨터가 읽을 수 있는 기록매체는 ROM, CD-ROM, 자기 테이프, 플로피디스크, 메모리 카드, 하드 디스크, 광자기 디스크, 스토리지 디바이스 등의 비휘발성(non-volatile) 또는 비일시적인(non-transitory) 매체일 수 있으며, 또한 데이터 전송 매체(data transmission medium)와 같은 일시적인(transitory) 매체를 더 포함할 수도 있다. 또한, 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다.
- [116] 본 명세서의 흐름도/타이밍도에서는 각 과정들을 순차적으로 실행하는 것으로 기재하고 있으나, 이는 본 개시의 일 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것이다. 다시 말해, 본 개시의 일 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 개시의 일 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 흐름도/타이밍도에 기재된 순서를 변경하여 실행하거나 각 과정들 중 하나 이상의 과정을 병렬적으로 실행하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이므로, 흐름도/타이밍도는 시계열적인 순서로 한정되는 것은 아니다.
- [117] 이상의 설명은 본 실시예의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 다양한 수정 및 변형이 가능할 것이다. 따라서, 본 실시예들은 본 실시예의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.
- [118] CROSS-REFERENCE TO RELATED APPLICATION
- [119] 본 특허출원은, 본 명세서에 그 전체가 참고로서 포함되는, 2022년 08월 02일에 한국에 출원한 특허출원번호 제10-2022-0095862호에 대해 우선권을 주장한다.

청구범위

- [청구항 1] 컴퓨터에 의해 구현되는 감정 인식 방법에 있어서,
 사용자의 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 나타내는 입력 특징을 획득하는 단계;
 복수의 감정 클래스들 각각에 대해 적어도 하나의 특징(feature)을 저장한 큐(queue)로부터 상기 복수의 감정 클래스들에 대응되는 대표 특징들을 추출하는 단계 - 상기 대표 특징들 각각은 상기 복수의 감정 클래스들 각각에 대응되고, 상기 큐에 저장된 특징들은 상기 사용자의 이전 입력 특징들이 상기 복수의 감정 클래스들 중 어느 하나로 분류된 것임 -;
 상기 입력 특징 및 상기 대표 특징들을 이용하여 상기 사용자의 감정을 판단하는 단계; 및
 상기 입력 특징을 상기 사용자의 감정에 대응되는 감정 클래스의 특징으로서 상기 큐에 저장하는 단계를 포함하는 감정 인식 방법.
- [청구항 2] 제1항에 있어서,
 상기 대표 특징들을 추출하는 단계는,
 각 감정 클래스마다 적어도 하나의 특징을 인코딩함으로써, 상기 대표 특징들을 추출하는 단계를 포함하는 감정 인식 방법.
- [청구항 3] 제1항에 있어서,
 상기 대표 특징들을 추출하는 단계는,
 각 감정 클래스마다 복수의 특징들이 저장되어 있는 경우, 각 감정 클래스마다 저장된 상기 복수의 특징들 중 하나를 선택함으로써, 상기 대표 특징들을 추출하는 단계를 포함하는 감정 인식 방법.
- [청구항 4] 제1항에 있어서,
 상기 사용자의 감정을 판단하는 단계는,
 상기 입력 특징 및 상기 대표 특징들 간 상관관계들에 기초하여 상기 사용자의 감정을 판단하는 단계를 포함하는 감정 인식 방법.
- [청구항 5] 제1항에 있어서,
 상기 사용자의 감정을 판단하는 단계는,
 상기 입력 특징 및 상기 대표 특징들을 처리하여 결합 특징(concatenated feature)을 생성하는 단계;
 감정 인식 모델을 이용하여 상기 결합 특징으로부터 상기 감정 클래스들에 관한 확률 정보를 획득하는 단계 - 상기 감정 인식 모델은 훈련용 결합

특징들을 포함하는 훈련 데이터의 입력에 응답하여 상기 복수의 감정 클래스들 각각에 대응되는 확률 정보를 출력하도록 훈련된 것임 -; 및 상기 확률 정보에 기초하여 상기 사용자의 감정을 판단하는 단계를 포함하는 감정 인식 방법.

[청구항 6]

제5항에 있어서,
상기 감정 인식 모델은,
훈련용 결합 특징들에 어텐션 메커니즘(attention mechanism)을 적용하는 것에 기초하여 훈련된 것인 감정 인식 방법.

