



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0146067
(43) 공개일자 2024년10월07일

(51) 국제특허분류(Int. Cl.)
A61B 18/20 (2006.01) A61B 17/00 (2022.01)
A61B 18/00 (2022.01) G16H 40/63 (2018.01)
(52) CPC특허분류
A61B 18/203 (2013.01)
G16H 40/63 (2021.08)
(21) 출원번호 10-2024-7030753
(22) 출원일자(국제) 2023년01월30일
심사청구일자 없음
(85) 번역문제출일자 2024년09월12일
(86) 국제출원번호 PCT/EP2023/052131
(87) 국제공개번호 WO 2023/156174
국제공개일자 2023년08월24일
(30) 우선권주장
22156800.9 2022년02월15일
유럽특허청(EPO)(EP)

(71) 출원인
코닌클리케 필립스 엔.브이.
네덜란드 아인트호벤 5656 에이취 하이 테크 캠퍼스 52
(72) 발명자
구자탈, 라시, 쿠마라
네덜란드 아인트호벤 5656 에이취 하이 테크 캠퍼스 52 필립스 인터내셔널 비.브이. 인텔렉추얼 프로퍼티 앤 스탠다드 내
파틸, 라빈드라, 발라사헤브
네덜란드 아인트호벤 5656 에이취 하이 테크 캠퍼스 52 필립스 인터내셔널 비.브이. 인텔렉추얼 프로퍼티 앤 스탠다드 내
구타, 스리니바스
네덜란드 아인트호벤 5656 에이취 하이 테크 캠퍼스 52
(74) 대리인
장훈

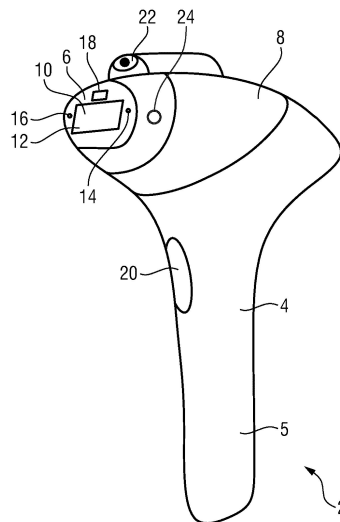
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 치료 디바이스와 함께 사용하기 위한 장치 및 방법

(57) 요약

본 발명은 대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스와 함께 사용하기 위한 장치 및 방법에 관한 것으로, 여기서 치료 디바이스는 치료 동작을 수행하기 위해 신체 부위의 피부에 광 펄스들을 인가하도록 그리고 광 펄스마다 플래시 사운드(flash sound)를 출력하도록 구성된다. 플래시 사운드를 신뢰성 있게 검출하고 과잉치료를 피하도록 사용자에게 알리기 위해 트레이닝된 알고리즘 또는 컴퓨팅 시스템이 실시예에서 사용된다.

대표도 - 도1



(52) CPC특허분류

A61B 2017/00128 (2013.01)

A61B 2018/00476 (2013.01)

A61B 2018/00773 (2013.01)

A61B 2018/00898 (2013.01)

명세서

청구범위

청구항 1

대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스(2)와 함께 사용되도록 구성된 장치(42)로서, 상기 치료 디바이스는 치료 동작을 수행하기 위해 상기 신체 부위의 피부에 광 펄스들을 인가하고, 상기 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드(flash sound)를 생성하도록 구성되고, 상기 장치(42)는

- 상기 치료 동작 동안 감지된 오디오 신호를 획득하도록 구성된 입력부(48);
- 상기 오디오 신호를 오디오 세그먼트들로, 특히 설정된 또는 사전 결정된 지속기간의 오디오 세그먼트들로 세그먼트화하고,
- 오디오 세그먼트마다, 상기 오디오 세그먼트들로부터 하나 이상의 특성 특징부들을 추출하고,
- 오디오 세그먼트마다, 상기 오디오 세그먼트들에서 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드들을 검출하기 위해 상기 특성 특징부들 중 하나 이상을 각각 포함하는 복수의 오디오 세그먼트들에 대해 트레이닝되었던 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 적용함으로써 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드를 상기 오디오 세그먼트가 포함하는지 여부를 검출하고,
- 상기 치료 동작의 시작 이후 또는 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 플래시 사운드 카운터를 변경하도록 구성되는, 프로세싱 유닛(50); 및
- 상기 플래시 사운드 카운터,
- 상기 플래시 사운드 카운터가 플래시 사운드 임계치를 초과하는지 여부를 표시, 및
- 상기 치료 동작이 완료되었다는 표시 중 하나 이상을 출력하도록 구성된 출력부(52)를 포함하는, 장치.

청구항 2

제1항에 있어서,

상기 프로세싱 유닛(50)은 트레이닝된 알고리즘 또는 컴퓨팅 시스템으로서 러닝(learning) 시스템, 머신 러닝, 분류 모델, 이진 분류 모델, 뉴럴 네트워크, 또는 컨볼루션 뉴럴 네트워크를 사용하도록 구성되는, 장치.

청구항 3

제1항 또는 제2항에 있어서,

상기 프로세싱 유닛(50)은 상기 오디오 세그먼트의 시간, 진폭 및 주파수와 관련된 하나 이상의 특징부들을 특성 특징부로서 추출하도록 구성되는, 장치.

청구항 4

제3항에 있어서,

상기 프로세싱 유닛(50)은, 오디오 세그먼트의 특성 특징부로서, 시작 시간, 종료 시간, 진폭, 그레이디언트(gradient), 기기, 키 코드(key chord)들, 멜로디, 리듬, 피치, 에너지, 노트 오프셋(note offset), 하나 이상의 멜-주파수 캡스트럼 계수(mel-frequency cepstral coefficient, MFCC)들, 멜 스펙트로그램(mel spectrogram), 스펙트럼 도심(centroid), 스펙트럼 플럭스(flux), 제로 교차율(zero crossing rate), 비트(beat), 음색(timbre), 피치, 하모니, 제곱 평균 제곱근 에너지, 진폭 엔벨로프(envelope), 대역 에너지 비, 스펙트로그램, 상수 Q 변환(constant Q transform), 스펙트럼 롤 오프(spectral roll off), 및 스펙트럼 확산(spectral spread) 중 하나 이상을 추출하도록 구성되는, 장치.

청구항 5

제3항 또는 제4항에 있어서,

상기 프로세싱 유닛(50)은 상기 오디오 세그먼트의 시간, 진폭 및 주파수와 관련된 상기 추출된 특징부들을 어레이로 배열하고 상기 오디오 세그먼트가 플래시 사운드를 포함하면 오디오 세그먼트마다 상기 어레이를 상기 트레이닝된 알고리즘 또는 컴퓨팅 시스템에 적용하여 검출되도록 구성되는, 장치.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서,

상기 프로세싱 유닛(50)에 의해 오디오 세그먼트에 상기 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 적용하는 것은 상기 하나 이상의 특성 특징부들에 기초하여 분류 경계들을 구축하는 것을 포함하는, 장치.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서,

상기 프로세싱 유닛(50)은 상기 알고리즘 또는 컴퓨팅 시스템의 추가 트레이닝을 위해 상기 검출의 결과를 사용하도록 구성되는, 장치.

청구항 8

대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스(2)와 함께 사용되도록 구성된, 특히 제1항 내지 제7항 중 어느 한 항에 청구된, 장치로서, 상기 치료 디바이스는 치료 동작을 수행하기 위해 상기 신체 부위의 피부에 광 펄스들을 인가하도록 그리고 상기 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드를 생성하도록 구성되고, 상기 장치(42)는

- 상기 치료 동작 동안 감지되거나 인위적으로 생성된 오디오 신호를 획득하도록 구성된 입력부(48);
- 프로세싱 유닛(50)을 포함하고, 상기 프로세싱 유닛(50)은,
- 상기 오디오 신호를 오디오 세그먼트들로, 특히 설정된 또는 사전 결정된 지속기간의 오디오 세그먼트들로 세그먼트화하도록,
- 오디오 세그먼트마다, 상기 오디오 세그먼트들로부터 하나 이상의 특성 특징부들을 추출하도록,
- 오디오 세그먼트마다, 오디오 세그먼트가 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드를 포함하는지 여부를 나타내는 상기 오디오 신호 플래시 사운드 정보의 주석(annotation)으로부터, 상기 오디오 신호 내에서 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드들을 나타내는 상기 오디오 신호의 주석을, 특히 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드들의 시간 정보를 유도하도록, 그리고
- 오디오 세그먼트마다, 상기 추출된 하나 이상의 특성 특징부들 및 상기 유도된 플래시 사운드 정보를 알고리즘 또는 컴퓨팅 시스템에 제공하여 하나 이상의 특성 특징부들에 기초하여 오디오 세그먼트 내에서 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드의 검출을 위해 상기 알고리즘 또는 컴퓨팅 시스템을 트레이닝시키도록 구성되는, 장치.

청구항 9

제8항에 있어서,

상기 입력부(48)는 상기 오디오 신호의 주석을 획득하도록 구성되고 및/또는

상기 프로세싱 유닛(50)은, 특히 자동으로 또는 사용자 입력에 기초하여, 상기 오디오 신호의 주석을 생성하도록 구성되는, 장치.

청구항 10

제1항 내지 제7항 중 어느 한 항에 있어서, 사용자 인터페이스를 제어하여

- 상기 치료 동작에 관한 사용자 입력을 요청하는 제1 스크린을 디스플레이하고;
- 상기 제1 스크린을 통해, 치료될 상기 대상체의 신체 부위의 이미지 및/또는 치료될 상기 대상체의 신체 부위

의 크기의 크기 정보를 포함하는 상기 사용자 입력을 획득하고;

- 상기 장치 또는 상기 장치를 포함하는 전자 디바이스와 상기 치료 디바이스의 연결을 나타내는 제2 스크린을 디스플레이하고;
- 상기 피부 치료 디바이스의 동작 상태,
- 상기 치료 동작의 시작 이후 또는 상기 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 상기 플래시 사운드 카운터,
- 상기 플래시 사운드 카운터가 상기 플래시 사운드 임계치를 초과하는지 여부의 상기 표시, 및
- 상기 치료 동작이 완료되었다는 상기 표시 중 적어도 하나를 나타내는 제3 스크린을 디스플레이하도록 구성되는 출력부(52)를 포함하는, 장치.

청구항 11

시스템으로서,

치료 동작을 수행하기 위해 대상체의 신체의 피부에 인가하기 위한 광 펄스들을 생성하도록 구성된 하나 이상의 광원들(14, 16)을 포함하는 치료 디바이스(2)로서, 상기 치료 디바이스(2)는 상기 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드를 생성하도록 구성되는, 상기 치료 디바이스(2);

상기 치료 동작 동안 오디오 신호를 감지하도록 구성된 오디오 센서(44); 및

제1항 내지 제10항 중 어느 한 항에 청구된 장치(42)를 포함하는, 시스템.

청구항 12

대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스(2)와 함께 사용되도록 구성된 방법으로서, 상기 치료 디바이스는 치료 동작을 수행하기 위해 상기 신체 부위의 피부에 광 펄스들을 인가하도록 그리고 상기 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드를 생성하도록 구성되고, 상기 방법은

- 상기 치료 동작 동안 감지된 오디오 신호를 획득하는 단계,
- 상기 오디오 신호를 오디오 세그먼트들로, 특히 설정된 또는 사전 결정된 지속기간의 오디오 세그먼트들로 세그먼트화하는 단계,
- 오디오 세그먼트마다, 상기 오디오 세그먼트들로부터 하나 이상의 특성 특징부들을 추출하는 단계,
- 오디오 세그먼트마다, 상기 오디오 세그먼트들에서 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드들을 검출하기 위해 상기 특성 특징부들 중 하나 이상을 각각 포함하는 복수의 오디오 세그먼트들에 대해 트레이닝되었던 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 적용함으로써 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드를 상기 오디오 세그먼트가 포함하는지 여부를 검출하는 단계, 및
- 치료 동작의 시작 이후 또는 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 플래시 사운드 카운터를 변경하는 단계; 및
- 하기 중 하나 이상을 출력하는 단계를 포함하는, 방법:
- 상기 플래시 사운드 카운터,
- 상기 플래시 사운드 카운터가 플래시 사운드 임계치를 초과하는지 여부의 표시, 및
- 상기 치료 동작이 완료되었다는 표시.

청구항 13

대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스(2)와 함께 사용되도록 구성된, 특히 제12항에 청구된, 방법으로서, 상기 치료 디바이스는 치료 동작을 수행하기 위해 상기 신체 부위의 피부에 광 펄스들을 인가하도록 그리고 상기 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드를 생성하도록 구성되고, 상기 방법은

- 치료 동작 동안 감지되거나 인위적으로 생성된 오디오 신호를 획득하는 단계,

- 상기 오디오 신호를 오디오 세그먼트들로, 특히 설정된 또는 사전 결정된 지속기간의 오디오 세그먼트들로 세그먼트화하는 단계,
- 오디오 세그먼트마다, 상기 오디오 세그먼트들로부터 하나 이상의 특성 특징부들을 추출하는 단계,
- 오디오 세그먼트마다, 오디오 세그먼트가 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드를 포함하는지 여부를 나타내는 상기 오디오 신호 플래시 사운드 정보의 주석(annotation)으로부터, 상기 오디오 신호 내에서 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드들을 나타내는 상기 오디오 신호의 주석을, 특히 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드들의 시간 정보를 유도하는 단계, 및
- 오디오 세그먼트마다, 상기 추출된 하나 이상의 특성 특징부들 및 상기 유도된 플래시 사운드 정보를 알고리즘 또는 컴퓨팅 시스템에 제공하여 하나 이상의 특성 특징부들에 기초하여 오디오 세그먼트 내에서 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드의 검출을 위해 상기 알고리즘 또는 컴퓨팅 시스템을 트레이닝시키는 단계를 포함하는, 방법.

청구항 14

제12항에 있어서, 상기 방법은 사용자 인터페이스를 제어하여

- 상기 치료 동작에 관한 사용자 입력을 요청하는 제1 스크린을 디스플레이하도록;
- 상기 제1 스크린을 통해, 치료될 상기 대상체의 신체 부위의 이미지 및/또는 치료될 상기 대상체의 신체 부위의 크기의 크기 정보를 포함하는 상기 사용자 입력을 획득하고;
- 상기 사용자 인터페이스를 포함하는 전자 디바이스 또는 장치와 상기 치료 디바이스의 연결을 나타내는 제2 스크린을 디스플레이하고;
- 상기 피부 치료 디바이스의 동작 상태,
- 상기 치료 동작의 시작 이후 또는 상기 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 상기 플래시 사운드 카운터,
- 상기 플래시 사운드 카운터가 상기 플래시 사운드 임계치를 초과하는지 여부를 상기 표시, 및
- 상기 치료 동작이 완료되었다는 상기 표시 중 적어도 하나를 나타내는 제3 스크린을 디스플레이하도록 구성되는, 방법.

