



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0113968
(43) 공개일자 2019년10월08일

(51) 국제특허분류(Int. Cl.)
A61B 5/00 (2006.01) A61B 5/02 (2006.01)
G10L 17/26 (2013.01) G10L 25/66 (2013.01)
(52) CPC특허분류
A61B 5/7275 (2013.01)
A61B 5/0022 (2013.01)
(21) 출원번호 10-2019-7026858
(22) 출원일자(국제) 2018년02월12일
심사청구일자 없음
(85) 번역문제출일자 2019년09월11일
(86) 국제출원번호 PCT/IL2018/050162
(87) 국제공개번호 WO 2018/146690
국제공개일자 2018년08월16일
(30) 우선권주장
62/457,914 2017년02월12일 미국(US)

(71) 출원인
카디오콜 엘티디.
이스라엘 7178094 모디인 스테룻 하미크트조웃 13
(72) 발명자
하우프먼, 일미야후
이스라엘 7528653 리손-레지온 아파트 2 유샤 스트리트 14
고렌, 알론
이스라엘 7311500 모샤브 벤-쉬맨 하틴아 스트리트 154
(뒷면에 계속)
(74) 대리인
특허법인 무한

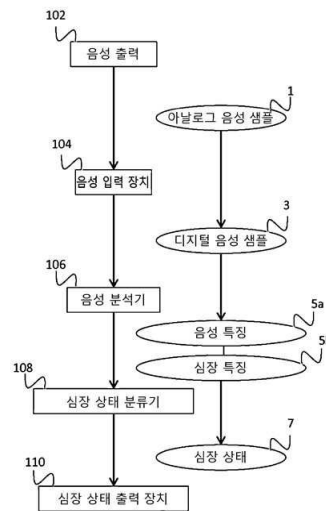
전체 청구항 수 : 총 50 항

(54) 발명의 명칭 심장병에 대한 언어적 정기 검사

(57) 요약

심장 상태, 예를 들어 부정맥을 결정하기 위해 대상의 음성 샘플을 분석하는 방법이 개시되며, 상기 방법은 상기 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 단계, 상기 적어도 하나의 음성 특징에 대한 상기 심장 상태의 영향을 검출하는 단계, 및 상기 영향에 기초하여 심장 상태를 결정하는 단계를 포함한다. 또한, 관련 장치 및 방법이 개시된다.

대표도 - 도1a



(52) CPC특허분류

A61B 5/02 (2013.01)

A61B 5/4803 (2013.01)

A61B 5/7264 (2013.01)

G10L 17/26 (2013.01)

G10L 25/66 (2013.01)

(72) 발명자

아타르, 엘리

이스라엘 4858706 로쉬 하야인 레이한 스트리트 6

사박, 핀하스

이스라엘 9641328 예루살렘 아파트 1 이즈착 샬이
브 스트리트 7

명세서

청구범위

청구항 1

부정맥 심장 상태(arrhythmic cardiac condition)를 결정하기 위해 대상(subject)의 음성 샘플(voice sample)을 분석하는 방법에 있어서,

상기 음성 샘플로부터 적어도 하나의 음성 특징(voice feature)을 추출하는 단계;

상기 적어도 하나의 음성 특징에 대한 심장 상태(cardiac condition)의 영향(effect)을 검출하는 단계;

상기 영향에 기초하여 부정맥 심장 상태를 결정하는 단계

를 포함하는 음성 샘플 분석 방법.

청구항 2

제1항에 있어서,

상기 적어도 하나의 음성 특징에 대한 심장 상태의 영향을 검출하는 단계는,

상기 적어도 하나의 음성 특징의

혈류(blood flow);

심실 박동(ventricular beat); 및

심방 박동(atrial beat)

을 포함하는 그룹으로부터 선택된 적어도 하나의 음성에 영향을 미치는 원인(cause affecting voice)에 의해 영향을 받는 심장 상태의 영향을 검출하는 단계

를 포함하는 음성 샘플 분석 방법.

청구항 3

제2항에 있어서,

상기 영향을 검출하는 단계는,

상기 음성에 영향을 미치는 원인의 타이밍(timing), 상기 음성에 영향을 미치는 원인의 주기(periodicity), 및 상기 음성에 영향을 미치는 원인의 크기(magnitude) 또는 크기의 변화(change in magnitude) 중 적어도 하나의 영향을 식별하는 단계

를 포함하는 음성 샘플 분석 방법.

청구항 4

제1항에 있어서,

상기 부정맥 심장 상태를 결정하는 단계는,

복수의 상기 식별된 영향을 통합하는(integrating) 단계

를 포함하는 음성 샘플 분석 방법.

청구항 5

제1항에 있어서,

유성 세그먼트(voiced segments)를 식별하기 위해 상기 음성 샘플을 분류하는(classifying) 단계
를 더 포함하는 음성 샘플 분석 방법.

청구항 6

제5항에 있어서,

적어도 일부 세그먼트가 제거되고 나머지 세그먼트 엔드(segment ends)가 평탄화 되도록(smoothed) 상기 유성
세그먼트를 연결하는(concatenating) 단계
를 더 포함하는 음성 샘플 분석 방법.

청구항 7

제2항에 있어서,

상기 음성 샘플로부터 제거된 세그먼트에서 상기 음성에 영향을 미치는 원인에 관련되는(pertaining) 데이터를
외삽하는(extrapolating) 단계
를 더 포함하는 음성 샘플 분석 방법.

청구항 8

제1항에 있어서,

상기 검출된 영향은,
심장 펄스(heart pulses)의 타이밍
을 포함하는 음성 샘플 분석 방법.

청구항 9

제8항에 있어서,

상기 영향을 검출하는 단계는,
세 번의 연속적인 심장 펄스의 발생(three consecutive occurrences of heart pulses)을 식별하는 단계
를 포함하는 음성 샘플 분석 방법.

청구항 10

제9항에 있어서,

상기 부정맥 심장 상태를 결정하는 단계는,
상기 연속적인 심장 펄스의 발생 사이의 시간 간격(time interval)을 계산하는 단계 및 기준 심장 상태

(reference heart condition)로부터 획득된 기준 시간 간격(reference time interval)과 상기 시간 간격을 매칭하는 단계

를 포함하는 음성 샘플 분석 방법.

청구항 11

제10항에 있어서,

상기 결정하는 단계는,

상기 매칭의 확률을 결정하는 단계; 및

상기 확률을 상기 부정맥 심장 상태를 결정하기 위한 임계값과 비교하는 단계

를 더 포함하는 음성 샘플 분석 방법.