[청구항 7]

제1항에 있어서,
상기 입력 특징을 획득하는 단계는,
하나의 문장을 위한 적어도 하나의 추가 텍스트 데이터를 입력받는 단계;
및
상기 텍스트 데이터와 상기 추가 텍스트 데이터의 결합으로부터 입력 특징을 추출하는 단계를 포함하는 감정 인식 방법.

[청구항 8]

제1항에 있어서,
상기 텍스트 데이터는 하나의 문장을 나타내고, 상기 오디오 데이터는 하나의 발화를 나타내고, 상기 이미지 데이터는 하나의 이미지 프레임을 나타내는 것인 감정 인식 방법.

[청구항 9]

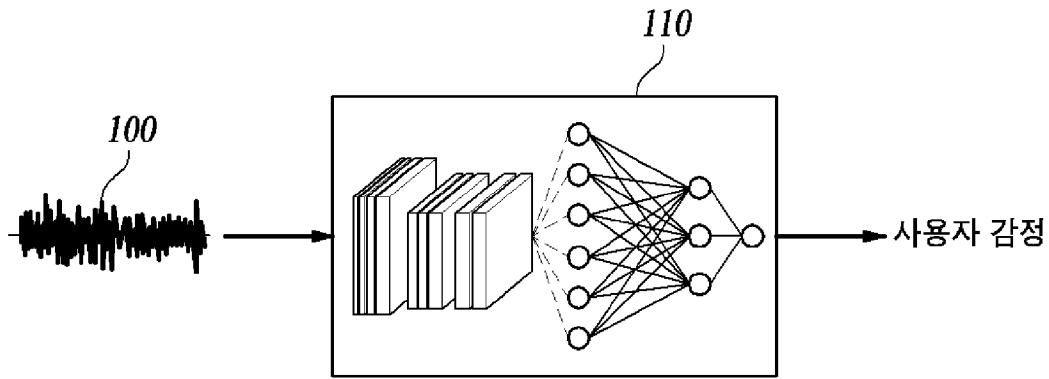
감정 인식 장치에 있어서,
명령어들을 저장하는 메모리; 및
적어도 하나의 프로세서를 포함하되,
상기 적어도 하나의 프로세서는 상기 명령어들을 실행함으로써,
사용자의 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 나타내는 입력 특징을 획득하고,
복수의 감정 클래스들 각각에 대해 적어도 하나의 특징(feature)을 저장한 큐(queue)로부터 상기 복수의 감정 클래스들에 대응되는 대표 특징들을 추출하고 - 상기 대표 특징들 각각은 상기 복수의 감정 클래스들 각각에 대응되고, 상기 큐에 저장된 특징들은 상기 사용자의 이전 입력 특징들이 상기 복수의 감정 클래스들 중 어느 하나로 분류된 것임 -,
상기 입력 특징 및 상기 대표 특징들을 이용하여 상기 사용자의 감정을 판단하고,
상기 입력 특징을 상기 사용자의 감정에 대응되는 감정 클래스의 특징으로서 상기 큐에 저장하는, 감정 인식 장치.

[청구항 10]

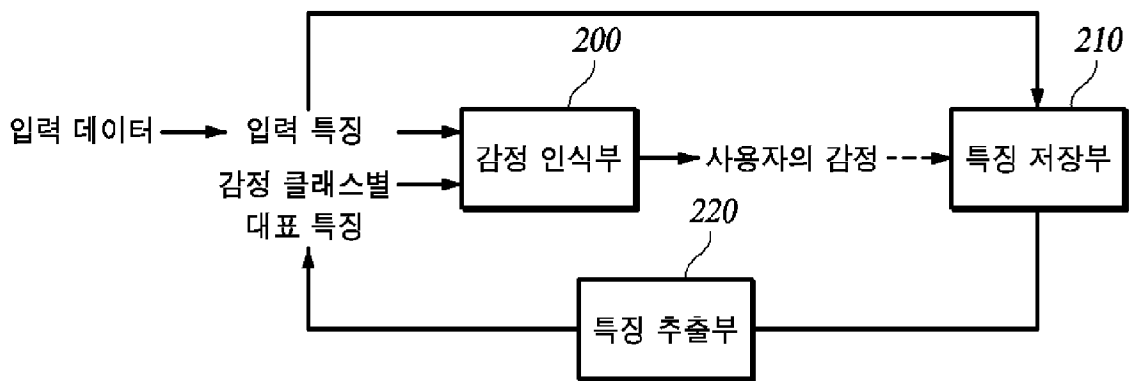
명령어가 저장된, 컴퓨터로 읽을 수 있는 기록매체로서, 상기 명령어는 상기 컴퓨터에 의해 실행될 때 상기 컴퓨터로 하여금,
사용자의 텍스트 데이터, 오디오 데이터 또는 이미지 데이터 중 어느 하나를 나타내는 입력 특징을 획득하는 단계;