청구항 15

컴퓨터 프로그램으로서, 컴퓨터로 하여금 상기 컴퓨터 프로그램이 상기 컴퓨터에서 실행될 때 제12항, 제13항 또는 제14항에 청구된 방법의 단계들을 실행하게 하게 하는 프로그램 코드 수단을 포함하는, 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스와 함께 사용되도록 구성된 장치 및 방법에 관한 것으로, 여기서 치료 디바이스는 치료 동작을 수행하기 위해 신체 부위의 피부에 광 펄스들을 인가하도록 그리고 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드(flash sound)를 생성하도록 구성된다. 본 발명은 추가로, 그러한 장치의 컴퓨터 구현 방법, 치료 디바이스를 포함하는 시스템, 및 컴퓨터 프로그램에 관한 것이다.

배경 기술

[0002] 원치 않는 모발을 제거하는 기술은 면도, 전기 분해, 뽑기, 레이저 및 광 요법(광제모로 알려짐) 및 치료용 항-안드로겐 주입을 포함한다. 광-기반 기술은 또한 모발 성장 감소 및 여드름 치료를 포함한 다른 유형의 피부과 치료에도 사용된다.

[0003] 광 에너지의 적절한 구성 즉, 파장, 강도 및/또는 펄스 지속기간(광이 펄스되는 경우)의 관점에서 광 에너지의 적절한 구성의 사용을 통해, 모근의 선택적 가열 및 모낭에 후속적인 일시적 또는 영구적 손상이 달성될 수 있

다. 가정용 광제모 디바이스, 예를 들어 Philips Lumea 디바이스는 고강도 광원, 예를 들어 광대역 광의 고출력 버스트를 생산하는 크세논 플래시 램프의 강한 펄스 광(IPL)을 사용한다.

[0004] 광제모 치료는 광제모 디바이스 사용자가 제모를 목적으로 피부의 상대적으로 작은 영역을 치료하는 것이 특징이다. 광제모 치료는 강한 빛을 사용하여 모발과 모근의 멜라닌을 가열하여, 모낭을 휴식 단계에 넣어 모발의 재성장을 방지한다. 제모에 이 기술을 효과적으로 사용하기 위해, 사용자는 피부에 틈을 남기지 않고 완전히 치료해야 한다. 이 효과는 지속시간이 제한되기 때문에 정기적으로 치료를 반복해야 한다: 전형적으로 2주에 한 번씩 치료를 수행하는 약 2개월의 초기 기간 이후 유지 단계에서 4 내지 8주에 한 번씩 반복해야 한다.

[0005] Lumea 디바이스는 Lumea 앱(즉, 컴퓨터, 랩톱, 태블릿 및/또는 스마트폰에서 실행될 수 있는 애플리케이션 프로그램 또는 소프트웨어 애플리케이션, 이는 예를 들어, 치료 디바이스에 무선으로 결합될 수 있음)에 의해 지원되어, 사용자의 제모 치료를 돕는다. 이는 Lumea 디바이스에 대한 사용자 경험 및 사용자 참여를 향상시킨다.

[0006] 많은 광제모 디바이스들에서, 치료를 위해 방출된 광의 각각의 플래시/버스트(burst)에는 플래시 사운드가 바로 후속되는데, 이는 (예컨대, 대략 1초의) 시간 및 (예컨대, 대략 9 kHz)의 주파수 둘 모두에서 특징적인 짧은 사운드 임펄스이다. 치료에 필요한 플래시들의 수는 펄스들이 인가되는 신체 부위의 크기에 의존한다. 권장량을 초과하는 지속기간 동안 고강도 레이저 펄스에 대한 노출은 피부에 대한 심각한 손상 및 과잉치료를 야기할 수 있다.

[0007] WO 2019/079972 A1호는 하드웨어 장비에 대한 간단한 알고리즘, 적은 계산량, 낮은 요건으로 특정 음성들이 인식될 수 있게 하는 특정 음성 인식 방법을 개시한다. 방법은 음성 신호를 샘플링하는 단계, 음성 신호의 멜(Me1)-주파수 캡스트럼 계수 특성 파라미터 매트릭스를 획득하는 단계, 및 특성 파라미터 매트릭스로부터 특성 파라미터들을 추출하는 단계를 포함한다. 특성 파라미터들은 식별을 위해 딥 뉴럴 네트워크-기반의 특정 음성 특징부 모델에 공급되어, 음성 신호가 특정 음성 특징부를 포함하는지 여부를 결정하게 한다. 특정 음성 인식 디바이스는 사용자의 음성을 수신하도록 그리고 음성이 특정 음성 특징부를 포함하는지 여부를 결정하도록 구성된다. WO 2019/079972 A1호에서 언급된 특정 음성 특징부의 예는 기침 소리이다.

발명의 내용

[0008] 본 발명의 목적은 과잉치료의 위험을 회피하거나 적어도 감소시키기 위한 조치를 제공하는 것이다.

[0009] 본 발명의 제1 태양에서, 대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스와 함께 사용되도록 구성된 장치가 제시되는데, 치료 디바이스는 치료 동작을 수행하기 위해 신체 부위의 피부에 광 펄스들을 인가하도록 그리고 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드를 생성하도록 구성되고, 장치는

[0010] - 치료 동작 동안 감지된 오디오 신호를 획득하도록 구성된 입력부;

[0011] - 프로세싱 유닛으로서,

[0012] - 오디오 신호를 오디오 세그먼트들로, 특히 설정된 또는 사전결정된 지속기간의 오디오 세그먼트들로 세그먼트화하도록,

[0013] - 오디오 세그먼트마다, 오디오 세그먼트들로부터 하나 이상의 특성 특징부들을 추출하도록,

[0014] - 오디오 세그먼트마다, 오디오 세그먼트들에서 광 펄스마다 치료 디바이스에 의해 생성되는 플래시 사운드들을 검출하기 위해 특성 특징부들 중 하나 이상을 각각 포함하는 복수의 오디오 세그먼트들에 대해 트레이닝되었던 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 적용함으로써 광 펄스마다 치료 디바이스에 의해 생성되는 플래시 사운드를 오디오 세그먼트가 포함하는지 여부를 검출하도록, 그리고

[0015] - 치료 동작의 시작 이후 또는 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 플래시 사운드 카운터를 변경하도록 구성되는, 프로세싱 유닛; 및

[0016] - 하기 중 하나 이상을 출력하도록 구성된 출력부를 포함한다:

[0017] - 플래시 사운드 카운터,

[0018] - 플래시 사운드 카운터가 플래시 사운드 임계치를 초과하는지 여부를 표시, 및

[0019] - 치료 동작이 완료되었다는 표시.

- [0020] 본 발명의 제2 태양에서, 대상체의 신체 부위에 대한 치료 동작을 수행하도록 구성된 치료 디바이스와 함께 사용되도록 구성된 장치가 제시되는데, 치료 디바이스는 치료 동작을 수행하기 위해 신체 부위의 피부에 광 펄스들을 인가하도록 그리고 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드를 생성하도록 구성되고, 장치는
- [0021] - 치료 동작 동안 감지되거나 인위적으로 생성된 오디오 신호를 획득하도록 구성된 입력부;
- [0022] - 프로세싱 유닛으로서,
- [0023] - 오디오 신호를 오디오 세그먼트들로, 특히 설정된 또는 사전결정된 지속기간의 오디오 세그먼트들로 세그먼트화하도록,
- [0024] - 오디오 세그먼트마다, 오디오 세그먼트들로부터 하나 이상의 특성 특징부들을 추출하도록,
- [0025] - 오디오 세그먼트마다, 오디오 세그먼트가 광 펄스마다 치료 디바이스에 의해 생성되는 플래시 사운드를 포함하는지 여부를 나타내는 오디오 신호 플래시 사운드 정보의 주석(annotation)으로부터, 오디오 신호 내에서 광 펄스마다 치료 디바이스에 의해 생성되는 플래시 사운드들을 나타내는 상기 오디오 신호의 주석을, 특히 광 펄스마다 치료 디바이스에 의해 생성되는 플래시 사운드들의 시간 정보를 유도하도록, 그리고
- [0026] - 오디오 세그먼트마다, 추출된 하나 이상의 특성 특징부들 및 상기 유도된 플래시 사운드 정보를 알고리즘 또는 컴퓨팅 시스템에 제공하여 하나 이상의 특성 특징부들에 기초하여 오디오 세그먼트 내에서 광 펄스마다 상기 치료 디바이스에 의해 생성되는 플래시 사운드의 검출을 위해 상기 알고리즘 또는 컴퓨팅 시스템을 트레이닝시키도록 구성되는, 상기 프로세싱 유닛을 포함한다.
- [0027] 본 발명의 실시예에서, 장치는 사용자 인터페이스를 제어하여
- [0028] - 치료 동작에 관한 사용자 입력을 요청하는 제1 스크린을 디스플레이하도록;
- [0029] - 제1 스크린을 통해, 치료될 대상체의 신체 부위의 이미지 및/또는 치료될 대상체의 신체 부위의 크기의 크기 정보를 포함하는 사용자 입력을 획득하도록;
- [0030] - 장치 또는 장치를 포함하는 전자 디바이스와 치료 디바이스의 연결을 나타내는 제2 스크린을 디스플레이하도록;
- [0031] - 하기 중 적어도 하나를 나타내는 제3 스크린을 디스플레이하도록 구성되는 출력부를 포함한다:
- [0032] - 피부 치료 디바이스의 동작 상태,
- [0033] - 치료 동작의 시작 이후 또는 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 플래시 사운드 카운터,
- [0034] - 플래시 사운드 카운터가 플래시 사운드 임계치를 초과하는지 여부를 표시, 및
- [0035] - 치료 동작이 완료되었다는 표시.
- [0036] 본 발명의 제3 태양에서, 시스템이 제시되는데, 시스템은
- [0037] 치료 동작을 수행하기 위해 대상체의 신체의 피부에 인가하기 위한 광 펄스들을 생성하도록 구성된 하나 이상의 광원들을 포함하는 치료 디바이스로서, 치료 디바이스는 치료 디바이스에 의해 생성되는 광 펄스마다 플래시 사운드를 생성하도록 구성되는, 치료 디바이스;
- [0038] 치료 동작 동안 오디오 신호를 감지하도록 구성된 오디오 센서; 및
- [0039] 본 명세서에 개시된 바와 같은 상이한 태양들 중 어느 하나에 따른 장치를 포함한다.
- [0040] 본 발명의 또 다른 태양에서, 대응하는 (컴퓨터 구현) 방법, 컴퓨터 프로그램이 컴퓨터 상에서 실행될 때 본 명세서에 개시된 각각의 방법들의 단계들을 컴퓨터가 실행하게 하기 위한 프로그램 코드 수단을 포함하는 컴퓨터 프로그램, 및 프로세서가 실행할 때, 본 명세서에 개시된 각각의 방법들이 실행되게 하는 컴퓨터 프로그램 제품을 저장하는 비일시적 컴퓨터 판독가능 기록 매체가 제공된다.
- [0041] 본 발명의 바람직한 실시예는 종속 청구항에 정의되어 있다. 청구된 방법, 시스템, 컴퓨터 프로그램 및 매체는 청구된 장치와 유사하거나 및/또는 동일한 바람직한 실시예를 가지며, 특히 종속 청구항에 정의되고 본 명세서에 개시된 바와 같다는 것이 이해될 것이다.

- [0042] 본 발명의 모든 태양들은, 레이저 제모 치료의 불리한 태양을 고려하여, 사용자가 치료에 필요한 플래시들의 수를 정확하게 알아야 하고/하거나 원하는 영역의 치료가 완료되고 충분한 수의 플래시들로 치료된 때를 알아야 한다는 공통 아이디어에 기초한다. 단일 치료는 전형적으로, 예컨대, (원하는 치료 영역의 길이에 따라) 약 30 내지 40개의 플래시들을 취한다. 사용자는 한 번의 치료 세션 동안 적용된 플래시들의 수를 수동으로 카운트할 수 있고, 플래시들의 어떠한 권장하는 수가 없어서 이는 열악한 사용자 경험으로 이어진다. 그러나, 사용자는 달리 다른 작업들에 열중해 있거나 배경 노이즈에 의해 방해받을 수 있어서, 사용자는 플래시들의 총 수를 정확하게 추적하는 데 어려움을 겪을 수 있다. 따라서, 치료 세션 동안 플래시 사운드들의 (그리고, 그에 따른 방출된 플래시들의 수의) 카운트를 자동화하는 것은 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 사용하여 본 발명(제1 태양)에 의해 제공된다. 이러한 트레이닝된 알고리즘 또는 컴퓨팅 시스템은 각각의 치료 동안 플래시들의 총 수를 정확하게 카운트하고 치료 디바이스 및 환경 둘 모두로부터 발생하는 배경 노이즈를 처리하도록 트레이닝될 수 있다. 추가로, 신호들의 비주기적 특성이 고려될 수 있다. 따라서, 치료 세션 동안 피부 치료 디바이스에 의해 방출되는 플래시들의 수는 향상된 정확도로 획득될 수 있다.
- [0043] 본 발명의 맥락에서, 치료 동안 플래시 사운드들의 수를 카운트하기 위해 카운터를 변경하는 것은, 플래시 사운드가 검출될 때마다, 카운터 값이 (예컨대, 0에서 시작하여) 증분되거나 (예컨대, 소정 치료 영역을 치료하기 위한 추정된 최대 카운터 값에서 시작하여) 감소되도록 하는 것으로 이해하여야 한다.
- [0044] 트레이닝은 오디오 세그먼트들에서 플래시 사운드들을 검출하기 위해 오디오 세그먼트(예컨대, 예를 들어, 1초의 지속기간의 세그먼트)로부터 유도된 하나 이상의 특성 특징부들을 각각 포함하는 복수의 오디오 세그먼트들의 사용에 의해 본 발명(제2 태양)에 따라 이루어진다.
- [0045] 추가로, 본 발명(제3 태양)에 따르면, 치료 세션 동안 방출된 플래시들의 수를 포함한 다양한 정보를 효율적으로 획득하여 사용자에게 제시하는 사용자 인터페이스가 제공된다. 이는 사용자가 현재 치료 경과를 용이하게 이해하고 과잉치료를 회피하는 것을 돕는다.
- [0046] 바람직하게는, 본 발명은 트레이닝된 알고리즘 또는 컴퓨팅 시스템으로서 러닝(learning) 시스템, 머신 러닝, 분류 모델, 이진 분류 모델, 뉴럴 네트워크, 또는 컨볼루션 뉴럴 네트워크를 사용한다. 예를 들어, 구현예에서, 종래에 알려진 바와 같이, XGBoost 분류기가 사용될 수 있다. XGBoost는 그레이디언트 부스팅 프레임워크(gradient boosting framework)를 사용하는 결정 트리-기반 앙상블 머신 러닝 알고리즘이다. 이는, 예컨대, 회귀, 분류, 랭킹, 및 사용자 정의 예측 문제들을 해결하기 위해, 광범위한 응용들에 사용될 수 있다. 이는 트레이닝 단계(시스템/알고리즘이 트레이닝될 때)뿐만 아니라 동작 단계(시스템/알고리즘이 플래시 사운드 검출에 사용될 때) 둘 모두에서 최소 프로세싱 시간에 높은 예측 성능을 제공할 수 있다.
- [0047] 다른 실시예에서, 오디오 세그먼트의 시간, 진폭 및 주파수와 관련된 하나 이상의 특징부들이 각각의 오디오 세그먼트로부터 추출될 수 있는 특성 특징부(들)로서 사용될 수 있다.
- [0048] 대체적으로, 시작 시간, 종료 시간, 진폭, 그레이디언트, 기기, 키 코드(key chord)들, 멜로디, 리듬, 피치, 에너지, 노트 오프셋(note offset), 하나 이상의 멜-주파수 캡스트럼 계수(mel-frequency cepstral coefficient, MFCC)들, 멜 스펙트로그램(mel spectrogram), 스펙트럼 도심(centroid), 스펙트럼 플럭스(flux), 제로 교차율(zero crossing rate), 비트(beat), 음색(timbre), 피치, 하모니, 제곱 평균 제곱근 에너지, 진폭 엔벨로프(envelope), 대역 에너지 비, 스펙트로그램, 상수 Q 변환(constant Q transform), 스펙트럼 롤 오프(spectral roll off), 및 스펙트럼 확산(spectral spread) 중 하나 이상을 포함하는, 광범위한 특징부들이 오디오 세그먼트의 특성 특징부(들)로서 사용하는 데 이용가능하다. 구현예에서, 8 내지 13개의 캡스트럼 계수들을 갖는 선택된 MFCC 특징부들이 오디오 세그먼트에 대체적으로 이용가능한 훨씬 더 많은 수의, 예컨대, 39개의 MFCC 특징부들 대신 사용될 수 있다. 바람직하게는, 단지 작은 화자-특정 정보만을 전달하는 입력 신호의 평균 로그 에너지를 나타내는 0차 계수는, 그것이 배경 노이즈의 검출로 이어질 수 있기 때문에, 사용되지 않는다(플래시 사운드는 대체적으로 약 0.1초 동안 일시적으로 발생한다). 따라서, 본 발명에 따르면, 가변 배경 노이즈의 존재 하에서 광제모 디바이스에 의해 생성되는 낮은 진폭 및 짧은 지속기간을 갖는 플래시 사운드들이 효과적으로 검출될 수 있다. 추가로, 요구되는 낮은 레이턴시 및 메모리가 제공될 수 있다.
- [0049] 플래시 램프와 같은 광원을 트리거함으로써 플래시 사운드가 생성될 수 있거나, 그것은 LED와 같은 광원이 사용되는 경우에 인위적으로 합성될 수 있다. 예를 들어, 동작 동안 플래시 램프 또는 제논 아크 램프가 펄스 방전으로 인해 (플래시 사운드를 나타내는) 사운드들을 생성한다. 레이저 방사선은 공기 중에 전기 고장을 야기할 수 있고, 이는 사운드를 생성한다. 마찬가지로, 강한 펄스형 광이 피부와 접촉할 때, 사운드가 생성된다. 그러나, 치료 디바이스의 다른 실시예들은 방전 사운드를 생성하지 않는 LED들과 같은 반도체 광원들을 사용할 수