청구항 12

제11항에 있어서,

상기 기준 심장 상태는,

건강이거나(healthy) or 부정맥(arrhythmic)인 음성 샘플 분석 방법.

청구항 13

제9항에 있어서,

상기 세 번의 심장 펄스의 발생(three occurrences of a heart pulse)은,

동일한 유성 세그먼트에서 식별되지 않은 것인 음성 샘플 분석 방법.

청구항 14

제1항에 있어서,

상기 영향을 검출하는 단계는,

두 번의 연속적인 심장 펄스의 발생을 식별하는 단계

를 포함하는 음성 샘플 분석 방법.

청구항 15

제1항에 있어서,

상기 검출된 영향은,

심장 펄스의 크기 또는 크기의 변화의 영향

을 포함하는 음성 샘플 분석 방법.

청구항 16

제15항에 있어서,

상기 영향을 검출하는 단계는,
상기 적어도 하나의 음성 특징의 복수의 값의 분포를 계산하는 단계
를 포함하는 음성 샘플 분석 방법.

청구항 17

제16항에 있어서,
상기 부정맥 심장 상태를 결정하는 단계는,
상기 분포의 특성화 파라미터(characterizing parameter)를 임계값과 비교하는 단계
를 포함하는 음성 샘플 분석 방법.

청구항 18

제17항에 있어서,
상기 특성화 파라미터는,
상기 분포의 형태(shape of the distribution)
를 포함하는 음성 샘플 분석 방법.

청구항 19

제18항에 있어서,
상기 값은,
상기 형태의 폭(width of the shape)
를 포함하는 음성 샘플 분석 방법.

청구항 20

제1항에 있어서,
상기 검출된 영향은,
심장 펄스의 주기성(periodicity)의 영향
을 포함하는 음성 샘플 분석 방법.

청구항 21

제20항에 있어서,
상기 대상의 심박수(heart rate)를 추정하는 단계
를 더 포함하는 음성 샘플 분석 방법.

청구항 22

제21항에 있어서,

상기 부정맥 심장 상태를 결정하는 단계는,

상기 심박수의 주파수 주위로 미리 결정된 범위(predetermined range around a frequency of the heart rate)에서 상기 주기성을 특성화하는 단계를 포함하고,

상기 특성화하는 단계는,

상기 미리 결정된 범위에서 피크의 대역폭(band width of a peak)을 계산하는 단계를 포함하고,

상기 부정맥 심장 상태를 결정하는 단계는,

상기 대역폭을 임계값과 비교하는 단계를 포함하는 음성 샘플 분석 방법.

청구항 23

제1항에 있어서,

상기 부정맥 심장 상태를 결정하는 단계를 검증하는 단계

를 더 포함하는 음성 샘플 분석 방법.

청구항 24

제23항에 있어서,

상기 검증하는 단계는,

미리 결정된 발성(predetermined vocalization)을 포함하는 제2 음성 샘플을 획득하는 단계

를 포함하는 음성 샘플 분석 방법.

청구항 25

제23항에 있어서,

상기 검증하는 단계는,

상기 대상에 대한 심전도 검사(electrocardiogram test) 및 광혈류측정 검사(photoplethysmography test) 중 적어도 하나를 포함하는 음성 샘플 분석 방법.

청구항 26

제1항에 있어서,

상기 부정맥 심장 상태는,

심방 세동(atrial fibrillation)을 포함하는 음성 샘플 분석 방법.

청구항 27

제1항에 있어서,

상기 음성 샘플은,

자발적인 스피치(spontaneous speech)인 음성 샘플 분석 방법.

청구항 28

제1항에 있어서,
상기 음성 특징은,
피치(pitch)를 포함하는 음성 샘플 분석 방법.

청구항 29

제1항에 있어서,
상기 음성 특징을 추출하는 단계는,
상기 음성 샘플의 제1 세그먼트 및 상기 음성 샘플의 제2 세그먼트 사이 교차 일관성을 수행하는(performing cross coherence) 단계
를 포함하는 음성 샘플 분석 방법.

청구항 30

제1항에 있어서,
상기 음성 특징을 추출하는 단계는,
웨이블릿 분석(wavelet analysis);
Mel frequency Cepstral Coefficient(MFCC) 분석;
성문 펄스 분석(glottal pulse analysis); 및
선형 예측 코딩(Linear Predictive Coding(LPC)) 분석
을 포함하는 그룹으로부터 선택된 피치 시그널(speech signal)의 분석을 수행하는 단계
를 포함하는 음성 분석 방법.

청구항 31

대상에 의해 제공된 음성 샘플에서 부정맥 심장 상태를 결정하기 위한 시스템에 있어서,
상기 음성 샘플의 디지털 음성 샘플을 획득하기 위한 음성 입력 장치(voice input);
상기 디지털 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 음성 특징 추출기(voice feature extractor);
상기 적어도 하나의 음성 특징에 대한 심장 활동(cardiac activity)의 영향을 식별하는 음성 특징 프로세서(voice feature processor);
상기 영향에 기초하여 부정맥 심장 상태를 결정하는 심장 상태 분류기(cardiac condition classifier)
를 포함하는 음성 분석 시스템.

청구항 32

제31항에 있어서,

상기 음성 입력 장치는 휴대폰(mobile phone), 디지털 어시스턴트(digital assistant), 및 차량 멀티미디어 시스템(car multimedia system) 중 적어도 하나의 일부를 구성하는 음성 분석 시스템.

청구항 33

제31항에 있어서,

상기 음성 특징 추출기, 상기 음성 특징 프로세서, 및 상기 심장 상태 분류기 중 적어도 하나는, 서버에 위치한 음성 분석 시스템.

청구항 34

제33항에 있어서,

상기 서버는 전화 스위치와 함께 배치되어 상기 스위치로부터 데이터를 얻는 음성 분석 시스템.

청구항 35

제31항에 있어서,

상기 음성 입력 장치는,

미리 결정된 스케줄에 따라 상기 디지털 음성 샘플을 획득하는 음성 분석 시스템.

청구항 36

제31항에 있어서,

영구 메모리(permanent memory)를 더 포함하고,

상기 메모리는 부정맥 심장 상태 또는 건강한 심장 상태와 관련된 복수의 음성 샘플로부터 유도된(derived) 기준 음성 특징을 저장하는

음성 분석 시스템.

청구항 37

제36항에 있어서,

상기 심장 상태 분류기는,

상기 기준 음성 특징을 특성화하는 훈련 단계(training stage characterizing)에 기초하여 상기 부정맥 심장 상태를 결정하는

음성 분석 시스템.