복수의 감정 클래스들 각각에 대해 적어도 하나의 특징(feature)을 저장한 큐(queue)로부터 상기 복수의 감정 클래스들에 대응되는 대표 특징들을 추출하는 단계 - 상기 대표 특징들 각각은 상기 복수의 감정 클래스들 각각에 대응되고, 상기 큐에 저장된 특징들은 상기 사용자의 이전 입력 특징들이 상기 복수의 감정 클래스들 중 어느 하나로 분류된 것임 -; 및
상기 입력 특징 및 상기 대표 특징들을 이용하여 상기 사용자의 감정을 판단하는 단계
상기 입력 특징을 상기 사용자의 감정에 대응되는 감정 클래스의 특징으로서 상기 큐에 저장하는 단계
를 실행하는, 컴퓨터로 읽을 수 있는 기록매체.

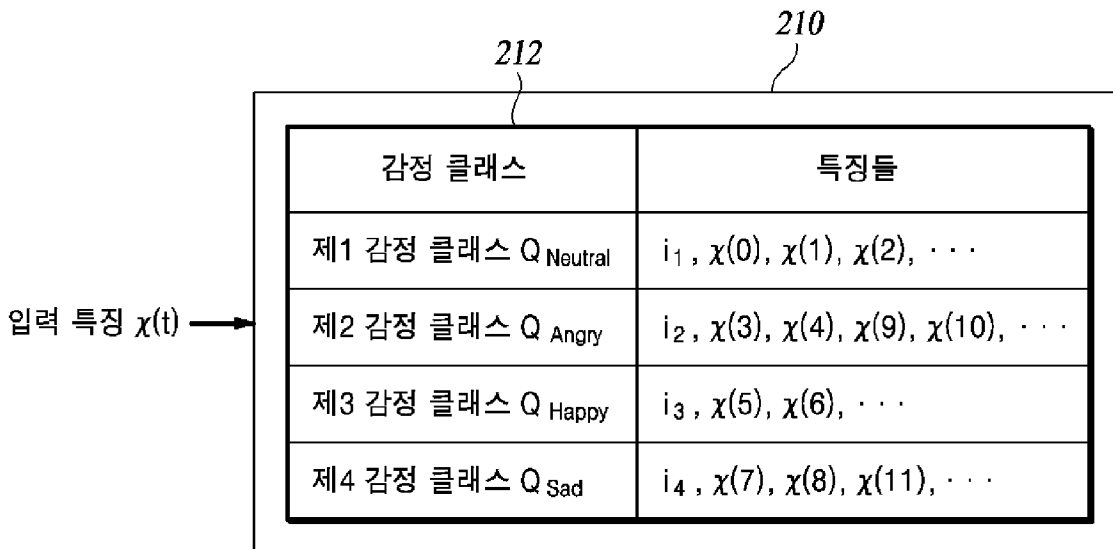
[도1]