있는데, 이는 그들이 상이하게 기능하기 때문이다. 이어서, 치료를 위한 광 펄스가 방출될 때마다 플래시 사운드를 합성하기 위해 추가 사운드 생성기가 사용될 수 있다.

[0050] 제1 태양의 다른 실시예에 따르면, 프로세싱 유닛은 오디오 세그먼트의 시간, 진폭 및 주파수와 관련된 추출된 특징부들을 어레이로 배열하도록, 그리고 오디오 세그먼트가 플래시 사운드를 포함하는지 여부를, 오디오 세그먼트마다 검출하기 위해 어레이를 트레이닝된 알고리즘 또는 컴퓨팅 시스템에 거치게 하도록 구성될 수 있다. 어레이 포맷은, 예컨대, 추출된 특징부들이 시간에 관련되고, 진폭에 관련되고, 주파수에 관련된 특징부들에 따라 배열 또는 정렬되도록 구성될 수 있으며, 이는 XGBoost 분류 모델과 같은 트레이닝된 알고리즘 또는 컴퓨팅 시스템으로 그들을 프로세싱하는 것을 더 용이하게 하거나 심지어 가능하게 한다.

[0051] 프로세싱 유닛에 의해 오디오 세그먼트에 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 적용하는 것은 하나 이상의 특성 특징부들에 기초하여 분류 경계들을 구축하는 것을 포함할 수 있다. 일 예로서, XGBoost는 전달된 입력 특징부에 기초하여 분류 경계들을 구축하기 위해 앙상블 트리-기반 기법을 사용한다. 그것은 병렬 프로세싱, 트리 프루닝(tree-pruning) 및 정규화를 사용하여 과적합성을 회피하고 균형 잡힌 모델을 제공한다. 분류 경계들에 기초하여, 각각의 오디오 세그먼트는 플래시들의 총 수가 치료 동안 최종적으로 예측될 수 있도록 플래시들 및 비-플래시들 중 어느 하나로 분류될 수 있다.

[0052] 제1 태양의 다른 실시예에서, 프로세싱 유닛은 알고리즘 또는 컴퓨팅 시스템의 추가 트레이닝을 위해 검출의 결과를 사용하도록 구성된다. 따라서, 동작 동안, 검출 결과들은 플래시 사운드 검출을 위해 사용될 뿐만 아니라, 동시에 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 추가로 개선하기 위한 추가 트레이닝을 위해 사용된다.

[0053] 본 발명의 제2 태양에 따르면, 치료 동작 동안 감지되거나 인위적으로 생성된 오디오 신호가 알고리즘 또는 컴퓨팅 시스템을 트레이닝하기 위해 사용된다. 실제 치료 동작 동안 행해지는 바와 동일한 방식으로, 오디오 신호는 오디오 세그먼트들로 세그먼트화되고 하나 이상의 특성 특징부들은 오디오 세그먼트들로부터 추출된다. 오디오 신호의 주석으로부터, 오디오 세그먼트가 플래시 사운드를 포함하는지 여부를 나타내는 플래시 사운드 정보가 하나 이상의 특성 특징부들로부터 유도된다. 이러한 맥락에서, 오디오 신호의 주석은 오디오 신호 내의 플래시 사운드들의 표시, 특히 플래시 사운드들의 시간 정보(예컨대, 시작 시간, 종료 시간, 피크 시간, 시간 지속기간 등과 같은, 오디오 신호가 어느 시점에서 플래시 사운드를 포함하는지를 나타내는 타임 스탬프 또는 다른 표시)로서 대략적으로 이해되어야 한다. 이어서, 추출된 특성 특징부들 및 유도된 플래시 사운드 정보는 하나 이상의 특성 특징부들에 기초하여 오디오 세그먼트 내에서 플래시 사운드의 검출을 위한 알고리즘 또는 컴퓨팅 시스템을 트레이닝하는 데 사용된다.

[0054] 오디오 신호의 주석은, 입력이, 예컨대, 오디오 신호와 함께 또는 그와 별개로 저장된 바와 같은, 오디오 신호의 주석을 획득할 수 있도록 사용자에게 의해 미리 만들어질 수 있다. 대안적으로 또는 추가적으로, 주석은, 예컨대, 알고리즘 또는 컴퓨팅 시스템 또는 다른 프로세싱 엔티티에 의해, 자동으로 만들어질 수 있다. 예를 들어, 프로세싱 유닛은, 예컨대, 자동으로 또는 사용자 입력에 기초하여, 오디오 신호의 주석을 생성하도록 구성될 수 있다.

[0055] 본 발명의 제3 태양에 따르면, 상이한 발명이 획득될 수 있고 사용자에게 제시될 수 있다. 이러한 정보는 다양한 스크린들 상에 디스플레이될 수 있으며, 여기서 제1 스크린은 치료 동작에 관한 사용자 입력을 요청하고, 제2 스크린은 장치 또는 장치를 포함하는 전자 디바이스와 치료 디바이스의 연결을 나타내고, 제3 스크린은 피부 치료 디바이스의 동작 상태, 치료 동작의 시작 이후 또는 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 플래시 사운드 카운터, 플래시 사운드 카운터가 플래시 사운드 임계치를 초과하는지 여부를 표시, 및 치료 동작이 완료되었다는 표시 중 하나 이상을 나타낸다. 이는 사용자가 치료 동작의 현재 진행 및 언제 치료를 완료하는지를 알 수 있게 하여 과잉치료가 방지될 수 있거나 적어도 감소될 수 있게 한다.

[0056] 본 발명에 따른 시스템은 치료 디바이스, 오디오 센서, 및 전술한 바와 같은 장치들 중 하나 이상을 포함한다. 시스템의 실시예에 따르면, 오디오 센서, 예컨대, 마이크로폰은 장치 또는 장치를 포함하는 전자 디바이스의 일부, 특히 컴퓨터, 랩톱, 태블릿 또는 스마트폰의 일부(또는 장치만을 나타내는 별개의 엔티티)이다. 다른 실시예에서, 장치는 치료 디바이스 또는 전자 디바이스, 특히 컴퓨터, 랩톱, 태블릿 또는 스마트폰의 일부이다. 일반적으로, 전술된 태양들의 장치들은 동일한 디바이스 상에서 또는 상이한 디바이스들 상에서 구현될 수 있다. 실제 구현예에서, 모든 태양들의 장치들 및 방법들은 태블릿 또는 스마트폰과 같은 동일한 디바이스 상에서 구현된다.

도면의 간단한 설명

[0057]

본 발명의 이들 및 다른 태양들이 이하에 설명되는 실시예(들)로부터 명백할 것이며 그것을 참조하여 설명될 것이다. 다음 도면에서

도 1은 본 발명이 사용될 수 있는 치료 디바이스의 예시적인 실시예의 예시를 도시한다.

도 2는 본 발명에 따른 시스템의 예시적인 실시예의 개략도를 도시한다.

도 3은 본 발명의 제1 태양에 따른 방법의 예시적인 실시예의 흐름도를 도시한다.

도 4는 본 발명의 제2 태양에 따른 방법의 실시예의 흐름도를 도시한다.

도 5는 플래시 사운드들 및 배경 노이즈의 표시들에 의한 오디오 신호의 시각화를 도시한다.

도 6은 2개의 플래시 사운드들에 대한 zoom된 시각화를 도시한다.

도 7은 본 발명에 따른 프리-프로세싱을 예시하는 개략도를 도시한다.

도 8은 본 발명에 따른 특징부 추출을 예시하는 개략도를 도시한다.

도 9는 본 발명에 따른 모델 배치부터 생성까지의 프로세스의 흐름도를 도시한다.

도 10은 본 발명의 제3 태양에 따른 방법의 실시예의 흐름도를 도시한다.

도 11은 본 발명에 따른 사용자 인터페이스를 예시하는 상이한 스크린들을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0058]

도 1은 피부 영역에 (플래시로 또한 지칭되는) 광 펄스를 인가하는 데 사용될 수 있는 예시적인 치료 디바이스(2)의 예시이다. 도1의 치료 디바이스(2)는 단지 본 발명이 사용될 수 있는 휴대용 치료 디바이스(2)의 예로서 제시된 것이며, 치료 디바이스(2)는 도1에 도시된 형태 또는 휴대용 치료 디바이스인 것으로 한정되지 않는다는 것이 인식될 것이다. 치료 디바이스(2)는 대상체(예를 들어, 사람 또는 동물)의 신체에 사용하기 위한 것이며, 사용 동안 사용자가 한손 또는 양손에 쥐어야 한다. 치료 디바이스(2)는 치료 디바이스(2)가 대상체의 신체 부위에 접촉하거나 근접할 때 하나 이상의 광 펄스를 사용하여 대상체의 피부 또는 신체에 대한 일부 치료 동작을 수행하는 것이다. 치료 동작은 레이저 및/또는 광 요법(광제모 치료 또는 Intense Pulsed Light, IPL 치료로 공지됨)으로 원치 않는 모발을 제거하는 것일 수 있다. 대안적으로, 치료 동작은 모발의 제거 또는 성장 방지 이외의 이유, 예컨대 피부과(피부) 치료, 여드름 치료, 다른 유형의 피부 상태에 대한 광 요법 치료, 피부 회춘, 피부 타이팅, 또는 포트 와인 얼룩 치료로 레이저 및/또는 광요법일 수 있다.

[0059]

본 명세서에서 설명되는 바와 같이, 치료 디바이스(2)는 '사용자'에 의해 동작 또는 사용되며, 치료 디바이스(2)는 '대상체'의 신체에 사용된다. 일부 경우, 사용자와 대상체가 동일인, 즉 치료 디바이스(2)를 손에 쥐고 사용자가 스스로 사용한다(예를 들어, 다리의 피부에 사용). 다른 경우에, 사용자와 대상체는 다른 사람, 예를 들어, 사용자가 치료 디바이스(2)를 손에 쥐고 다른 사람에게 사용한다. 어느 경우든, 사용자가 신체 부위의 완전한 커버리지를 달성하거나 신체 부위의 특정 영역의 과도한 치료를 피하는 것은 어려운데, 이는 광 펄스를 인가할 때 또는 광 펄스를 인가한 직후에 피부에 사용자가 인지할 수 있는 변화가 거의 또는 전혀 없기 때문이다.

[0060]

예시적인 치료 디바이스(2)는 적어도 손잡이 부분(5) 및 헤드 부분(6)을 포함하는 하우징(4)을 포함한다. 손잡이 부분(5)은 사용자가 한 손으로 치료 디바이스(2)를 잡을 수 있는 형태로 되어 있다. 헤드 부분(6)은 하우징(4)의 헤드 단부(8)에 위치하고, 헤드 부분(6)이 신체 또는 피부에 접촉하는 위치에서 대상체의 신체 또는 피부에 치료 동작이 수행되도록 하기 위해 헤드 부분(6)이 대상체와 접촉하도록 배치되어야 한다.

[0061]

치료 디바이스(2)는 광 펄스를 이용하여 치료 동작을 수행하기 위한 것이다. 따라서, 도 1에서, 헤드 부분(6)은 하우징(4) 내부 또는 위에 배열된 조리개(10)를 포함하여, 조리개(10)가 대상체의 피부에 인접하거나 위에(즉, 접촉) 배치될 수 있다. 치료 디바이스(2)는 조리개(10)를 통해 대상체의 피부에 인가되고 치료 동작에 영향을 미치는 광 펄스를 발생시키기 위한 하나 이상의 광원(12)을 포함한다. 하나 이상의 광원(12)은 하우징(4)에 배열되어, 광 펄스가 하나 이상의 광원(12)으로부터 조리개(10)를 통해 제공된다. 조리개(10)는 하우징(4)의 헤드 단부(8)에서 개구 형태이거나, 광 펄스에 대해 투명하거나 반투명한(즉, 광 펄스가 원도구를 통과할

수 있는) 윈도우(도파관 포함)의 형태일 수 있다.