청구항 38

제36항에 있어서,

상기 영구 메모리는,

상기 대상으로부터 미리 수집된 적어도 하나의 음성 샘플을 저장하는 음성 분석 시스템.

청구항 39

제38항에 있어서,
상기 심장 상태 분류기는,
상기 미리 수집된 음성 샘플 및 상기 디지털 음성 샘플 사이 변화에 기초하여 상기 부정맥 심장 상태를 결정하는 음성 분석 시스템.

청구항 40

대상에 의해 제공된 음성 샘플에서 부정맥 심장 상태를 결정하기 위한 시스템에 있어서,
상기 음성 샘플의 디지털 음성 샘플을 획득하기 위한 음성 입력 장치(voice input);
상기 디지털 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 음성 특징 추출기(voice feature extractor);
상기 적어도 하나의 음성 특징에 대한 심장 상태(cardiac condition)의 영향을 식별하는 음성 특징 프로세서(voice feature processor);
상기 영향에 기초하여 상기 심장 상태를 결정하는 심장 상태 분류기(cardiac condition classifier)
를 포함하는 음성 분석 시스템.

청구항 41

심장 상태를 결정하기 위해 대상의 음성 샘플을 분석하는 방법에 있어서,
상기 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 단계;
상기 적어도 하나의 음성 특징에 대한 혈류의 영향을 검출하는 단계;
상기 영향에 기초하여 심장 상태를 결정하는 단계
를 포함하고,
상기 영향은 상기 혈류의 타이밍, 상기 혈류의 주기성, 및 상기 혈류의 크기 또는 상기 혈류의 크기의 변화 중 적어도 하나인 음성 샘플 분석 방법.

청구항 42

제41항에 있어서,
상기 결정된 심장 상태는,
비정상적인 심박수(abnormal heart rate)를 포함하는 음성 샘플 분석 방법.

청구항 43

제42항에 있어서,
상기 비정상적인 심박수는,
심실 빈맥(ventricle tachycardia)을 포함하는 음성 샘플 분석 방법.

청구항 44

제41항에 있어서,
상기 결정된 심장 상태는,
이소성 박동(ectopic beats) 및 조기 심실 수축(premature ventricular contraction) 중 적어도 하나를 포함하는 음성 샘플 분석 방법.

청구항 45

심장 행동(cardiac behavior)를 재구성(reconstruct)하기 위해 대상의 음성 샘플을 분석하는 방법에 있어서,
상기 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 단계;
상기 적어도 하나의 음성 특징에 대한 혈류의 영향을 검출하는 단계; 및
상기 검출된 영향으로부터 상기 심장 행동의 하나 이상의 파라미터를 재구성하는 단계를 포함하는 음성 샘플 분석 방법.

청구항 46

제45항에 있어서,
상기 심장 행동은,
심장 세동(cardiac fibrillation)의 발생을 포함하는 음성 샘플 분석 방법.

청구항 47

제46항에 있어서,
미리 결정된 횟수보다 많은 심장 세동의 발생을 식별할 때 부정맥 심장 상태의 존재를 결정하는 단계를 포함하는 음성 샘플 분석 방법.

청구항 48

제46항에 있어서,
미리 결정된 횟수보다 적은 심장 세동의 발생을 식별할 때 부정맥 심장 상태의 존재를 배제하는 단계를 포함하는 음성 샘플 분석 방법.

청구항 49

제45항에 있어서,
상기 심장 행동의 하나 이상의 파라미터는,

비 부정맥 심장 상태(non-arrhythmic cardiac condition)를 나타내는 음성 샘플 분석 방법.

청구항 50

제45항에 있어서,

상기 하나 이상의 파라미터는,

펄스 웨이브 형태(pulse wave shape)을 포함하는 음성 샘플 분석 방법.

발명의 설명

기술 분야

[0001] 아래 실시예들은 심장병에 대한 언어적 정기 검사에 관한 것이다.

배경 기술

[0002] 본 발명은 일부 실시 예에서 음성을 이용한 의료 진단(medical diagnosis)에 관한 것으로, 이에 한정되지는 않지만, 보다 상세하게는 인간 음성(human voice)에 기초한 심장 상태 추정(estimating cardiac conditions)에 관한 것이다.

[0003] 미국특허출원공개공보 US2012/0220899는 사용자의 건강 상태(health state)를 진단하기 위한 방법 및 장치를 개시한다. 상기 장치는 사용자의 음성(voice)를 검출하고 모니터링하는 음성 검출 유닛(voice detecting unit); 진단될 건강 상태에 기초하여, 상기 음성 검출 유닛에 의해 검출된 음성으로부터 음성 특징(voice feature)를 추출하는 음성 분석 유닛(voice analyzing unit); 상기 음성 분석 유닛에서 추출된 음성 특징과 비정상 상태(abnormal state)의 기준을 비교하여 사용자의 건강 상태를 진단하고, 상기 건강 상태의 변화를 모니터링하는 음성 진단 유닛(voice diagnosing unit); 및 상기 음성 진단 유닛에서 진단된 건강 상태 및 건강 상태 변화에 관한 정보를 출력하는 진단 출력 유닛(diagnosis outputting unit)을 포함한다.

[0004] 미국특허출원공개공보 US2014/0122063는 일련의 발성 파라미터(series of parameters of phonation)이 계산될 수 있는 음성을 녹음 및 분석하기 위한 방법 및 컴퓨팅 시스템을 개시한다. 이들은 유기 장애(organic disorders, 후두의 생리학(physiology of the larynx)에 영향을 미침) 또는 신경계 장애(neurological disorders, 뇌의 언어 능력 담당 영역(cerebral centers of speech)에 영향을 미침)에 의해 야기된 영향에 관한 관련 정보를 전달한다. 또한, 이 분류 방법은 기존의 기능 장애의 추정이 얻어지고 성격의 할당을 가능하도록 하는 이 발명의 필수 부분으로 간주된다(The classification methods are also considered an essential part of the invention which allow estimations of the existing dysfunction to be obtained and for the allocation of personality).