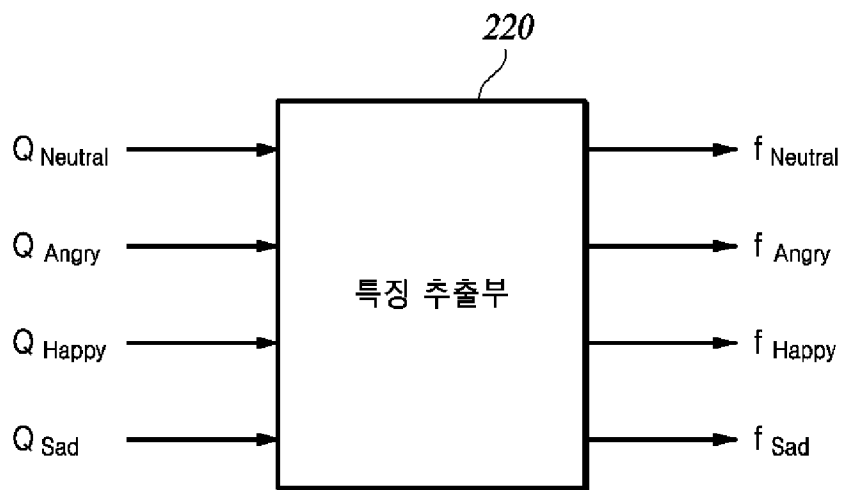
[도2]



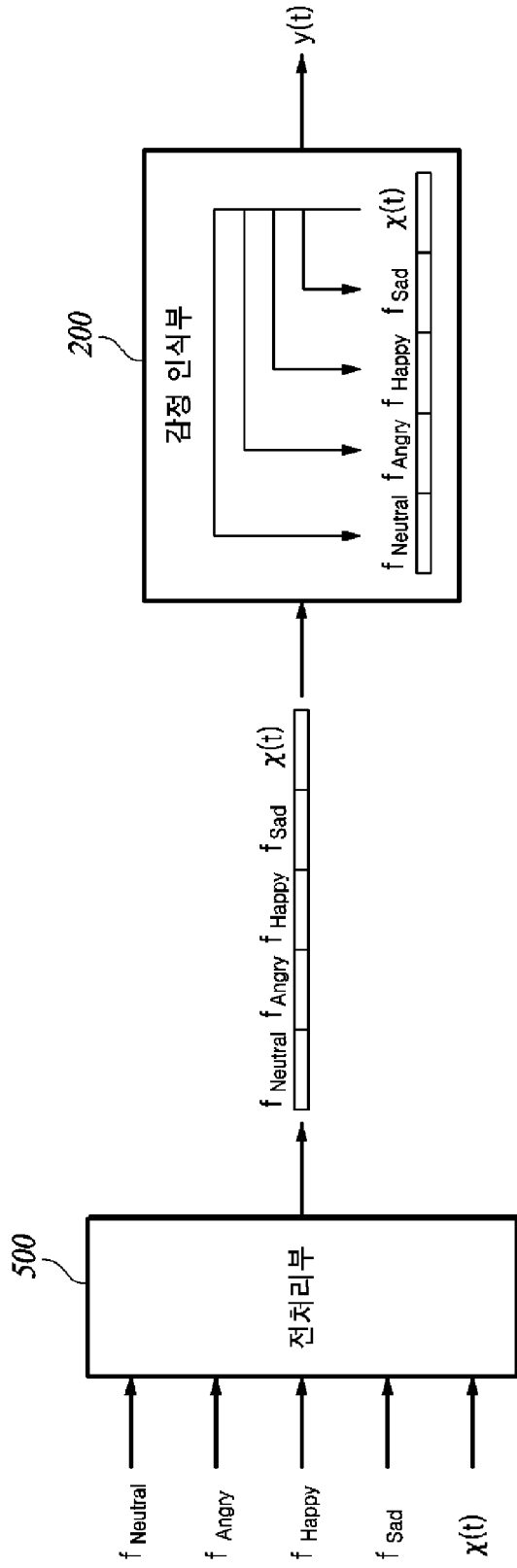
[도3]