- [0062] 도 1에 도시된 예시적인 실시예에서, 조리개(10)는 일반적으로 직사각형 형상을 가지며, 이는 피부 상에 일반적으로 직사각형 형상의 피부 치료 영역을 초래한다. 조리개(10)는 임의의 다른 원하는 형상을 가질 수 있다는 것을 인식할 수 있을 것이다. 예를 들어, 조리개(10)는 정사각형, 타원형, 원형, 또는 임의의 다른 다각형 형상일 수 있다.
- [0063] 도 1에 도시되지 않았지만, 치료 디바이스(2)는 각각 조리개(10)를 포함하는 치료 디바이스(2)의 헤드 부분(6)에 대해 하나 이상의 제거가능한 부착물을 가질 수 있다. 부착물은 상이한 신체 부위에 사용하기 위한 것일 수 있으며, 상이한 크기의 조리개(10)를 갖는다. 예를 들어, 하나의 부착물은 큰 신체 부위, 예컨대 다리에 사용하기 위한 것이고, 따라서 큰 조리개(10)를 가질 수 있는 반면, 다른 부착물은 윗입술 위의 피부/모발에 사용하기 위한 것이고, 따라서 작은 조리개(10)를 가질 수 있다.
- [0064] 하나 이상의 광원(12)은 임의의 적합하거나 원하는 파장(또는 파장의 범위) 및/또는 강도의 광 펄스를 발생시킬 수 있다. 예를 들어, 광원(12)은 가시광선, 적외선(IR) 광 및/또는 자외선(UV) 광을 발생시킬 수 있다. 각각의 광원(12)은 임의의 적합한 유형의 광원, 예컨대 하나 이상의 발광 다이오드(LED), (크세논) 플래시 램프, 레이저 또는 레이저들 등을 포함할 수 있다. 광원(들)(12)은 약 2.5ms의 지속시간 동안 450-1200nm 범위의 스펙트럼을 갖는 광 펄스를 제공할 수 있는데, 이는 이들 파장이 흡수에 의해 모발 및 모근 내의 멜라닌을 가열함으로써, 모낭을 휴식 단계에 두어 모발 재생장을 방지한다.
- [0065] 하나 이상의 광원(12)은 광의 펄스를 제공하도록 구성된다. 즉, 광원(들)(12)은 짧은 지속시간(예를 들어, 1초 미만) 동안 높은 강도로 광을 발생시키도록 구성된다. 광 펄스의 강도는 조리개(10)에 인접한 피부 또는 신체 부위에 대한 치료 동작에 영향을 미칠 만큼 충분히 높아야 한다.
- [0066] 예시된 치료 디바이스(2)는 선택적으로 헤드 부분(6) 위에 또는 안에 위치하는 2개의 피부 접촉 센서(14, 16)를 포함할 수 있으며, 이는 헤드 부분(6)이 피부와 접촉하고 있는지 여부를 결정하는 데 사용된다. 피부 접촉 센서(14, 16)는 헤드 부분(6)이 피부에 접촉되어 있는지 여부를 표시하는 파라미터를 측정하고, 파라미터의 측정의 시계열을 포함하는 각각의 측정 신호를 발생시킨다. 측정 신호는 헤드 부분(6)이 피부와 접촉하고 있는지를 결정하기 위해 프로세싱될 수 있다. 전형적으로 피부 접촉 센서는 광 펄스가 사용자 또는 대상체의 눈으로 향하는 것을 피하기 위해 광 펄스가 발생되기 전에 치료 디바이스(2)가 피부와 정확하게 접촉하고 있는지 확인하기 위해 치료 디바이스(2), 특히 광제모기에서 사용된다.
- [0067] 일부 실시예에서, 파라미터는 커패시턴스일 수 있고, 따라서 피부 접촉 센서(14, 16)는 헤드 부분(6)의 표면 상의 전기적 접촉 또는 전극의 각각의 쌍을 통해 커패시턴스를 측정할 수 있으며, 측정된 커패시턴스는 피부 접촉이 있는지를 표시한다. 대안적인 실시예에서, 파라미터는 빛의 강도 또는 레벨일 수 있고, 따라서 피부 접촉 센서들(14, 16)은 광 센서에 입사되는 빛의 강도 또는 레벨을 측정하는 광 센서들일 수 있고, 측정된 강도 또는 레벨은 피부 접촉이 있는지 여부(예를 들어, 피부가 광 센서들(14, 16)을 가리기 때문에 더 적은/어떤 빛도 피부 접촉을 표시할 수 없고, 그 반대의 경우도 마찬가지)를 표시한다. 다른 대안적인 실시예에서, 파라미터는 접촉 압력의 척도일 수 있고, 따라서 피부 접촉 센서(14, 16)는 각각의 압력 센서 또는 기계적 스위치를 통해 접촉 압력을 측정할 수 있으며, 측정된 접촉 압력은 피부 접촉이 있는지 여부를 표시한다.
- [0068] 예시된 치료 디바이스(2)는 헤드 부분(6) 상에 또는 안에 위치하여 헤드 부분(6)이 접촉하는 피부의 피부 톤을 결정하기 위해 사용되는 선택적인 피부 톤 센서(18)를 포함할 수 있다. 피부 톤 센서(18)는 피부의 피부 톤을 표시하는 파라미터를 측정할 수 있고, 파라미터의 측정치들의 시계열을 포함하는 측정 신호를 발생시킬 수 있다. 측정 신호는 헤드 부분(6)이 접촉하는 피부의 피부 톤을 결정하기 위해 프로세싱될 수 있다. 전형적으로 피부 톤 센서는 치료 디바이스(2), 특히 광제모기에서 사용되는데, 이는 광 펄스가 치료되는 피부 유형에 적합한 강도를 갖는지 확인하거나, 또는 피부 유형이 광 펄스에 적합하지 않은 경우(예를 들어, 멜라닌 함량이 훨씬 더 높은 어두운 피부) 광 펄스가 발생하는 것을 방지하기 위해서도 사용된다.
- [0069] 일부 실시예에서, 선택적인 피부 톤 센서(18)는 광 센서일 수 있고, 광 센서에 의해 측정되는 파라미터는 피부로부터 반사되는 특정 파장 또는 다중 파장에서의 광의 강도 또는 레벨일 수 있다. 특정 파장(들)에서 측정된 반사광의 강도 또는 레벨은 피부 톤을 표시할 수 있다. 측정된 반사광의 강도 또는 레벨은 피부 내의 멜라닌 농도에 기초할 수 있으며, 따라서 측정된 강도 또는 레벨은 멜라닌 농도를 표시할 수 있다. 멜라닌 농도는 예를 들어 660nm(적색) 및 880nm(적외선) 파장에서의 광 반사 측정에서 유래할 수 있다.
- [0070] 예시된 치료 디바이스(2)는 또한 치료 디바이스(2)를 활성화시켜 헤드 부분(6)이 대상체의 신체에 대해 요구되

는 치료 동작(예를 들어, 하나 이상의 광원(들)(12)에 의한 하나 이상의 광 펄스의 발생)을 수행하게 하기 위해 사용자에게 의해 동작할 수 있는 사용자 제어부(20)를 포함한다. 사용자 제어부(20)는 스위치, 버튼, 터치 패드 등의 형태일 수 있다.

[0071] 앞서 언급된 바와 같이, 적절한 피드백이 결정될 수 있도록 시간에 따른 치료 디바이스의 위치(배향 포함) 및/또는 이동에 대한 정보를 제공하기 위해 하나 이상의 센서가 선택적으로 사용될 수 있으며, 일부 실시예에서 하나 이상의 센서는 치료 디바이스(2)의 일부이거나 그 안에 있을 수 있다. 센서(들)가 하나 이상의 이동 센서, 예컨대 가속도계, 자이로스코프 등인 경우, 하나 이상의 이동 센서는, 치료 디바이스(2)의 일부, 예를 들어 내부에 있을 수 있다(따라서 도 1에 도시되지 않음). 하나 이상의 센서가 이미징 유닛(예를 들어, 카메라)을 포함하는 실시예에서, 이미징 유닛은 치료 디바이스(2) 상에 또는 그 안에 있을 수 있다. 도 1에 도시된 바와 같이, 치료 디바이스(2)의 하우징(4)에는 이미징 유닛(22)이 배열되어, 치료 디바이스(2)의 전방에 있는 피부 또는 신체 부위의 이미지 또는 비디오 시퀀스를 획득할 수 있다. 즉, 이미징 유닛(22)은 치료 디바이스(2)의 하우징(4) 상에 배열되어, 광원(들)(12)이 조리개(10)를 통해 광 펄스가 방출되는 방향과 일반적으로 평행한 방향으로부터 이미지 또는 비디오 시퀀스를 획득한다. 일부 실시예에서, 이미징 유닛(22)은, 광 펄스가 광원(들)(12)이 조리개(10)를 통해 보이는 피부의 일부에 인가될 수 있도록 헤드 부분(6)이 대상체와 접촉하도록 배치될 때, 이미징 유닛(22)이 조리개(10) 및 헤드 부분(6)과 접촉하는 피부의 일부분을 볼 수 없도록 하우징(4) 상에 배치된다. 이미징 유닛(22)이 치료 디바이스(2)의 하우징(4) 상에 또는 안에 배치되는 실시예에서, 이미징 유닛(22)은 치료 디바이스(2)의 나머지 부분에 대하여 고정된 배열로 되어 있다.

[0072] 치료 디바이스(2)는 광 펄스(플래시)가 생성될 때마다 (플래시 사운드로 또한 지칭되는) 짧은 오디오 펄스를 발생시키기 위해, 라우드스피커와 같은, 오디오 소스(24)(사운드 생성기)를 추가로 포함할 수 있다. 이는 특히 광원(들)(12)이 아크 램프들의 특징적인 펄스 방전 플래시 사운드를 생성하지 않는 LED들 또는 VCSEL들과 같은 반도체 소스일 때의 경우이다. 다른 실시예에서, 플래시 사운드들은 광 펄스가 생성될 때 광원에 의해 (예컨대, 자동으로) 생성될 수 있고; 예를 들어, 광 펄스가 방출될 때마다 톱 사운드와 같은 사운드 펄스가 생성될 수 있다. 플래시 사운드는, 예컨대, 소정 주파수(예컨대, 9 kHz 또는 임의의 다른 선택된 주파수) 또는 주파수 범위(예컨대, 1 내지 10 kHz의 범위)에서 그리고 사용자가 플래시 사운드를 들을 수 있게 하는 강도로 짧은 지속기간(예컨대, 0.1 내지 2초, 예컨대 약 1초)을 갖는 짧은 오디오 펄스이다.

[0073] 도 2는 치료 디바이스(2), 및 치료 디바이스(2)와 함께 사용하기 위한, 예컨대, 도 1에 예를 들어 도시된 바와 같이, 치료 디바이스(2)를 제어하기 위한 장치(42)를 포함하는 본 발명에 따른 예시적인 시스템(40)의 개략도이다. 시스템(40)은, 바람직하게는 장치(42)의 일부일 수 있지만, 또한 외부 요소일 수 있는, 치료 동작 동안 오디오 신호를 감지하도록 구성된 오디오 센서(44)(예컨대, 마이크로폰; 다수의 오디오 센서가 또한 있을 수 있음)를 추가로 포함한다. 추가로, 시스템(40)은, 특히 치료 동작의 진행 또는 사용자가 따라야 하는 임의의 명령어들에 관한, 피드백을 사용자 및/또는 대상체에게 제공하기 위한 피드백 유닛(46)(예컨대, 그래픽 사용자 인터페이스 또는 사운드 출력)을 선택적으로 포함할 수 있다.

[0074] 앞서 언급한 바와 같이, 일부 실시예에서, 장치(42)는 치료 디바이스(2)와 별개의 디바이스일 수 있으며, 따라서 장치(42)는 전자 디바이스의 형태, 예컨대 스마트폰, 스마트 워치, 태블릿, 개인 정보 단말기(PDA), 노트북, 데스크톱 컴퓨터, 원격 서버, 스마트 미러 등 일 수 있다. 다른 실시예에서, 장치(42), 특히 장치(42)에 의해 제공된 발명에 따른 기능성은 치료 디바이스(2)의 일부일 수 있다.

[0075] 장치(42)는 일반적으로 장치(42)의 동작을 제어하고, 장치(42)가 본 명세서에 설명된 방법 및 기술을 수행할 수 있게 하는 프로세싱 유닛(50)을 포함한다. 프로세싱 유닛(50)은 오디오 센서(44)로부터 측정 신호를 수신하고, 측정 신호를 프로세싱하여 치료 동작에 대해 제공할 피드백을 결정하고, 피드백 제어 신호를 피드백 유닛(46)에 제공하도록 추가로 구성될 수 있다. 따라서, 프로세싱 유닛(50)은 센서(44)가 장치(42)의 일부인 실시예에서 직접, 또는 센서(44)가 장치(42)로부터 분리되는 실시예에서 다른 컴포넌트를 통해, 오디오 센서(44)로부터 측정 신호를 수신하도록 구성될 수 있다. 어느 경우에도, 프로세싱 유닛(50)은 입력 유닛(48)을 통해 측정 신호를 수신하기 위한 하나 이상의 입력 포트들 또는 와이어들을 구비하거나 포함할 수 있다. 프로세싱 유닛(50)은 또한 출력 유닛(52)을 통해 피드백 제어 신호를 출력하기 위한 하나 이상의 출력 포트들 또는 와이어들을 구비하거나 포함할 수 있다.

[0076] 프로세싱 유닛(50)은 본 명세서에 설명된 다양한 기능을 수행하기 위해 소프트웨어 및/또는 하드웨어와 함께 수많은 방식으로 구현될 수 있다. 프로세싱 유닛(50)은 필요한 기능을 수행하거나 프로세싱 유닛(50)의 컴포넌트를 제어하여 필요한 기능에 영향을 미치게 하기 위해 소프트웨어 또는 컴퓨터 프로그램 코드를 사용하여 프로그램

래밍될 수 있는 하나 이상의 마이크로프로세서 또는 디지털 신호 프로세서(DSP)를 포함할 수 있다. 프로세싱 유닛(50)은 일부 기능(예를 들어, 증폭기, 전치-증폭기, 아날로그-디지털 변환기(ADC) 및/또는 디지털-아날로그 변환기(DAC))을 수행하기 위한 전용 하드웨어와 다른 기능을 수행하기 위한 프로세서(예를 들어, 하나 이상의 프로그래밍된 마이크로프로세서, 제어기, DSP 및 관련 회로부)의 조합으로 구현될 수 있다. 본 발명의 다양한 실시예에 채용될 수 있는 구성요소의 예는 통상적인 마이크로프로세서, DSP, 주문형 집적 회로(ASIC), 필드 프로그래밍가능 게이트 어레이(FPGA), 뉴럴 네트워크를 구현하기 위한 하드웨어 및/또는 소위 인공 지능(AI) 하드웨어 가속기(즉, 메인 프로세서와 함께 사용될 수 있는 AI 응용을 위해 특별히 설계된 프로세서(들) 또는 다른 하드웨어) 및/또는 하드웨어를 포함하지만, 이로 한정되지 않는다.