[0005] 추가 배경 기술은 다음을 포함한다:

[0006] An article titled: "Pulse Estimation - A Frequency Domain Approach" by Sandra Dias and Anibal Ferreira, of the Department of Electrical and Computer Engineering, University of Porto - Faculty of Engineering, Porto, Portugal, published on the World Wide Web in www.ortra.com/events/Portals/16/Dias%20et%20al_AfekaConfSP2014.pdf;

[0007] a Master's Thesis submitted to the School of Engineering of the University of Porto titled: "of the glottal pulse from speech or singing voice" by Sandra de Oliveira Dias, in July 2012;

[0008] an article titled "Wavelet Transform Decomposition Level Determination Exploiting Sparseness Measurement" by Lei Lei, Chao Wang, and Xin Liu, published in International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering Vol:7, No:9, 2013;

[0009] A PhD thesis titled "in Glottal Analysis and its Applications" by Thomas Drugmore to obtain the title of PhD in Applied Sciences of University of Mons;

- [0010] U.S. Patent No. US 8,078,470;
- [0011] U.S. Patent Application Publication No. US 2008/0045805;
- [0012] U.S. Patent No. US 7,917,366;
- [0013] U.S. Patent No. US 8,249,875;
- [0014] U.S. Patent No. US 8,768,708;
- [0015] International Application Publication No. WO 2016/035070;
- [0016] International Application Publication No. WO 2014/037937; 및
- [0017] U.S. Patent No. US 7,398,213.

발명의 내용

해결하려는 과제

과제의 해결 수단

- [0018] 예 1. 부정맥 심장 상태(arrhythmic cardiac condition)를 결정하기 위해 대상(subject)의 음성 샘플(voice sample)을 분석하는 방법에 있어서,
- [0019] 상기 음성 샘플로부터 적어도 하나의 음성 특징(voice feature)을 추출하는(extracting) 단계;
- [0020] 상기 적어도 하나의 음성 특징에 대한 혈류(blood flow)의 영향(effect)을 검출하는(detecting) 단계;
- [0021] 상기 영향에 기초하여 부정맥 심장 상태를 결정하는(determining) 단계
- [0022] 를 포함하는 방법.
- [0023] 예 2. 예 1에 있어서, 상기 부정맥 심장 상태를 결정하는 단계는, 복수의 식별된 영향(a plurality of the identified effect)을 통합하는(integrating) 단계를 포함하는 방법.
- [0024] 예 3. 예 1 내지 예 2 중 어느 하나에 있어서, 상기 영향을 식별하는 단계는, 상기 혈류의 타이밍(timing), 상기 혈류의 주기성(periodicity), 및 상기 혈류의 크기(magnitude) 또는 크기의 변화(change in magnitude) 중 적어도 하나의 영향을 식별하는 단계를 포함하는 방법.
- [0025] 예 4. 예 1에 있어서, 유성 세그먼트(voiced segments)를 식별하기 위해 상기 음성 샘플을 분류하는 단계를 더 포함하는 방법.
- [0026] 예 5. 예 4에 있어서, 적어도 일부 세그먼트가 제거 되고 나머지 세그먼트 엔드(segment ends)가 평탄화 되도록(smoothed) 상기 유성 세그먼트를 연결하는(concatenating) 단계를 더 포함하는 방법.
- [0027] 예 6. 예 1 내지 예 5 중 어느 하나에 있어서, 상기 음성 샘플로부터 제거된 세그먼트에서 상기 혈류에 관련되는(pertaining) 데이터를 외삽하는(extrapolating) 단계를 더 포함하는 방법.
- [0028] 예 7. 예 1 내지 예 6 중 어느 하나에 있어서, 상기 검출된 영향은 심장 펄스(heart pulses)의 타이밍을 포함하는 방법.
- [0029] 예 8. 예 7에 있어서, 상기 영향을 검출하는 단계는 적어도 세 번의 연속적인 심장 펄스의 발생(three consecutive occurrences of heart pulses)을 식별하는 단계를 포함하는 방법.
- [0030] 일부 실시예에서, 세 번의 연속적인 심장 펄스의 발생을 식별함으로써 영향을 검출하는 것은 세 번의 연속적인 심장 펄스의 발생과 대략적으로 동등한 시간, 또는 대략 2초 내지 3초가 잠재적으로 소요된다. 2초 내지 3초 안에 심장 상태의 영향을 검출하는 것은, 일반적으로 설정(setup; 예를 들어, 전극 배치)에만 2초 내지 3초 이상 필요한 ECG에 의한 검출과 비교된다.
- [0031] 예 9. 예 8에 있어서, 상기 부정맥 심장 상태를 결정하는 단계는 상기 적어도 세 번의 연속적인 심장 펄스의 발생 사이의 시간 간격(time intervals)을 계산하는 단계를 포함하는 방법.