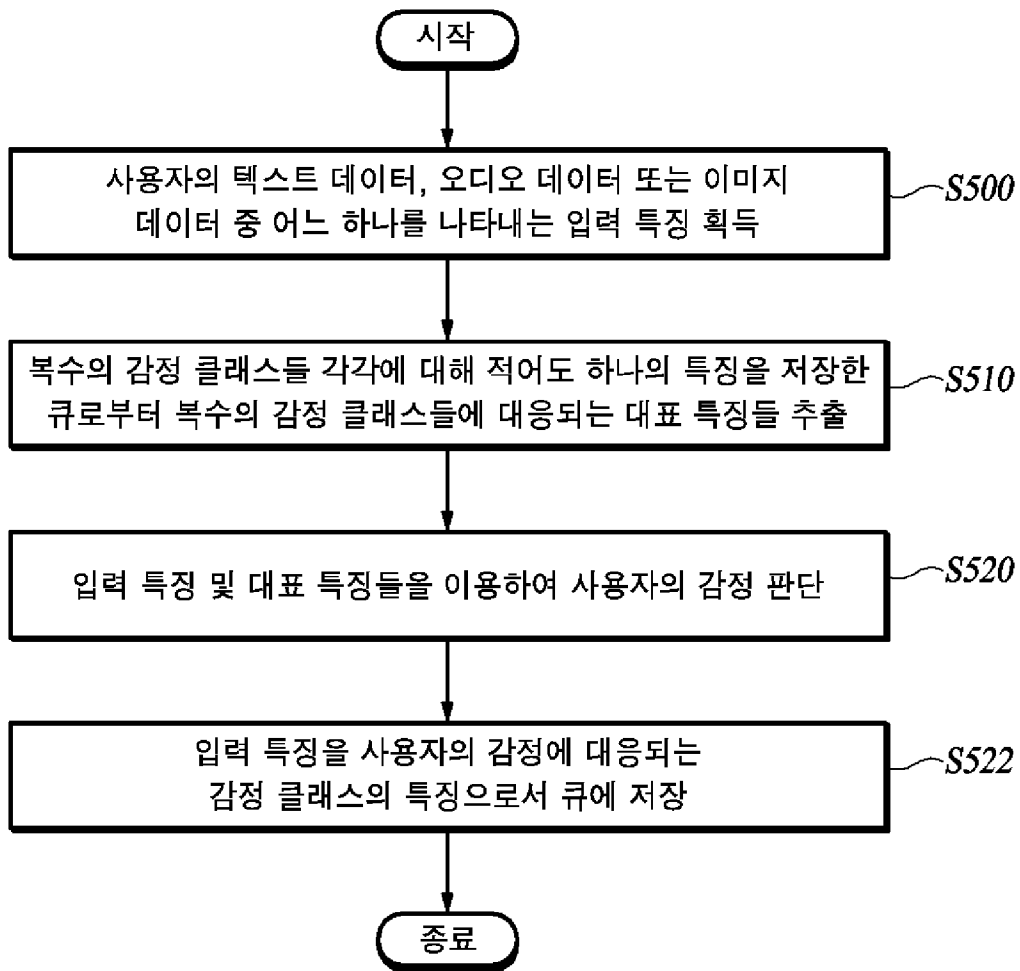
[도4]



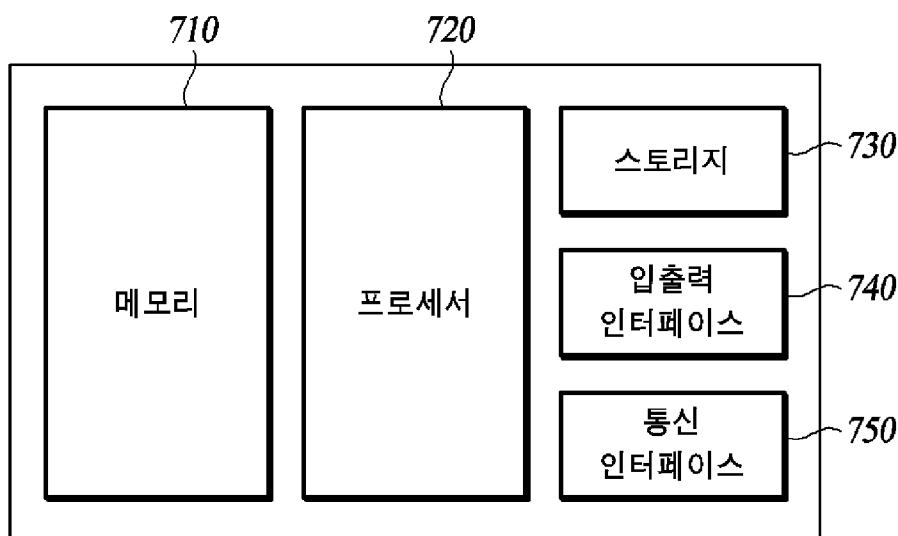
[도5]



[도6]



[도7]