[0077] 프로세싱 유닛(50)은 선택적으로 메모리 유닛(도시되지 않음)을 포함하거나 이와 연관될 수 있다. 메모리 유닛은 장치(42)의 동작을 제어하고/하거나 본 명세서에 설명된 방법을 실행 또는 수행함에 있어서 프로세싱 유닛(50)에 의해 사용되기 위한 데이터, 정보 및/또는 신호(이미지(들)를 포함)를 저장할 수 있다. 일부 구현예에서, 메모리 유닛은 프로세싱 유닛(50)이 본 명세서에 기술된 방법들을 포함한 하나 이상의 기능을 수행하도록 프로세싱 유닛(50)에 의해 실행될 수 있는 컴퓨터 판독가능 코드를 저장한다. 특정 실시예에서, 프로그램 코드는 스마트폰, 태블릿, 노트북, 컴퓨터 또는 서버를 위한 애플리케이션의 형태일 수 있다. 메모리 유닛은 임의의 유형의 비일시적 기계 판독가능 매체, 예를 들어 랜덤 액세스 메모리(RAM), 정적 RAM(SRAM), 동적 RAM(DRAM), 판독 전용 메모리(ROM), 프로그래밍가능 ROM(PROM), 소거가능 PROM(EPROM), 및 전기적 소거가능 PROM(EEPROM)과 같은 휘발성 및 비휘발성 컴퓨터 메모리를 포함한 캐시 또는 시스템 메모리를 포함할 수 있고, 메모리 유닛은 메모리 칩; 광 디스크(예를 들어, 콤팩트 디스크(CD), 디지털 다기능 디스크(DVD) 또는 블루-레이 디스크(Blu-Ray disc)); 하드 디스크; 테이프 저장 솔루션; 또는 메모리 스틱, 솔리드 스테이트 드라이브(SSD), 메모리 카드 등을 포함한 솔리드 스테이트 디바이스의 형태로 구현될 수 있다.

[0078] 도 2에 도시된 실시예에서, 장치(42)가 센서(들)(44)와 분리되어 있는 것으로 도시되어 있음에 따라, 장치(42)는 장치(42)가 센서(들)(44)로부터 측정 신호(들)를 수신할 수 있게 하도록, 인터페이스 회로부와 같은 입력 유닛(48)을 포함할 수 있다. 장치(42) 내의 인터페이스 회로부(48)는 센서(들)(44), 치료 디바이스(2), 서버 및 데이터베이스 중 임의의 하나 이상을 포함하는 다른 디바이스와의 데이터 연결 및/또는 데이터 교환을 가능하게 한다. 센서(들)(44)(또는 치료 디바이스(2)와 같은 임의의 전자 디바이스)에 대한 연결은 직접 또는 간접적(예를 들어, 인터넷을 통해)일 수 있으며, 따라서 인터페이스 회로부(48)는 원하는 유선 또는 무선 통신 프로토콜을 통해 장치(42)와 네트워크 사이의 연결, 또는 장치(42)와 다른 장치(예컨대, 센서(44) 및/또는 치료 디바이스(2)) 사이의 직접 연결을 가능하게 할 수 있다. 예를 들어, 인터페이스 회로부(48)는 와이파이(WiFi), 블루투스(Bluetooth), 지그비(Zigbee), 또는 임의의 셀룰러 통신 프로토콜(Global System for Mobile Communications(GSM), Universal Mobile Telecommunications System(UMTS), Long Term Evolution(LTE), LTE-Advanced 등을 포함하지만 이에 제한되지 않음)을 사용하여 동작할 수 있다. 무선 연결의 경우, 인터페이스 회로부(48)(및 따라서 장치(42))는 송신 매체(예를 들어, 공기)를 통해 송신/수신하기 위한 하나 이상의 적합한 안테나를 포함할 수 있다. 대안적으로, 무선 연결의 경우, 인터페이스 회로부(48)는 인터페이스 회로부(48)가 송신 매체(예를 들어, 공기)를 통해 송신/수신하기 위한 장치(42) 외부의 하나 이상의 적합한 안테나에 연결될 수 있게 하는 수단(예를 들어, 커넥터 또는 플러그)을 포함할 수 있다. 인터페이스 회로부(50)는 프로세싱 유닛(50)에 연결된다.

[0079] 추가로, 장치(42)는 장치가 정보를, 특히 피드백 유닛(46)에, 출력할 수 있게 하는 출력 회로부와 같은 출력 유닛(52)을 포함할 수 있다. 출력 유닛은 일반적으로 구성될 수 있고 입력 유닛(48)과 유사하거나 동일한 기능들을 가질 수 있다.

[0080] 피드백 유닛(46)은 장치(42)의 사용자가 (예컨대, 입력 유닛(48)을 통해) 장치(42)에 정보, 데이터 및/또는 커맨드들을 입력할 수 있게 하고/하거나 장치(42)가 (출력 유닛(52)을 통해) 장치(42)의 사용자에게 정보 또는 데이터를 출력할 수 있게 하는 하나 이상의 컴포넌트들을 포함할 수 있다. 사용자 인터페이스는 키보드, 키패드, 하나 이상의 버튼, 스위치 또는 다이얼, 마우스, 트랙 패드, 터치스크린, 스타일러스, 카메라, 마이크로폰 등을 포함하지만 이에 제한되지 않는 임의의 적합한 입력 컴포넌트(들)를 포함할 수 있고, 사용자 인터페이스는 디스플레이 유닛 또는 디스플레이 스크린, 하나 이상의 조명 또는 조명 요소, 하나 이상의 확성기, 진동 요소 등을 포함하지만 이에 제한되지 않는 임의의 적합한 출력 컴포넌트(들)를 포함할 수 있다.

[0081] 일부 실시예에서, 피드백 유닛(46)은 치료 디바이스(2)에 부착되거나 그 일부이다. 다른 실시예에서, 피드백 유닛(46)은 장치(42)의 일부이다. 이러한 실시예에서, 피드백 유닛(46)은 장치(42)의 사용자 인터페이스 컴포넌트의 임의의 출력 컴포넌트의 일부일 수 있고, 또는 이를 사용할 수 있다. 추가 실시예에서, 피드백 유닛

(46)은 치료 디바이스(2) 및 장치(42)로부터 분리되어 있을 수 있다.

- [0082] 장치(42)의 실제적인 구현에는 도 2에 도시된 것들에 추가적인 컴포넌트를 포함할 수 있다는 것이 인식될 것이다. 예를 들어, 장치(42)는 또한 전원 공급장치, 예컨대 배터리, 또는 장치(42)가 주 전원에 연결될 수 있도록 하는 컴포넌트를 포함할 수 있다.
- [0083] 본 발명은 상이한 태양들, 특히 실제 치료 동작 동안 장치(42)를 실제로 사용하는 태양 및 치료 동작 동안 치료 디바이스(2)에 의해 방출된 플래시 사운드들을 신뢰성 있게 인식하도록 장치(42)를 트레이닝시키는 태양을 갖는다.
- [0084] 치료 디바이스의 실제 작동 동안 사용될 수 있는 제1 태양에 따른 방법(100)의 실시예의 흐름도가 도 3에 예시되어 있다. 방법(100)의 제1 단계(101)에서, (오디오 센서(44)에 의해) 치료 동작 동안 감지되는 오디오 신호가 (입력 유닛(48)에 의해) 획득되며, 즉 오디오 센서(44)로부터 수신되거나 추출된다. 획득된 오디오 신호는, 제2 단계(102)에서, 오디오 세그먼트들로, 특히 설정된 또는 사전결정된 지속기간(예컨대, 1초 또는 임의의 다른 적절한 지속기간)의 오디오 세그먼트들로 세그먼트화된다. 다른 (조합가능한) 실시예에서, 상기 사전결정된 지속기간의 오디오 세그먼트들은 치료 동작 동안 입력 유닛(48)에 의해 직접 획득된다. 오디오 세그먼트마다, (아래에서 더 상세히 설명되는 바와 같은) 하나 이상의 특성 특징부들이 제3 단계(103)에서 오디오 세그먼트들로부터 추출된다. 이어서, 제4 단계(104)에서, 오디오 세그먼트마다, 오디오 세그먼트가 플래시 사운드를 포함하는지 여부가 검출된다. 이는 오디오 세그먼트들에서 플래시 사운드들을 검출하기 위해 특성 특징부들 중 하나 이상을 각각 포함하는 복수의 오디오 세그먼트들에 대해 트레이닝되었던 트레이닝된 알고리즘 또는 컴퓨팅 시스템을 적용함으로써 수행된다. 제5 단계(105)에서, 치료 동작의 시작 이후 또는 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 플래시 사운드 카운터는 플래시 사운드가 검출될 때마다 증가된다. 마지막으로, 단계(106)에서, 정보가 출력 유닛(52)을 통해 피드백 유닛(46)으로 출력된다. 정보는 플래시 사운드 카운터(즉, 그의 현재 값), 플래시 사운드 카운터가 플래시 사운드 임계치를 초과하는지 여부의 표시, 및 치료 동작이 완료되었다는 표시 중 하나 이상을 포함할 수 있다.
- [0085] 플래시 사운드 임계치는 예를 들어 사전 설정될 수 있고, 치료될 소정 신체 영역에 대해 요구되거나 적절하거나 처방된 플래시들의 최대 수를 반영할 수 있다. 이러한 설정은 장치(42)에 저장될 수 있거나, 사용자에게 의해 입력될 수 있으며, 사용자는 그것을 예컨대 사용자 매뉴얼 내의 표로부터 취할 수 있거나, 그것을 피드백 유닛 또는 장치(42) 또는 치료 디바이스(2) 상의 사용자 안내에 의한 지원 하에 또는 수동으로 선택 또는 계산할 수 있다. 예를 들어, 플래시 사운드 임계치는 치료 영역의 크기뿐만 아니라 피부의 유형, 피부의 색상, 이전 치료들의 수, 마지막 치료 이후의 시간 등에 의존할 수 있다.
- [0086] 트레이닝을 위해 사용될 수 있는, 제2 태양에 따른 방법(200)의 실시예의 흐름도가 도 4에 예시되어 있다. 방법(200)의 제1 단계(201)에서, 치료 동작 동안 (오디오 센서(44)에 의해) 감지되거나 (예컨대, 외부 디바이스 또는 장치의 사운드 프로세서에 의해) 인위적으로 생성된 오디오 신호가 획득된다. 따라서, 오디오 신호는 그것이 치료 세션 동안 기록되었을 것 같은 "실제" 오디오 신호이거나 이를 시뮬레이션하고, 따라서 다수의 플래시 사운드들(및 선택적으로 배경 노이즈)을 포함한다. 따라서, 그것은 수 초, 수십 초 또는 심지어 몇 분의 길이를 가질 수 있다. 추가로, 오디오 신호는 주석을 포함하고, 이는 오디오 신호를 듣는 동안 사용자에게 의해 생성될 수 있거나 오디오 신호를 인위적으로 생성한 (그리고 동시에 주석을 생성한) 알고리즘 또는 디바이스에 의해 자동으로 생성될 수 있다. 오디오 신호의 주석은 오디오 신호 내의 플래시 사운드들을 나타내는 플래시 사운드 정보, 특히 플래시 사운드의 시작, 종료, 최고 피크, 지속기간 등과 같은 플래시 사운드들의 시간 정보를 포함한다.
- [0087] (주석이 달린) 오디오 신호는, 제2 단계(202)에서, 오디오 세그먼트들로, 특히 설정된 또는 사전 결정된 지속기간의 오디오 세그먼트들로 세그먼트화된다. 제3 단계(203)에서, 오디오 세그먼트마다, 하나 이상의 특성 특징부들, 예컨대, (아래에서 더 상세히 설명되는 바와 같은) 한 세트의 MFCC 계수들이 오디오 세그먼트들로부터 추출된다. 제4 단계(204)에서, 오디오 신호의 주석으로부터, 오디오 세그먼트가 플래시 사운드를 포함하는지 여부를 나타내는 플래시 사운드 정보가 오디오 세그먼트마다 유도되며, 예컨대, 플래시 사운드 정보는 (주석이 달린) 오디오 신호와 연관되거나 그에 내장된 주석을 나타내는 파일 또는 메타데이터로부터 판독된다. 제5 단계(205)에서, 오디오 세그먼트마다, 추출된 하나 이상의 특성 특징부들 및 유도된 플래시 사운드 정보는 알고리즘 또는 컴퓨팅 시스템에 제공되어 하나 이상의 특성 특징부들에 기초하여 오디오 세그먼트 내에서 플래시 사운드의 검출을 위해 알고리즘 또는 컴퓨팅 시스템을 트레이닝시킨다.
- [0088] 몇몇 머신 러닝 및 딥 러닝 아키텍처들이 일반적으로 오디오 분류를 위해 알려져 있다. 하나의 그러한 사용 사

레는 공개 데이터 세트인 UrbanSound8K이다. 이러한 데이터세트는 자동차 경적, 개 짖음, 총 소리, 어린이 놀이, 에어 컨디셔너, 드릴링, 차량 엔진 주행, 잭 해머, 사이렌, 및 거리 음악 등 10개의 클래스들의 도시 사운드의 8732개의 라벨링된 사운드 추출물을 포함한다. 멜-주파수 캡스트럼 계수(MFCC)들, 특징부 추출 및 그 위에 제로 패딩(Zero padding)을 적용하는 것을 사용하여, 이러한 입력은 컨볼루션 뉴럴 네트워크(Convolutional Neural Network, CNN) 모델로 전달된다. MFCC 특징부들을 사용하여, 98.12%의 트레이닝 정확도 및 91.23%의 검증 정확도가 달성된다. 정확도를 개선하기 위해, 제로 패딩보다는 오히려 특징부 패딩이 적용되고, MFCC와 함께 제로 교차율, 피치, RMSE, 크로마(chroma)와 같은 다른 특징부들이 사용된다. 100개의 에포크(epoch)들에 대한 트레이닝 후, 99.21%의 트레이닝 정확도 및 97.29%의 검증 정확도가 달성된다.

[0089] 다른 데이터 세트는 50개의 클래스들로 이루어진 ESC-50이다. 데이터 세트는 5초의 지속기간을 각각 갖는 2000개의 오디오 조각들로 이루어진다. ESC-50에 대한 성능은 94% 정확도이다. 오디오 스펙트로그램 변환기를 사용하여, 컨볼루션-프리 아키텍처가 오디오 분류에 이용된다. 오디오 스펙트로그램 변환기를 사용하여 94.7%의 트레이닝 정확도가 달성된다.