- [0032] 예 10. 예 9에 있어서, 부정맥 심장 상태를 결정하는 단계는 상기 시간 간격을 기준 심장 상태(reference heart condition)으로부터 획득된 기준 시간 간격(reference time intervals)과 매칭시키는 단계를 포함하는 방법.
- [0033] 예 11. 예 10에 있어서, 상기 결정 단계는,
- [0034] 상기 매칭의 확률을 결정하는 단계; 및
- [0035] 상기 확률을 상기 부정맥 심장 상태를 결정하기 위한 임계값과 비교하는 단계를 포함하는 방법.
- [0036] 예 12. 예 11에 있어서, 기준 심장 상태가 건강(healthy)이거나 부정맥(arrhythmic)인 방법.
- [0037] 예 13. 예 7 내지 예 12 중 어느 하나에 있어서, 상기 적어도 세 번의 연속적인 심장 펄스의 발생은 동일한 유성 세그먼트에서 식별되지 않은 것인 방법.
- [0038] 예 14. 예 7 내지 예 13 중 어느 하나에 있어서, 상기 영향을 검출하는 단계는 적어도 두 번의 연속적인 심장 펄스 발생을 식별하는 단계를 포함하는 방법.
- [0039] 일부 실시예에서 두 번의 연속적인 심장 펄스의 발생을 식별함으로써 영향을 검출하는 것은 두 번의 연속적인 심장 펄스의 발생과 대략적으로 동등한 시간, 또는 대략 2초가 잠재적으로 소요된다.
- [0040] 예 15. 예 1 내지 14 중 어느 하나에 있어서, 검출된 영향은 심장 펄스의 크기 또는 크기의 변화의 영향(an effect of a magnitude, or a change in a magnitude of a heart pulse)를 포함하는 방법.
- [0041] 예 16. 예 15에 있어서, 상기 영향을 검출하는 단계는 상기 적어도 하나의 음성 특징의 복수의 값(value)의 분포(distribution)를 계산하는 단계를 포함하는 방법.
- [0042] 예 17. 예 16에 있어서, 상기 분포는 표준 편차(standard deviation)를 포함하는 방법.
- [0043] 예 18. 예 16 내지 예 17 중 어느 하나에 있어서, 상기 영향에 기초하여 부정맥 심장 상태를 결정하는 단계는 상기 분포의 특성화 파라미터(characterizing parameter)를 임계값과 비교하는 단계를 포함하는 방법.
- [0044] 예 19. 예 18에 있어서, 상기 특성화 파라미터는 상기 분포의 형태(shape of the distribution)을 포함하는 방법.
- [0045] 예 20. 예 19에 있어서, 상기 값은 상기 형태의 폭(width of the shape)을 포함하는 방법.
- [0046] 예 21. 예 1 내지 예 20 중 어느 하나에 있어서, 검출된 영향은 심장 펄스의 주기성의 영향을 포함하는 방법.
- [0047] 예 22. 예 1 내지 21 중 어느 하나에 있어서, 상기 대상의 심박수(heart rate)를 추정하는 단계를 더 포함하는 방법.
- [0048] 예 23. 예 22에 있어서, 상기 부정맥 심장 상태를 결정하는 단계는 상기 심박수의 주파수 주위로 미리 결정된 범위(predetermined range around a frequency of the heart rate)에서 상기 주기성을 특성화하는(characterizing) 단계를 포함하는 방법.
- [0049] 예 24. 예 23에 있어서, 상기 특성화하는 단계는 상기 미리 결정된 범위에서 피크의 대역폭(band width of a peak)을 계산하는 단계를 포함하는 방법.
- [0050] 예 25. 예 24에 있어서, 상기 부정맥 심장 상태를 결정하는 단계는 상기 대역폭을 임계값과 비교하는 단계를 포함하는 방법.
- [0051] 예 26. 예 1 내지 예 25 중 어느 하나에 있어서, 상기 부정맥 심장 상태를 결정하는 단계를 검증하는(validating) 단계를 더 포함하는 방법.
- [0052] 예 27. 예 26에 있어서, 상기 검증하는 단계는 미리 결정된 발성(predetermined vocalization)을 포함하는 제2 음성 샘플을 획득하는 단계를 포함하는 방법.
- [0053] 예 28. 예 26에 있어서, 상기 검증하는 단계는 상기 대상에 대한 심전도 검사(electrocardiogram test) 및 광혈류측정 검사(photoplethysmography test) 중 적어도 하나를 포함하는 방법.
- [0054] 예 29. 예 1 내지 예 28 중 어느 하나에 있어서, 상기 부정맥 심장 상태는 심방 세동(atrial fibrillation)을 포함하는 방법.
- [0055] 예 30. 예 1 내지 예 29 중 어느 하나에 있어서, 음성 샘플은 자발적인 스피치(spontaneous speech)인 방법.

- [0056] 예 31. 예 1 내지 예 30 중 어느 하나에 있어서, 상기 음성 샘플은 복수의 대상의 음성을 포함하는 방법.
- [0057] 예 32. 예 31에 있어서, 상기 음성 샘플로부터 상기 대상의 음성 세그먼트(voice segments)를 추출하는 단계를 더 포함하는 방법.
- [0058] 예 33. 예 1 내지 예 32 중 어느 하나에 있어서, 상기 음성 특징은 피치(pitch)를 포함하는 방법.
- [0059] 예 34. 예 1 내지 예 33에 있어서, 상기 음성 특징을 추출하는 단계는 Mel frequency Cepstral Coefficient(MFCC)를 추출하는 단계를 포함하는 방법.
- [0060] 예 34.1 예 1 내지 34 중 어느 하나에 있어서, 상기 음성 특징을 추출하는 단계는 교차 일관성(cross coherence)를 계산하는 단계를 포함하는 방법. 일부 실시예에서, 교차 일관성은 피치 및/또는 포먼트 주파수(formant frequencies)에서 계산된다. 일부 실시예에서, 교차 일관성은 스피치 시그널의 두 개의 세그먼트의 하나 이상의 세트(one or more sets of two segments in the speech signal) 사이에서 선택적으로 계산된다. 일부 실시예에서, 더 작은 일관성을 갖는 세트(두 개의 세그먼트)는 세그먼트 동안 발생하는 심장 펄스의 표시로서 선택적으로 기능할 수 있다(In some embodiments a set(two segments) with lower coherence can optionally serve as indication of heart pulse that occurs during the segments).
- [0061] 일부 실시예에서, 교차 일관성은 다음과 같이 계산된다:
- [0062] $Speech_frame(m) =$
- [0063] $\{Speech(t+Ts*m), Speech((t+1)+Ts*m) \dots\}$
- [0064] Where:
- [0065] "Ts" - is a time step between frames [sec]
- [0066] "Tw" - is a duration of each frame [sec]
- [0067] "m"- is a frame number
- [0068] "t" - time
- [0069] "Speech" - are the speech samples
- [0070] "Speech_frame(m)" - is a frame of speech samples at time $Ts*m$
- [0071] $Coh(m, m+T) = CrossCoherence(Speech_frame(m), Speech_frame(m+T))$
- [0072] $Coh(m, m+T)$ is a cross coherence between frame "m" (at time $m*Ts$) and frame "m+T" at time $((m+T)*Ts)$
- [0073]
- [0074] 일관성의 값(Values of the coherence)은 0에서 1 범위이며, 1은 완벽한 일치치를 의미한다.
- [0075] 상기 방법은 선택적으로, 각 프레임 m에 대해 ($Ts * m$ [sec]의 시간에) 일관성이 임계치 아래로 떨어지는 프레임 거리 T(frame distance T)를 찾는 단계를 포함한다. 일부 실시예에서, 임계값은 0.1 내지 0.98, 예를 들어 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.97, 및 0.98이다.
- [0076] 일부 실시예에서, 시간 $T * T$ 는 RR 간격의 추정치(estimated of RR interval)이다.
- [0077] 일부 실시예에서, 전체 교차 일관성 섭동(overall cross coherence perturbation; 분포의 폭) 선택적으로 AF의 표시로 기능할 수 있다. 즉, AF가 포함된 스피치는 더 넓은 분포, 더 큰 표준 편차를 가질 수 있다.
- [0078] 예 34.2 예 1 내지 예 33 중 어느 하나에 있어서, 상기 음성 특징을 추출 하는 단계는 변조 스펙트럼의 특정 특성(specific properties of the modulation spectrum)을 포함하는 방법.
- [0079] 일부 실시예에서, 선택적으로 각 프레임에서 계산되는 음향 특징의 스펙트럼의 특성(the properties of the spectrum of the acoustic feature)은 병리학 적 스피치(pathological speech)의 표시로서 기능한다.
- [0080] 34.3 예 1 내지 예 34.2 중 어느 하나에 있어서, 상기 음성 특징을 추출하는 단계는, 웨이블릿 분석(wavelet analysis);