70

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2023/008107

A. CLASSIFICATION OF SUBJECT MATTER G10L 25/63(2013.01)i; G10L 15/04(2006.01)i; G06F 40/30(2020.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G10L 25/63(2013.01); G06F 18/00(2023.01); G06F 40/284(2020.01); G06F 40/30(2020.01); G06K 9/00(2006.01); G06K 9/48(2006.01); G06N 20/00(2019.01); G10L 15/16(2006.01) Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models: IPC as above Japanese utility models and applications for utility models: IPC as above Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) eKOMPASS (KIPO internal) & keywords: 감정(emotion), 인식(recognition), 분류(classification), 업데이트(update), 어텐션 메커니즘(attention mechanism)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KR 10-2021-0106884 A (AJOU UNIVERSITY INDUSTRY-ACADEMIC COOPERATION FOUNDATION) 31 August 2021 (2021-08-31) See paragraphs [0001]-[0095].	1-10
Y	KR 10-2020-0057309 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 26 May 2020 (2020-05-26) See paragraphs [0026]-[0097].	1-10
Y	KR 10-2021-0114137 A (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 23 September 2021 (2021-09-23) See claim 1.	5-6
Y	US 2022-0189502 A1 (TQINTELLIGENCE, INC.) 16 June 2022 (2022-06-16) See paragraph [0052]; and claims 1 and 5.	6
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 26 September 2023		Date of mailing of the international search report 26 September 2023
Name and mailing address of the ISA/KR Korean Intellectual Property Office Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208 Facsimile No. +82-42-481-8578		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2023/008107

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KR 10-2022-0098991 A (INDUSTRY ACADEMY COOPERATION FOUNDATION OF SEJONG UNIVERSITY) 12 July 2022 (2022-07-12) See paragraphs [0007]-[0042].	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No. PCT/KR2023/008107

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
KR 10-2021-0106884 A	31 August 2021	KR 10-2360309 B1	08 February 2022
KR 10-2020-0057309 A	26 May 2020	None	
KR 10-2021-0114137 A	23 September 2021	None	
US 2022-0189502 A1	16 June 2022	US 11508396 B2	22 November 2022
KR 10-2022-0098991 A	12 July 2022	KR 10-2541660 B1	12 June 2023

A. 발명이 속하는 기술분류(국제특허분류(IPC)) G10L 25/63(2013.01)i; G10L 15/04(2006.01)i; G06F 40/30(2020.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) G10L 25/63(2013.01); G06F 18/00(2023.01); G06F 40/284(2020.01); G06F 40/30(2020.01); G06K 9/00(2006.01); G06K 9/48(2006.01); G06N 20/00(2019.01); G10L 15/16(2006.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) eKOMPASS(특허청 내부 검색시스템) & 키워드: 감정(emotion), 인식(recognition), 분류(classification), 업데이트(update), 어텐션 매커니즘(attention mechanism)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	KR 10-2021-0106884 A (아주대학교산학협력단) 2021.08.31 단락 [0001]-[0095]	1-10
Y	KR 10-2020-0057309 A (한국전자통신연구원) 2020.05.26 단락 [0026]-[0097]	1-10
Y	KR 10-2021-0114137 A (한국전자통신연구원) 2021.09.23 청구항 1	5-6
Y	US 2022-0189502 A1 (TQINTELLIGENCE, INC.) 2022.06.16 단락 [0052]; 및 청구항 1, 5	6
A	KR 10-2022-0098991 A (세종대학교산학협력단) 2022.07.12 단락 [0007]-[0042]	1-10
<input type="checkbox"/> 추가 문헌이 C(계속)에 기재되어 있습니다. <input checked="" type="checkbox"/> 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: "A" 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 "D" 본 국제출원에서 출원인이 인용한 문헌 "E" 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 "L" 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 "O" 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 "P" 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌		
국제조사의 실제 완료일 2023년09월26일(26.09.2023)		국제조사보고서 발송일 2023년09월26일(26.09.2023)
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 고재용 전화번호 +82-42-481-8131

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2021-0106884 A	2021/08/31	KR 10-2360309 B1	2022/02/08
KR 10-2020-0057309 A	2020/05/26	없음	
KR 10-2021-0114137 A	2021/09/23	없음	
US 2022-0189502 A1	2022/06/16	US 11508396 B2	2022/11/22
KR 10-2022-0098991 A	2022/07/12	KR 10-2541660 B1	2023/06/12