[0090] TUT Rare Sound Events 2017은 각각의 타깃 클래스에 대한 격리된 사운드 이벤트 및 배경으로서 역할을 하는 일상적인 음향 장면들의 기록들로 이루어진다. 타깃 사운드 이벤트 카테고리들은 아기 울음, 유리 깨짐, 총 소리이다. 몇몇 이벤트들이 상기 문제를 해결하기 위해 조직화되었으며, 평가 데이터세트에서 약 93%의 그리고 개발 데이터세트에서 약 96%의 최상의 정확도가 달성되었다. 이러한 문제를 해결하기 위해 제안된 접근법은 하기의 4단계로 나뉜다: 특징부 추출(로그 크기 및 멜 스펙트로그램), 1D convNet에 의한 스펙트럼 특징부의 변환, RNN-LSTM에 의한 시간 의존성, 및 포스트-프로세싱에 의한 오디오 이벤트의 존재 및 시작 시간의 검출.

[0091] 문헌[Uyen Diep (Sally), "Classification of Cough Events", ECE Senior Capstone Project 2017, Tech Notes]은

[0092] 2개의 주요 스테이지들, 즉 이벤트 검출 및 이벤트 분류로 나뉘지는, 환자들을 위한 자동화된 기침 모니터링 시스템의 일부로서 기침 분석을 개시한다. 이러한 방법에서, 사운드 이벤트들은 32ms 길이 또는 16ms 길이의 중첩된 시간 프레임들로 분할된다. 각각의 프레임의 특징부들은 프레임이 기침의 특성을 갖는지 여부를 결정하기 위해 계산되고 분류기로 공급된다. 이어서, 결과들이 조합되어 이벤트에 대한 전반적인 결정을 내린다. MFCC가 이러한 방법에 사용될 수 있다.

[0093] 신호 프로세싱에 기초한 오디오 분류를 위한 공지된 접근법은 주로 딥 뉴럴 네트워크를 이용한 다중클래스 분류이다. 주로 분류되고 있는 사운드들은 몇 초 동안 지속되며 특성이 연속적이다. 대조적으로, 광제모의 현재 시나리오에서의 플래시 검출 시, 플래시들은 대체적으로 수 밀리초로 지속되고 비주기적이다. 일 실시예에서, 플래시가 존재하는지 여부를 검출하기 위해 이진 분류 모델을 구축하기 위한 고전적인 머신 러닝 접근법이 플래시 사운드 검출 및 트레이닝을 위해 사용되고 있다. 딥 러닝에 비해 이러한 머신 러닝(ML) 접근법을 사용하는 것의 중요성은 모델의 앱 배치에 용이한 모델의 최적화된 크기이다.

[0094] 예컨대, 집에서 제모 치료를 하는 동안, 우수한 제모 치료, 적절한 권장 및 마찰을 일으키지 않는 경험에 대한 다수의 문제들이 있다. 특히 치료 디바이스의 사운드 패턴을 검출할 필요가 있다. 사운드 패턴 신호는 치료 디바이스의 사용 동안 방출되는 플래시 사운드들로 이루어지지만 TV 사운드, 라디오, 아기 울음, 교통 노이즈, 사람 소통, 도어가 열리거나 닫힘 등과 같은 잠재적 배경 노이즈가 있을 수 있다. 따라서, 오디오 신호 내에서 의도된 사운드 패턴을 정확하게 식별할 필요가 있다. 본 발명은 임의의 배경 노이즈에도 불구하고 플래시 사운드들을 신뢰성 있고 정확하게 검출하기 위해 적절한 세트의 특성 특징부들(예컨대, 12 내지 13개의 선택된 MFCC 특징부들)과 함께, 신호 프로세싱 및 트레이닝된 알고리즘 또는 컴퓨팅 시스템, 특히 머신 러닝(ML)을 채용한다. 트레이닝된 알고리즘 또는 컴퓨팅 시스템, 예를 들어, ML 모델(예컨대, XGBoost)은 배경 노이즈(비-플래시들)를 플래시 사운드의 것과의 사이에서 구별하기 위해 트레이닝 프로세스에서 학습한다. 따라서, 주석은 모델이 플래시 및 비-플래시 오디오 세그먼트들을 포함하는 패턴을 이해하는 것을 돕는다.

[0095] 따라서, 본 발명은 하기 중 하나 이상을 포함하는 다수의 문제들을 해결한다:

[0096] • 50cm 내지 1m 거리로부터 캡처된 오디오에 기초하여 가변 진폭을 갖는 사운드를 정확하게 검출하는 능력

[0097] • TV, 도어 열림 또는 닫힘, 흐르는 수도물, 라디오 등과 같은 다양한 배경 노이즈들에 대한 강건성

[0098] • (치료 동안 사용되고 있는) 플래시들의 정확한 수의 자동 검출

- [0099] 주석 노력을 감소시키는 접근법
- [0100] 플래시 사운드와 중첩하는 다양한 배경 노이즈들(예컨대, TV, 낙수 등)의 분리
- [0101] 신호들을 프로세싱하고 플래시 사운드만을 검출하기 위한 정확한 검출
- [0102] 종래의 애플리케이션들 및 운영 체제들 둘 모두에 인공 지능(AI) 알고리즘을 통합하기 위한 경량 및 저-레이턴시 아키텍처를 갖는 능력
- [0103] 상기에 인용된 문헌에서 Uyen Diep에 의해 개시된 기침 검출 방법과 같은 종래 기술의 접근법들과 달리, 본 발명은, 플래시 사운드 특성들이 사람 기침들의 특성들과 매우 상이하기 때문에, 배경 노이즈의 존재 시 치료 디바이스에 의해 발생된 (인공) 플래시 사운드들을 신뢰성 있게 검출할 수 있다. 본 발명에 의해 해결되는 과제는 치료 동안 발생하는 플래시 사운드들의 정확한 수를 검출하는 것으로 간주될 수 있다. 플래시 사운드들은 본질적으로 연속적이지 않고 몇 초 동안 지속되며, 즉, 플래시 사운드는 (비주기적 연속 사운드들 및 주기적(단순 또는 복합) 사운드들과 대조적으로) 특성이 비주기적이며 일시적이다. 다리 및 겨드랑이와 같은 대상체의 다양한 신체 영역들의 치료 동안 획득되는 미가공 오디오 신호들은 상이한 지속기간들을 갖는다. 예를 들어, 겨드랑이의 오디오 길이는 30초에서 105초로 다양하며 다리의 오디오 길이는 30초에서 180초로 다양하다. 오디오 신호들은, 예컨대 0.8초의 최소 간격을 갖는 플래시 사운드들을 포함한다. 오디오 신호들은 또한 배경 노이즈를 포함한다.
- [0104] 결과적으로, 본 발명의 프로세싱 접근법은 종래 기술과 상이하다. 예를 들어, 기침과 같은 사운드 이벤트들이 중첩하는 프레임들로 그룹화될 수 있는 한편, 플래시들의 정확한 수를 카운트하기 위해, 본 접근법은 오디오 신호를 프리-프로세싱 동안 중첩하지 않는 오디오 세그먼트들로 분할한다. 중첩하지 않는 오디오 세그먼트들은 플래시들의 이중 카운트를 방지하는 데 특히 유용하다. MFCC 시스템들은 통상 8 내지 13개의 켈스트럼 계수들만을 사용한다. 예컨대, Uyen Diep에 의해 개시된 종래 기술에서 전형적으로 사용되는 0차 계수는 입력 신호의 평균 로그 에너지를 나타내며, 단지 화자-특정 정보를 거의 반송하지 않는다. 이는 그것이 배경 노이즈를 검출하는 것으로 이어질 것이기 때문에 본 발명의 바람직한 실시예들에서 사용되지 않는다. 따라서, 본 발명은 다양한 배경 노이즈와 함께 치료 디바이스에 의해 생성되는 낮은 진폭, 짧은 지속기간의 펄스들을 검출하는 능력을 제공한다. MFCC는 단지 특징부 추출 단계이며, 이러한 특징부들과 적절한 알고리즘/컴퓨팅 시스템(예컨대, 적절한 ML 모델)의 정확한 조합은 원하는 효과들을 달성하고 제기된 문제를 해결하는 데 기여하여, 특히 적절한 특징부 선택으로 인해 미묘한 플래시 사운드를 다양한 배경 노이즈와 구별하는 그의 능력으로 인해 낮은 레이턴시, 낮은 메모리 풋프린트 및 높은 정확도를 제공한다.
- [0105] 예시적인 구현예에서, 데이터(오디오 신호)는 치료 디바이스로부터 가까운 거리(예컨대, 0.5 내지 2.5m 범위)에서 스마트폰의 마이크로폰에 의해 수집된다. 오디오 신호는 임의의 기준의 배경 노이즈와 함께 플래시 사운드들을 포함한다.
- [0106] 트레이닝 동안, 플래시 사운드들이 존재하는 곳이라면 미가공 오디오 신호에 대해 주석이 수행될 수 있으며, 즉, 플래시 사운드들이 오디오 신호에 존재하는 순간들을 나타내는 코멘트 또는 메타데이터와 같은 표시가 오디오 신호에 추가된다. 오디오 신호의 프리-프로세싱은 미가공 오디오 신호 또는 주석이 달린 신호를 설정된 지속기간의 세그먼트들, 예컨대, 1초의 청크(chunk)들로 양자화하고 세그먼트들을 틱(tick)들(즉, 플래시 사운드들) 및 비-틱들(비-플래시 사운드들)로서 라벨링함으로써 수행된다. 추가로, 배경 노이즈 필터링이 수행될 수 있고, MFCC, 멜-스펙트로그램 등과 같은 시간 및 주파수 도메인 특징부들이 오디오 신호로부터 추출될 수 있고, 그에 따라서, 오디오 특징부 어레이를 생성할 수 있다. 일 실시예에서, 거의 실시간으로 플래시 사운드들을 검출하는 능력을 갖는 특징부 어레이를 사용함으로써 ML 모델이 구축된다. ML 모델은, 예컨대, 모델이 스마트폰과 같은 원하는 디바이스 상에서 실행되게 할 수 있는 프로그램 파일로서 앱(애플리케이션 또는 소프트웨어 프로그램)에 배치될 수 있다.
- [0107] 대체적으로 임의의 알고리즘 또는 컴퓨팅 시스템이 사용될 수 있고, 대체적으로 임의의 종류의 러닝 시스템, 머신 러닝, 분류 모델, 이진 분류 모델, 뉴럴 네트워크, 또는 컨볼루션 뉴럴 네트워크가 트레이닝된 알고리즘 또는 컴퓨팅 시스템으로서 사용될 수 있는 한편, 실시예에서 플래시 사운드가 오디오 신호에 존재하는지 여부를 검출하도록 XGBoost 이진 분류 모델이 구축된다. 이러한 XGBoost 분류기를 구축하는 방법은 당업계에 대체적으로 공지되어 있다. 예시적인 XGBoost 분류기는 하기의 하이퍼 파라미터들로 구축될 수 있다: `learning_rate = 0.02`, `n_estimators = 600`, `objective='binary:logistic'`. 다른 파라미터들 및 다른 분류기들 또는 ML 모델들

이 또한 사용될 수 있다. 이러한 맥락에서, 대체적으로, 트레이닝된 알고리즘 또는 컴퓨팅 시스템의 특정 구현에 또는 아키텍처 또는 레이아웃은 필수적이지 않지만, 예를 들어, 알려진 또는 종래의 아키텍처를 갖는 표준 러닝 시스템 또는 뉴럴 네트워크가 플래시 사운드들을 검출하는 작업에 사용될 수 있다는 점에 유의해야 한다.

[0108] (도 4에 예시된 바와 같은) 트레이닝을 위해, 오디오 신호들의 주석들이 사용된다. 이러한 주석들이 어떻게 생성되는지에 대한 실시예가 하기에 설명될 것이다. 전술된 구현예에서, 오디오 데이터는, 예컨대, 실시간으로 또는 임시 저장으로, 스마트폰의 마이크로폰에 의해, 바람직하게는 광제모 디바이스로부터 0.5m 내지 2.5m의 거리에서, 수집된다. 플래시 사운드들은 ML 모델에 의해 추출된다. 추출된 데이터는 오디오 파일 내에 각각의 플래시 사운드의 시작 및 종료 시간을 포함할 수 있다. 대체적으로, 오디오 파일에 주석이 추가될 수 있게 하는 임의의 주석 도구가 이러한 용도로 사용될 수 있다. 예시적인 구현예에서, 본 출원인의 주석 툴 "Barista"가 사용된다.

[0109] 일 실시예에서, 미가공 오디오 파일이 열리고, 피크들이 시각화되고, (예컨대, 단지 한 번만) 오디오가 청취된다. 오디오 파일의 일부의 이러한 시각화는 플래시 사운드들("플래시 톱") 및 배경 노이즈("도어 열림/닫힘 또는 임의의 다른 노이즈")가 표시된 도 5에 도시되어 있다. 이어서, 오디오 파일의 재생이 시작되고, 이는 플래시 사운드가 시각화될 수 있거나 들릴 수 있는 곳에서 일시정지된다. 이어서, 그것은 그러한 영역에서 확대될 수 있다. 도 6은 시작 시간 및 종료 시간이 표시된 2개의 플래시 사운드들에 대한 1초 줌을 갖는 이러한 시각화를 도시한다. 이어서, 주석이 달릴 영역이 선택될 수 있다. 선택된 영역은 교차 검증을 위해 재생되고 청취될 수 있다. 이러한 프로세스는 오디오 파일이 끝날 때까지 반복될 수 있다.

[0110] 일 실시예에서, 도 7에 개략적으로 예시된 바와 같이, 미가공 오디오 파일(301) 및/또는 (플래시 사운드의 표시를 갖는) 주석이 달린 오디오 파일(302)의 오디오 프리-프로세싱(300)이 오디오 파일을 소정 길이(예컨대, 0.1 내지 5초, 예컨대, 0.5 또는 1초의 범위)의 오디오 세그먼트들(청크들)(303)로 양자화함으로써 수행될 수 있다. 이어서, 각각의 오디오 세그먼트는 플래시 사운드 또는 비-플래시 사운드로서 라벨링될 수 있다. 추가로, 프리-프로세싱을 다수 회 반복하는 것, 오디오 파일의 오디오 신호의 피치를 변경하는 것 등과 같은 기본 오디오 조작들(304)이 이루어질 수 있다. 이는, 미가공 오디오 신호가 항상 충분한 또는 최적의 품질을 갖지 않을 수 있어서, 예컨대, 배경 노이즈로 손상될 수 있기 때문에, 수행된다.

[0111] 일 실시예에서, 도 8에 개략적으로 예시된 바와 같이, 오디오 데이터를 분석하기 위해 특징부 추출(400)이 수행된다. 오디오 데이터는 대체적으로 시간, 진폭 및 주파수를 나타내는 3D 신호로 여겨질 수 있다. 3D 포맷은 AI 모델들에 의해 직접 프로세싱될 수 없기 때문에(또는 용이하게 프로세싱되지 않기 때문에), 특징부 추출 접근법이 이용된다. 오디오 신호들을 XGBoost 분류 모델에 전달하기 전에, 신호들은 어레이 포맷으로 변환되는데, 이는 이어서 AI 모델에 의한 추가 프로세싱 및 분석을 위해 준비될 것이다. 대체적으로, 제1 단계(401)에서, 프리엠퍼시스(pre-emphasis), 프레임 블로킹(frame blocking) 및 윈도우잉(windowing)이 수행되고 나서, FFT가 제2 단계(402)에서 적용된다. 이어서, 필터 뱅크(403), 예컨대, 멜-스케일 필터 뱅크, 로그(404) 및 다시 필터 뱅크(405)가 적용되어 원하는 특징부들, 예컨대, MFCC 특징부들(406)을 획득한다.