- [0081] 성문 펄스 분석(glottal pulse analysis); 및
- [0082] 선형 예측 코딩(Linear Predictive Coding(LPC)) 분석
- [0083] 으로 구성된 그룹으로부터 선택된 스피치 시그널의 분석을 수행하는 단계를 포함한다.
- [0084] 예 35. 예 1 내지 예 33 중 어느 하나에 있어서, 상기 음성 특징은 상기 스피치 시그널의 주기와 상기 스피치 시그널의 다른 주기 사이에 교차 일관성을 수행하는 단계(performing cross coherence between a period of the speech signal and another period of the speech signal)를 포함하는 방법.
- [0085] 예 36. 예 1 내지 예 35 중 어느 하나에 있어서, 상기 음성 특징을 추출하는 단계는 웨이블릿 변환을 수행하는 단계(performing wavelet transform(WT))를 포함하는 방법.
- [0086] 예 35. 대상에 의해 제공된 음성 샘플에서 부정맥 심장 상태를 결정하기 위한 시스템에 있어서,
- [0087] 상기 음성 샘플의 디지털 음성 샘플(digital voice sample)을 획득하기 위한 음성 입력 장치(voice input);
- [0088] 상기 디지털 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 음성 특징 추출기(voice feature extractor);
- [0089] 상기 적어도 하나의 음성 특징에 대한 심장 활동(cardiac activity)의 영향을 식별하는 음성 특징 프로세서(voice feature processor);
- [0090] 상기 영향에 기초하여 부정맥 심장 상태를 결정하는 심장 상태 분류기(cardiac condition classifier)
- [0091] 를 포함하는 시스템.
- [0092] 예 36. 예 35에 있어서, 상기 음성 입력 장치 휴대폰((mobile phone)의 일부를 구성하는 시스템.
- [0093] 예 37. 예 35에 있어서, 상기 음성 입력 장치는 차량 멀티미디어 시스템(car multimedia system)의 일부를 구성하는 시스템.
- [0094] 예 38. 예 35 내지 예 37 중 어느 하나에 있어서, 상기 음성 특징 추출기, 상기 음성 특징 프로세서, 및 상기 심장 상태 분류기 중 적어도 하나는 서버(server)에 위치한 시스템.
- [0095] 예 39. 예 38에 있어서, 상기 서버는 전화 스위치와 함께 배치되어 상기 스위치로부터 데이터를 얻는(the server is collocated with a telephone switch and gets the data from the switch) 시스템.
- [0096] 예 40. 예 35 내지 예 39 중 어느 하나에 있어서, 상기 음성 입력 장치는 미리 결정된 스케줄에 따라 상기 디지털 음성 샘플을 획득하는 시스템.
- [0097] 예 41. 예 35 내지 예 40 중 어느 하나에 있어서, 영구 메모리(permanent memory)를 더 포함하는 시스템.
- [0098] 예 42. 예 41에 있어서, 상기 메모리는 부정맥 심장 상태 또는 건강한 심장 상태와 관련된 복수의 음성 샘플로부터 도출된 기준 음성 특징(reference voice features)들을 저장하는 시스템.
- [0099] 예 43. 예 42에 있어서, 상기 심장 상태 분류기는 상기 기준 음성 특징을 특성화하는 훈련 단계(training stage characterizing the reference voice features)에 기초하여 상기 부정맥 심장 상태를 결정하는 시스템.
- [0100] 예 44. 예 41에 있어서, 상기 영구 메모리는 상기 대상으로부터 미리 수집된 적어도 하나의 음성 샘플(stores at least one previously collected voice sample from the subject)을 저장하는 시스템.
- [0101] 예 45. 예 44에 있어서, 상기 심장 상태 분류기는 상기 미리 수집된 음성 샘플 및 상기 디지털 음성 샘플 사이 변화에 기초하여 상기 부정맥 심장 상태를 결정하는 시스템.
- [0102] 예 46. 심장 상태를 결정하기 위해 대상의 음성 샘플을 분석하는 방법에 있어서,
- [0103] 상기 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 단계;
- [0104] 적어도 하나의 음성 특징에 대한 혈류의 영향을 검출하는 단계로서, 상기 영향은 혈류의 타이밍, 혈류의 주기성, 및 혈류의 크기 또는 크기의 변화 중 적어도 하나이고;
- [0105] 상기 영향에 기초하여 심장 상태를 결정하는 단계
- [0106] 를 포함하는 방법.

- [0107] 예 47. 예 46에 있어서, 상기 결정된 심장 상태는 비정상 심박수(abnormal heart rate)를 포함하는 방법.
- [0108] 예 48. 예 47에 있어서, 상기 비정상 심박수는 심실 빈맥(ventricle tachycardia)을 포함하는 방법.
- [0109] 예 49. 예 46에 있어서, 결정된 심장 상태는 이소성 박동(ectopic beats)를 포함하는 방법.
- [0110] 예 50. 예 49에 있어서, 상기 이소성 박동을 갖는 결정된 심장 상태는 조기 심실 수축(premature ventricular contraction)을 포함하는 방법.
- [0111] 예 51. 심장 행동을 재구성하기 위해 대상의 음성 샘플을 분석하는 방법(a method for analyzing a voice sample of a subject to reconstruct a cardiac behavior)에 있어서,
- [0112] 상기 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 단계;
- [0113] 적어도 하나의 음성 특징에 대한 혈류의 영향을 검출하는 단계; 및
- [0114] 검출된 영향으로부터 상기 심장 행동의 하나 이상의 파라미터를 재구성하는 단계를 포함하는 방법.
- [0115] 예 52. 예 51에있어서, 상기 심장 행동은 심장 세동의 발생(an occurrence of cardiac fibrillation)을 포함하는 방법.
- [0116] 예 53. 예 52에 있어서, 미리 결정된 횟수보다 많은 심장 세동의 발생을 식별할 때 부정맥 심장 상태의 존재를 결정하는 단계(determining a presence of an arrhythmic cardiac condition when identifying more than a predetermined number of occurrences of cardiac fibrillation)를 포함하는 방법.
- [0117] 예 54. 예 52에 있어서, 미리 결정된 횟수보다 적은 심장 세동의 발생을 식별할 때 부정맥 심장 상태의 존재를 배제하는 단계(ruling out a presence of an arrhythmic cardiac condition when identifying less than a predetermined number of occurrences of cardiac fibrillation)를 포함하는 방법.
- [0118] 예 55. 예 51에있어서, 상기 심장 행동의 하나 이상의 파라미터는 비 부정맥 심장 상태(non-arrhythmic cardiac condition)를 나타내는 방법.
- [0119] 예 56. 예 51에 있어서, 상기 하나 이상의 파라미터는 펄스 웨이브 형태(pulse wave shape)을 포함하는 방법.
- [0120] 본 발명의 일부 실시예에 따르면, 부정맥 심장 상태(arrhythmic cardiac condition)를 결정하기 위해 대상(subject)의 음성 샘플(voice sample)을 분석하는 방법에 있어서, 상기 음성 샘플로부터 적어도 하나의 음성 특징(voice feature)을 추출하는 단계, 상기 적어도 하나의 음성 특징에 대한 심장 상태(cardiac condition)의 영향(effect)을 검출하는 단계, 상기 영향에 기초하여 부정맥 심장 상태를 결정하는 단계를 포함하는 방법이 제공된다.
- [0121] 본 발명의 일부 실시예에 따르면, 상기 적어도 하나의 음성 특징에 대한 심장 상태의 영향을 검출하는 단계는, 상기 적어도 하나의 음성 특징의 혈류(blood flow),
- [0122] 심실 박동(ventricular beat), 및 심방 박동(atrial beat)을 포함하는 그룹으로부터 선택된 적어도 하나의 음성 에 영향을 미치는 원인(cause affecting voice)에 의해 영향을 받는 심장 상태의 영향을 검출하는 단계를 포함한다.
- [0123] 본 발명의 일부 실시예에 따르면, 상기 음성에 영향을 미치는 원인의 타이밍(timing), 상기 음성에 영향을 미치는 원인의 주기(periodicity), 및 상기 음성에 영향을 미치는 원인의 크기(magnitude) 또는 크기의 변화(change in magnitude) 중 적어도 하나의 영향을 식별하는 단계를 포함한다.
- [0124] 본 발명의 일부 실시예에 따르면, 상기 부정맥 심장 상태를 결정하는 단계는, 복수의 상기 식별된 영향을 통합하는(integrating) 단계를 포함한다.
- [0125] 본 발명의 일부 실시예에 따르면, 유성 세그먼트(voiced segments)를 식별하기 위해 음성 샘플을 분류하는 단계를 더 포함한다.
- [0126] 본 발명의 일부 실시 예에 따르면, 적어도 일부 세그먼트가 제거되고 나머지 세그먼트 엔드(segment ends)가 평탄화되도록(smoothed) 유성 세그먼트를 연결하는(concatenating) 단계를 더 포함한다.
- [0127] 본 발명의 일부 실시예에 따르면, 음성 샘플로부터 제거되는 세그먼트에서 음성에 영향을 미치는 원인에 관련되는 데이터(data pertaining to the cause affecting voice)를 외삽하는(extrapolating) 단계를 더 포함한다.

- [0128] 본 발명의 일부 실시예에 따르면, 검출된 영향은 심장 펄스의 타이밍(a timing of heart pulses)을 포함한다.
- [0129] 본 발명의 일부 실시예에 따르면, 영향을 검출하는 단계는 세 번의 연속적인 심장 펄스의 발생(three consecutive occurrences of heart pulses)을 식별하는 단계를 포함한다.
- [0130] 본 발명의 일부 실시예에 따르면, 부정맥 심장 상태를 결정하는 단계는 연속적인 심장 펄스 발생 사이의 시간 간격(time intervals between the consecutive occurrences of heart pulses)을 계산하고 시간 간격을 기준 심장 조건으로부터 획득된 기준 시간 간격과 매칭시키는 단계를 포함한다.
- [0131] 본 발명의 일부 실시예에 따르면, 결정하는 단계는 매칭의 확률(a probability of the matching)을 결정하는 단계, 및 확률을 부정맥 심장 상태를 결정하기 위한 임계 값과 비교하는 단계를 더 포함한다.
- [0132] 본 발명의 일부 실시예에 따르면, 기준 심장 상태는 건강(healthy)이거나 부정맥(arrhythmic)이다.
- [0133] 본 발명의 일부 실시예에 따르면, 세 번의 심장 펄스의 발생은 동일한 음성 세그먼트에서 식별되지 않은 것이다(the three occurrences of a heart pulse are not identified in the same voiced segment).
- [0134] 본 발명의 일부 실시예에 따르면, 영향을 검출하는 단계는, 두 번의 연속적인 심장 펄스의 발생을 식별하는 단계를 포함한다.
- [0135] 본 발명의 일부 실시예에 따르면, 검출된 효과는 크기의 영향, 또는 심장 펄스의 크기 변화(an effect of a magnitude, or a change in a magnitude of a heart pulse)를 포함한다.
- [0136] 본 발명의 일부 실시예에 따르면, 영향을 검출하는 단계는 적어도 하나의 음성 특징의 복수의 값의 분포(a distribution of a plurality of values of the at least one voice feature)를 계산하는 단계를 포함한다.
- [0137] 본 발명의 일부 실시예에 따르면, 영향에 기초하여 부정맥 심장 상태를 결정하는 단계는 분포의 특성화 파라미터(a characterizing parameter of the distribution)를 임계 값과 비교하는 단계를 포함한다.
- [0138] 본 발명의 일부 실시예에 따르면, 특성화 파라미터는 분포의 형태(shape of the distribution)를 포함한다.
- [0139] 본 발명의 일부 실시예에 따르면, 값은 형태의 폭(width of the shape)을 포함한다.
- [0140] 본 발명의 일부 실시예에 따르면, 검출된 영향은 심장 펄스의 주기성(a periodicity of heart pulses)에 대한 영향을 포함한다.
- [0141] 본 발명의 일부 실시예에 따르면, 대상의 심박수(a heart rate of the subject)를 추정하는 단계를 더 포함한다.
- [0142] 본 발명의 일부 실시예에 따르면, 부정맥 심장 상태를 결정하는 단계는 심박수의 주파수 주위의 미리 결정된 범위(the periodicity at a predetermined range around a frequency of the heart rate)에서의 주기성을 특성화하는(characterizing) 단계를 포함하고, 특성화하는 단계는 미리 결정된 범위에서 피크의 대역폭(a band width of a peak at the predetermined range)을 계산하는 단계를 포함하고, 부정맥 심장 상태를 결정하는 단계는 대역폭을 임계 값과 비교하는 단계를 포함한다.
- [0143] 본 발명의 일부 실시예에 따르면, 부정맥 심장 상태를 결정하는 단계를 검증하는(validating) 단계를 더 포함한다.
- [0144] 본 발명의 일부 실시예에 따르면, 검증하는 단계는 미리 결정된 발성(a predetermined vocalization)을 포함하는 제2 음성 샘플을 획득하는 단계를 포함한다.
- [0145] 본 발명의 일부 실시예에 따르면, 검증하는 단계는 대상에 대한 심전도 검사(electrocardiogram test) 및 광혈류측정 검사(photoplethysmography test) 중 적어도 하나를 포함한다.
- [0146] 본 발명의 일부 실시예에 따르면, 부정맥 심장 상태는 심방 세동(atrial fibrillation)을 포함한다.
- [0147] 본 발명의 일부 실시예에 따르면, 음성 샘플은 자발적인 스피치(spontaneous speech)이다.
- [0148] 본 발명의 일부 실시예에 따르면, 음성 특징은 피치(pitch)를 포함한다.
- [0149] 본 발명의 일부 실시예에 따르면, 음성 특징을 추출하는 단계는 음성 샘플의 제1 세그먼트 및 음성 샘플의 제2 세그먼트 사이 교차 일관성을 수행하는(performing cross coherence) 단계를 포함한다.