[0112] MFCC는 피치에 대해 지각적으로(perceptually) 관련된 스케일인 멜 스케일에 관한 것이다. MFCC 특징부로부터, 오디오 세그먼트, 즉 사운드의 조각의 특성을 설명하는 다수의 계수 값들이 획득된다. 일 실시예에서, 제1 단계(401)에서, 주어진 오디오 신호들은 짧은 프레임들(오디오 세그먼트들)로 나뉜다. 이러한 단계는 신호의 주파수들이 시간에 따라 변하기 때문에 수행되고; 따라서, 전체 오디오 신호에 걸쳐 푸리에 변환을 계산하는 것은 시간에 따른 오디오 신호의 주파수 윤곽들의 손실로 이어질 것이다. 추가로, (예컨대, 해밍(Hamming) 윈도우를 사용하는) 윈도우잉은 데이터가 무한하다는 것을 교차 체크하기 위해 그리고 스펙트럼 누설을 감소시키기 위해 수행된다.

[0113] 제2 단계(402)에서, 이산 푸리에 변환은 푸리에 변환의 단순화된 버전을 얻기 위해 그리고 (예컨대, 프레임당 총 39개의 계수들로) 실수-값 계수들을 얻기 위해 계산된다. 제3 단계(403)에서, 멜 이격된 필터 뱅크들이 적용되고, 이는 가장 낮은 그리고 가장 높은 주파수를 멜로 변환하고, 사용될 멜 대역들의 수로서 동일하게 이격된 포인트들을 생성하고, 포인트들을 다시 헤르츠(Hertz)로 변환하고, (이산 신호들이 존재하기 때문에) 이러한 포인트들을 가장 가까운 주파수 빈(bin)들로 라운딩(rounding)하고, 삼각 필터들을 생성하는 것을 포함한다. 결과적으로, MFCC 특징부의 39개의 계수들이 획득된다.

[0114] 대체적으로, 모든 오디오 신호는 다수의 특징부들을 포함한다. 그러나, 해결되어야 하는 문제와 관련이 있는 특성들이 추출될 필요가 있다. 이러한 오디오 특징부들을 사용하여, 지능형 오디오 분류 모델이 구축되고 트레이닝될 수 있다.

[0115] 본 발명의 실시예에서, 멜 주파수 캡스트럼 계수(MFCC)들이 특성들로서 사용된다. 신호의 MFCC들은 스펙트럼 엔벨로프의 전체 형상을 간결하게 설명하는 작은 세트의 특징부들(통상 약 10 내지 20개)이다. MFCC는 사람 음성의 특성들을 모델링한다. 대체적으로, 프레임당 약 39개의 계수들이 있다. 일 실시예에서, 단지 첫 번째 계수들, 예컨대, 입력 신호의 평균 로그 에너지를 나타내는 0차 컴포넌트를 포함하지 않는 처음 12 내지 13개의 계수들(바람직하게는 처음 13개의 계수)만이 사용되는데, 이는 처음 몇몇 계수들이 정보의 대부분(예컨대, 스펙트럼 엔벨로프, 포먼트(formant)들 등)을 인코딩하기 때문이다. MFCC 특징부는 음악 및 음성 프로세싱에서 잘 작동하지만 노이즈에 대한 강건성이 부족하다.

[0116] 하기 표에 표시된 카테고리들에 기초하여 다수의 오디오 카테고리화들이 있다.

오디오 카테고리화	유형들	설명
추상의 수준	높음	기기, 키 코드들, 멜로디, 리듬
	중간	퍼치 및 최상의 관련 기술자들, 예컨대, 노트 오프셋들, MFCC 들
	낮음	에너지, 스펙트럼 도심, 스펙트럼 플럭스, 제로 교차율
시간 범주	즉각 (약 50ms)	오디오 신호에 대한 즉각적인 정보를 제공하는 특징부 이는 약 50 내지 100 ms 를 통상 머무르는 오디오 신호들의 매우 짧은 창크를 고려한다
	세그먼트	세그먼트들에 걸쳐 계산되는 오디오 특징부들은 10 초 또는 15 초부터의 범위에 있다.
	글로벌	전체 오디오 파일에 대한 집계된 정보를 제공하는 특징부
음악 양태		비트, 음색, 퍼치, 하모니 등과 같은, 오디오 특징부들 - 그들이 일별하는 음악적 양태에 기초하여 분류됨
신호 도메인	시간 도메인	파형, x 축 시간, y 축 진폭으로부터 추출됨(제로 교차율, RMSE 에너지, 진폭 엔벨로프)
	주파수 도메인	사운드는 주파수로 특징지어진다. 시간 도메인은 어느 것도 갖지 않는다. 다른 특징부는 주파수 도메인 하에 들어온다(스펙트럼 도심, 스펙트럼 플럭스, 대역 에너지 비)
	시간-주파수 도메인	스펙트로그램, 멜 스펙트로그램, 상수 Q 변환 스펙트로그램 - 시간 도메인으로 시작하여 단시간 푸리에 변환을 적용
ML 접근법	전통적인 ML	제로 교차율, RMSE 에너지, 진폭 엔벨로프, 스펙트럼 도심, 스펙트럼 플럭스, 대역 에너지 비, 스펙트럼 톨 오프, 스펙트럼 확산
	딥 러닝	스펙트로그램, 미가공 오디오

[0117]

[0118] 오디오 특징부 카테고리화는 하기 도메인들에 기초할 수 있다:

신호 도메인	시간 도메인	제로 교차율	$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} 1_{\mathbb{R}_{<0}}(s_t s_{t-1})$
		RMSE 에너지	$RMS(t) = \sqrt{\frac{1}{T} \sum_{t=1}^T w_t(t) x_t^2(t)}$
		진폭 엔벨로프	$AE_t = \max_{k=t-K}^{(t+1) \cdot K-1} s(k)$
	주파수 도메인	스펙트럼 도심	$SC_t = \frac{\sum_{n=1}^N m_t(n) \cdot n}{\sum_{n=1}^N m_t(n)}$
		대역 에너지 비	$B E R_t = \frac{\sum_{n=1}^{K-1} m_t(n)^2}{\sum_{n=K}^N m_t(n)^2}$
	시간-주파수 도메인	스펙트로그램	스펙트로그램은 시간에 따라 변화하는 신호의 주파수 스펙트럼의 시각적 표현이다.
		상수 Q 변환	$\delta f_k = 2^{1/n} \cdot \delta f_{k-1} = (2^{1/n})^k \cdot \delta f_{\min}$

[0119]

[0120] 프리-프로세싱 단계에서 오디오 세그먼트들을 생성한 후, 추출된 MFCC 특징부들은 모델로 넘겨진다.

[0121] 하기에서, AI 모델 트레이닝의 일 실시예가 더 상세히 설명된다. 모델에 대한 입력은, 예컨대, MFCC 계수들로서 13개의 열(column)들 및 오디오 파일의 길이와 등가인 수의 행(row)들을 갖는 매트릭스(또는 벡터)일 수 있다(예컨대, 1초 길이의 각각의 오디오 세그먼트의 경우, 하나의 행이 있음).

[0122] 일 실시예에서, 하기 13개의 MFCC 계수들은 1초 길이의 각각의 오디오 세그먼트의 경우에 사용된다(즉, 이들 13개의 MFCC 계수들은 매트릭스의 13개의 열들을 나타낸다):

[0123] mfcc_second_derivative_std_1

[0124] mfcc_second_derivative_std_2

[0125] mfcc_second_derivative_std_3

[0126] mfcc_second_derivative_std_4

[0127] mfcc_second_derivative_std_5

[0128] mfcc_second_derivative_std_6

[0129] mfcc_second_derivative_std_7

[0130] mfcc_second_derivative_std_8

[0131] mfcc_second_derivative_std_9

[0132] mfcc_second_derivative_std_10

[0133] mfcc_second_derivative_std_l1

[0134] mfcc_second_derivative_std_12

[0135] mfcc_second_derivative_std_13.

[0136] 이들은 대체적으로 당업자에게 공지된 일반적인 명칭들이라는 것에 유의하여야 한다. 이러한 특징부들은 플래시 사운드를 식별하기 위해 그들의 예측력에 기초하여 선택되고 랭크된다. 여기서, 계수 "mfcc_second_derivative_std_1"은 MFCC 2차 도함수 표준 편차(델타로 또한 지칭됨 - 델타 계수)이다. 2차 도함수에 도달하기 위해, MFCC 특징부들이 추출되고, 이어서 그러한 MFCC 특징부들에 대한 2차 도함수가 계산된다.

[0137] 델타 MFCC는 프레임마다 계산된다. 각각의 프레임에 대해, 그것은 현재 MFCC 값들 마이너스 이전 MFCC 프레임 값들이다. 차분 연산이 노이즈에 민감하여, 그것은 평활화 필터를 적용하는 것만큼 우수하여, 더 우수한 정확도로 이어진다. 이어서, 이러한 매트릭스 계수들은 (트레이닝 동안) 타깃 특징부와 함께 모델에 입력된다.

[0138] 1초 길이의 예시적인 오디오 세그먼트에 대한 13개의 계수들에 대한 예시적인 값들은 2.17, 2.14, 1.34, 1.60, 1.94, 1.71, 1.62, 1.67, 1.66, 1.24, 1.48, 1.66, 1.65이다.

[0139] 예시적인 매트릭스는 하기와 같이 보인다(여기서, 처음 13개의 열들은 특징부 벡터(X)이고 마지막 열(Tick)은 타깃 변수(Y)이며, 각각의 행은 상이한 오디오 세그먼트들에 대한 값들을 나타낸다):

[illegible]

[0140]

[0141] 사용된 예시적인 모델은 추출된 MFCC 특징부에 기초하여 예컨대 1초 길이의 각각의 오디오 세그먼트에 플래시 사운드가 존재하는지 여부를 분류하기 위한 XGBoost 분류기이다. 대체적으로, 앙상블 모델은 선형 회귀, 로지스틱 회귀 등과 같은 그의 비앙상블 상대물(counterpart)들보다 회귀 과제들에서의 분류가 더 잘 수행된다. XGBoost는 결정 트리 및 그래디언트 부스팅 프레임워크를 사용하는 앙상블 머신 러닝 알고리즘이다. 이는 다중 회귀, 분류, 랭킹, 및 사용자 정의 예측 문제를 해결하기 위해 사용될 수 있다. 따라서, MFCC 계산의 출력인 매트릭스(오디오 세그먼트들에 대한 값들을 갖는 행들)는 분류를 위해 ML 모델에 공급된다. 트레이닝 동안, 플래시인 것으로 또는 플래시가 아닌 것으로 주석이 달린 각각의 프레임은 타깃 변수로서 특징부 매트릭스와 함께 전달된다. 대체적으로 상이한 계수들(즉, 특성들)이 트레이닝과 실제 사용 사이에서 대체적으로 사용될 수 있는 한편, 동일한 계수들이 디바이스 및 방법의 트레이닝 및 실제 사용에 사용되는 것은 바람직하다.

[0142] 오디오 파일들은 다른 배경 노이즈와 함께 디바이스 팬 노이즈를 갖는 플래시 사운드 마스크들을 포함하는 치료 사운드를 포함한다. 검증 목적으로 AI 모델 및 다른 것들을 트레이닝하는 데 사용될 수 있는 많은 치료들의 수집된 오디오 파일들. 거드랑이의 오디오 길이는 30초에서 105초로 다양하며 다리의 오디오 길이는 30초에서 180초로 다양하다. 트레이닝 정확도 96% 및 검증 정확도 90%는 상기 알고리즘을 사용하여 달성된다.

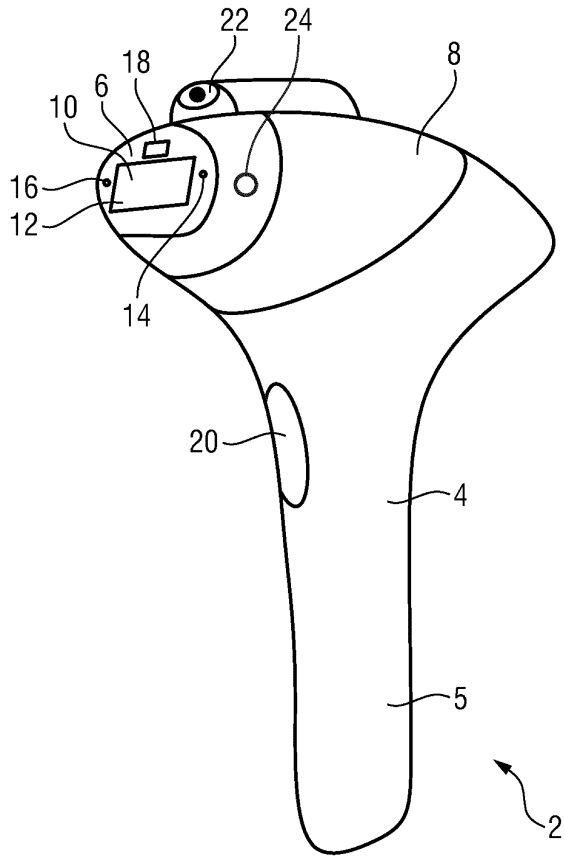
[0143] 플래시 검출 모델은 iOS, 안드로이드 등과 같은 상이한 운영 체제들에 배치될 수 있다. 도 9는 모델 배치부터 생성까지의 프로세스의 흐름도를 예시한다. 기본 모델을 고려하여 새로운 오디오들에 대해 모델을 재트레이닝하는 것은 중요한 단계로서 포함된다. 새로운 데이터 세트들을 사용하여 모델 및 그의 효율성의 미세 튜닝 및 최상의 AI 모델을 저장하는 것이 예측된다. 모델 업데이트는 모델의 연속적인 또는 반복된 재트레이닝 및 그의 최상의 상태로의 미세 튜닝 및 이를 모바일 애플리케이션에 배치하는 것을 포함하는 반복 프로세스이다.

- [0144] 생성 단계에서, 스마트폰(또는 다른 유용한 디바이스)의 마이크로폰을 사용함으로써 실시간 오디오가 입력으로서 취해지고, 이는 최신 모델로 전달되어 플래시들을 검출한다.
- [0145] 사용자 인터페이스를 제공하기 위해 사용될 수 있는, 제3 태양에 따른 방법(500)의 실시예의 흐름도가 도 10에 예시되어 있다. 방법(500)은 대체적으로, 제1 단계(501)에서, 치료 동작에 관한 사용자 입력을 요청하는 제1 스크린을 디스플레이하기 위해 사용자 인터페이스를 제어하도록 구성된다. 제2 단계(502)에서, 사용자 입력이 제1 스크린을 통해 획득되는데, 사용자 입력은 치료될 대상체의 신체 부위의 이미지 및/또는 치료될 대상체의 신체 부위의 크기의 크기 정보를 포함한다. 제3 단계(503)에서, 장치 또는 장치를 포함하는 전자 디바이스와 치료 디바이스의 연결을 나타내는 제2 스크린이 디스플레이된다. 제4 단계(504)에서, 하기 중 적어도 하나를 나타내는 제3 스크린이 디스플레이된다:
- [0146] - 피부 치료 디바이스의 동작 상태,
 - [0147] - 치료 동작의 시작 이후 또는 설정된 또는 사용자 지시된 순간 이후의 검출된 플래시 사운드들의 수를 카운트하는 플래시 사운드 카운터,
 - [0148] - 플래시 사운드 카운터가 플래시 사운드 임계치를 초과하는지 여부의 표시, 및
 - [0149] - 치료 동작이 완료되었다는 표시.
- [0150] 대응하는 사용자 인터페이스가 도 11에서 상이한 단계들로 예시된다. (단계(501)에서 제1 스크린을 나타내는) 스크린(601)은 치료를 효과적으로 수행하는 방법에 대한 사용자 가이드를 보여주고, 원하는 치료 영역의 크기, 예컨대, 다리 길이를 사용자에게 요청할 것이다. 스크린(602)은 (단계 502에서) 일단 사용자가 다리 길이를 입력한 것을 보여주고; 플래시들의 몇몇 이상적인 수가 적절한 치료를 위해 디스플레이될 것이다. 계속을 클릭한 후, 사용자는 스크린(603)에 도달할 것이고, 여기서 스크린은 치료를 시작하기 위해 시작 버튼을 클릭하도록 요청할 것이다. 치료의 시작이 (단계(503)에서 제2 스크린을 나타내는) 스크린(604) 상에서 마이크로폰 액세스를 (또는, 더 일반적으로, 연결 상태를 나타내도록) 요청할 것이다. 일단 그것이 승인되면, 플래시 사운드들의 검출이 스크린(605) 상에 디스플레이된 바와 같이 시작될 것이고, 전체 치료 동안 사용되었던 플래시들의 총 수는 (단계(504)에서 제3 스크린을 나타내는) 스크린(5) 상에 디스플레이되는 바와 같이 증분될 것이다. 다른 실시예에서, 원하는 치료 영역을 치료하는 데 필요한 플래시들의 남은 수가 디스플레이될 수 있거나, 적용된 플래시들의 총 수 및 플래시들의 남은 수 둘 모두가 디스플레이될 수 있다.
- [0151] 도 11에 도시된 스크린들은 예들로서 이해될 것이다. 다른 정보가 각각의 스크린 상에 또한 디스플레이될 수 있으며, 다른 디자인들 및 수들의 스크린들이 있을 수 있다.
- [0152] 본 발명은 사용자가 후속 치료들에 대해 카운트하는 것을 돕는 치료마다 플래시 카운트를 자동화하는 이점을 제공한다. 치료에는 전형적으로 약 30 내지 40개의 플래시들이 소요될 수 있다. 현재, 사용자는 플래시 카운트들을 수동으로 추적하여야 하며, 이는 열악한 사용자 경험으로 이어진다. 따라서, 디바이스의 적절한 권장은 디바이스의 사용 시 마찰을 일으키는 것을 감소시킬 것이고(즉, 사용자가 치료 동안 수동으로 카운트할 필요 없이 플래시들의 카운트를 자동화함), 더 나은 사용자 경험 및 만족으로 이어질 것이다. 제시된 방법은 상이한 배경 노이즈들 및 플래시 사운드를 식별하는 능력을 갖는 모델들로 시험되었다. 치료가 기대 수준을 충족하지 못할 때마다, 안내 또는 권장이 준수되었는지 여부가 확인될 수 있다.
- [0153] 본 발명은 도면 및 기술한 설명에서 자세히 예시되고 설명되었지만, 이러한 예시 및 설명은 예시적이거나 본보기인 것으로 간주되어야 하며 제한적이지 않다; 본 발명은 개시된 실시예에 제한되지 않는다. 개시된 실시예에 대한 다른 변형은 도면, 개시 및 첨부된 청구범위의 연구로부터 청구된 발명을 실시함에 있어서 당업자에 의해 이해되고 효력을 발휘할 수 있다.
- [0154] 청구항들에서, 단어 "포함하는"은 다른 요소들 또는 단계들을 배제하지 않으며, 단수 형태(부정 관사 "a" 또는 "an")는 복수를 배제하지 않는다. 단일 요소 또는 다른 유닛은 청구항에서 인용되는 여러 항목의 기능을 수행할 수 있다. 소정의 수단이 상호 상이한 종속항에 열거된다는 단순한 사실은 이들 수단의 조합이 유리하게 사용될 수 없다는 것을 나타내지는 않는다.
- [0155] 컴퓨터 프로그램은 적합한 비일시적 매체, 예컨대 광학 저장 매체 또는 다른 하드웨어와 함께 또는 다른 하드웨어의 일부로서 공급되는 솔리드 스테이트 매체 상에 저장/배포될 수 있지만, 인터넷이나 다른 유선 또는 무선 통신 시스템을 통해 다른 형태로 배포될 수도 있다.

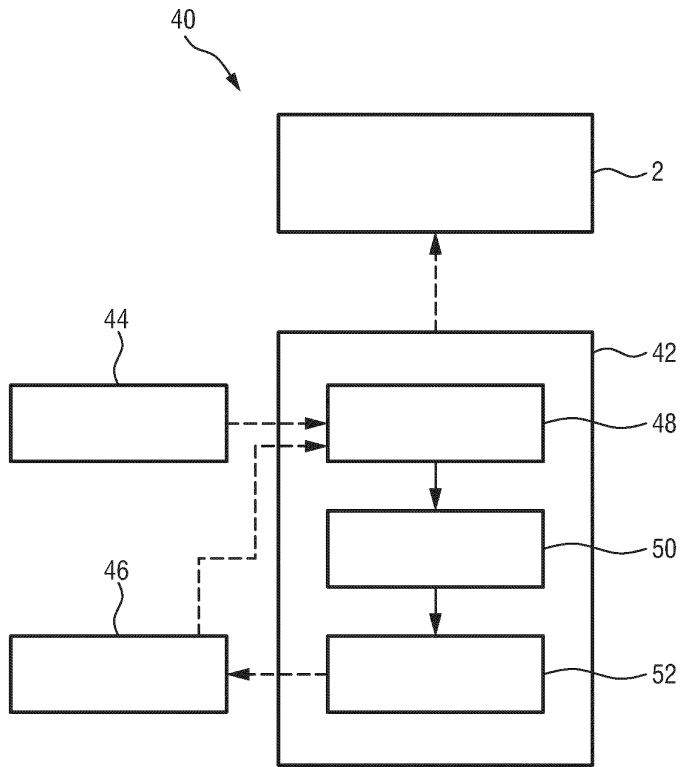
[0156] 청구항에서의 임의의 도면 부호는 범위를 제한하는 것으로 해석되어서는 안 된다.

도면

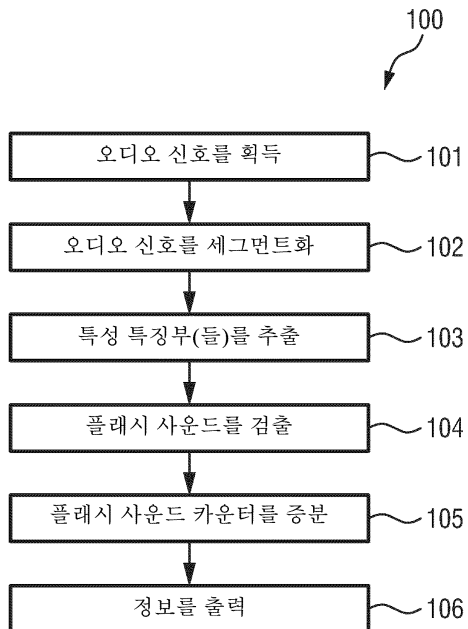
도면1



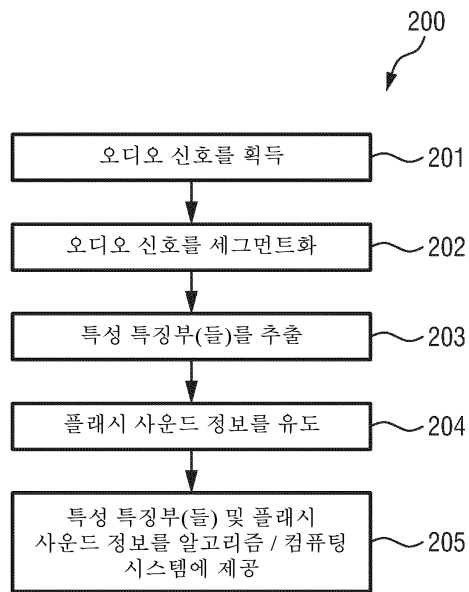
도면2



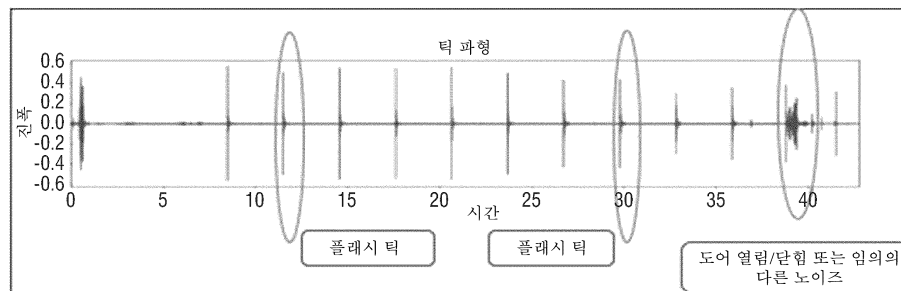
도면3



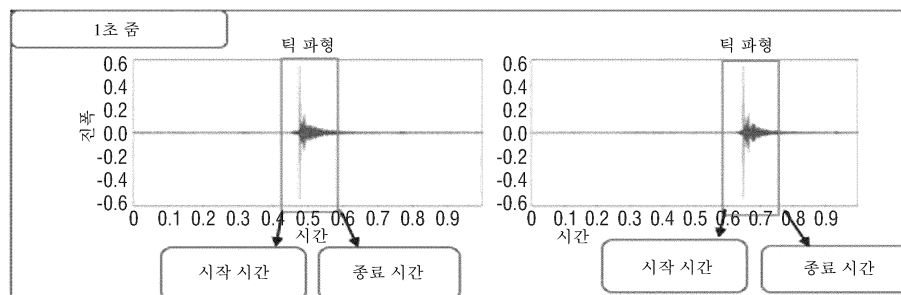
도면4



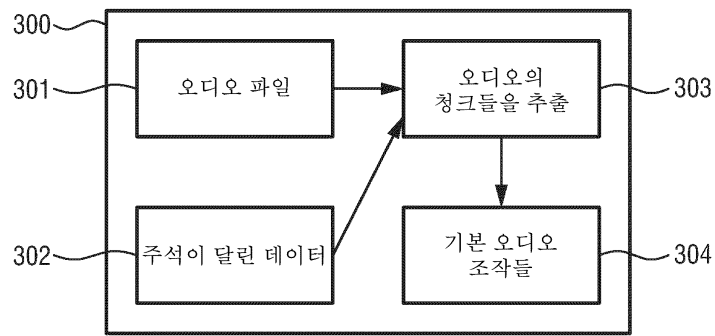
도면5



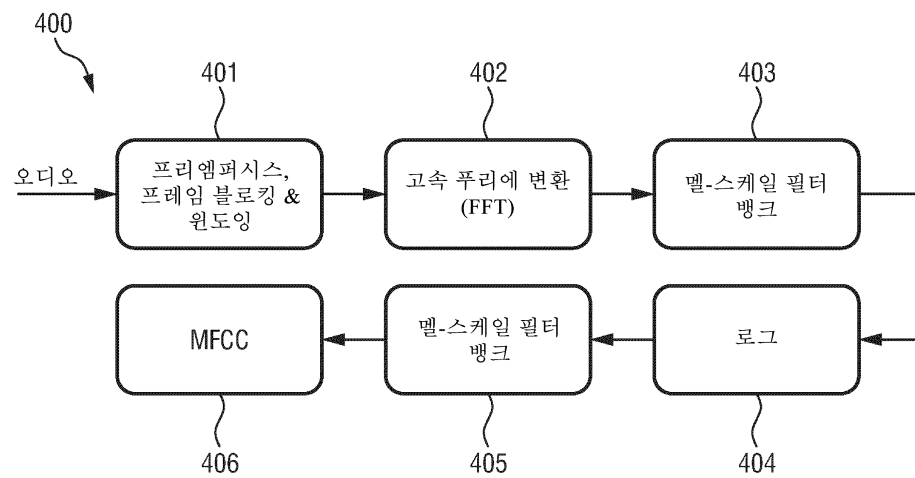
도면6



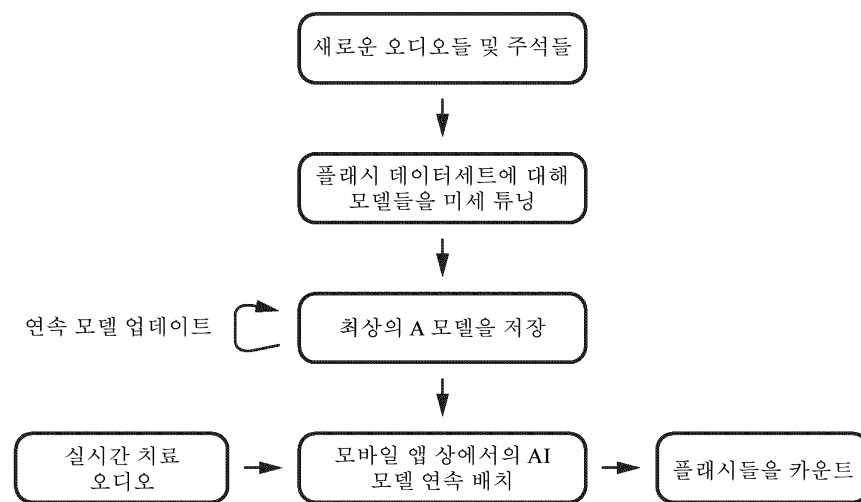
도면7



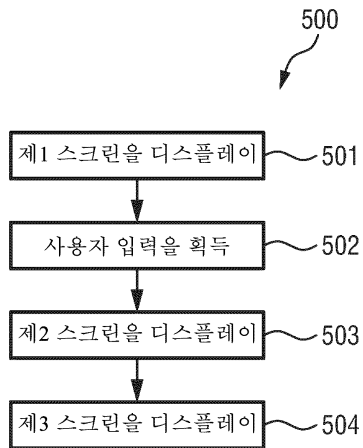
도면8



도면9



도면10



도면11