- [0150] 본 발명의 일부 실시예에 따르면, 음성 특징을 추출하는 단계는 웨이블릿 분석(wavelet analysis), Mel frequency Cepstral Coefficient(MFCC) 분석, 성문 펄스 분석(glottal pulse analysis), 선형 예측 코딩(Linear Predictive Coding(LPC)) 분석을 포함하는 그룹으로부터 선택된 피치 시그널(speech signal)의 분석을 수행하는 단계를 포함하는 포함한다.
- [0151] 본 발명의 일부 실시예에 따르면, 대상에 의해 제공된 음성 샘플에서 부정맥 심장 상태를 결정하기 위한 시스템에 있어서, 시스템은 음성 샘플의 디지털 음성 샘플을 획득하기 위한 음성 입력 장치(voice input), 디지털 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 음성 특징 추출기(voice feature extractor), 적어도 하나의 음성 특징에 대한 심장 활동(cardiac activity)의 영향을 식별하는 음성 특징 프로세서(voice feature processor), 영향에 기초하여 부정맥 심장 상태를 결정하는 심장 상태 분류기(cardiac condition classifier)를 포함한다.
- [0152] 본 발명의 일부 실시예에 따르면, 상기 음성 입력 장치는 휴대폰(mobile phone), 디지털 어시스턴트(digital assistant), 및 차량 멀티미디어 시스템(car multimedia system) 중 적어도 하나의 일부를 구성한다.
- [0153] 본 발명의 일부 실시예에 따르면, 상기 음성 특징 추출기, 상기 음성 특징 프로세서, 및 상기 심장 상태 분류기 중 적어도 하나는, 서버(server)에 위치한다.
- [0154] 본 발명의 일부 실시예에 따르면, 상기 서버는 전화 스위치와 함께 배치되어 상기 스위치로부터 데이터를 얻는다(the server is collocated with a telephone switch and gets the data from the switch).
- [0155] 본 발명의 일부 실시예에 따르면, 상기 음성 입력 장치는, 미리 결정된 스케줄에 따라 상기 디지털 음성 샘플을 획득한다.
- [0156] 본 발명의 일부 실시예에 따르면, 영구 메모리(permanent memory)를 더 포함하고, 상기 메모리는 부정맥 심장 상태 또는 건강한 심장 상태와 관련된 복수의 음성 샘플로부터 유도된(derived) 기준 음성 특징을 저장한다.
- [0157] 본 발명의 일부 실시예에 따르면, 상기 심장 상태 분류기는, 상기 기준 음성 특징을 특성화하는 훈련 단계(a training stage characterizing the reference voice features)에 기초하여 상기 부정맥 심장 상태를 결정한다.
- [0158] 본 발명의 일부 실시예에 따르면, 상기 영구 메모리는, 상기 대상으로부터 미리 수집된 적어도 하나의 음성 샘플을 저장한다.
- [0159] 본 발명의 일부 실시예에 따르면, 상기 심장 상태 분류기는, 상기 미리 수집된 음성 샘플 및 상기 디지털 음성 샘플 사이 변화에 기초하여 상기 부정맥 심장 상태를 결정한다.
- [0160] 본 발명의 일부 실시예에 따르면, 대상에 의해 제공된 음성 샘플에서 부정맥 심장 상태를 결정하기 위한 시스템에 있어서, 상기 시스템은, 상기 음성 샘플의 디지털 음성 샘플을 획득하기 위한 음성 입력 장치(voice input), 상기 디지털 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 음성 특징 추출기(voice feature extractor), 상기 적어도 하나의 음성 특징에 대한 심장 상태(cardiac condition)의 영향을 식별하는 음성 특징 프로세서(voice feature processor), 상기 영향에 기초하여 상기 심장 상태를 결정하는 심장 상태 분류기(cardiac condition classifier)를 포함한다.
- [0161] 본 발명의 일부 실시예에 따르면, 심장 상태를 결정하기 위해 대상의 음성 샘플을 분석하는 방법에 있어서, 상기 방법은 상기 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 단계, 상기 적어도 하나의 음성 특징에 대한 혈류의 영향을 검출하는 단계, 상기 영향에 기초하여 심장 상태를 결정하는 단계를 포함하고, 상기 영향은 상기 혈류의 타이밍, 상기 혈류의 주기성, 및 상기 혈류의 크기 또는 상기 혈류의 크기의 변화 중 적어도 하나이다.
- [0162] 본 발명의 일부 실시예에 따르면, 상기 결정된 심장 상태는, 비정상적인 심박수(abnormal heart rate)를 포함한다.
- [0163] 본 발명의 일부 실시예에 따르면, 상기 비정상적인 심박수는, 심실 빈맥(ventricle tachycardia)을 포함한다.
- [0164] 본 발명의 일부 실시예에 따르면, 상기 결정된 심장 상태는, 이소성 박동(ectopic beats) 및 조기 심실 수축(premature ventricular contraction) 중 적어도 하나를 포함한다.
- [0165] 본 발명의 일부 실시예에 따르면, 심장 행동(cardiac behavior)을 재구성(reconstruct)하기 위해 대상의 음성

샘플을 분석하는 방법에 있어서, 상기 방법은 상기 음성 샘플로부터 적어도 하나의 음성 특징을 추출하는 단계, 상기 적어도 하나의 음성 특징에 대한 혈류의 영향을 검출하는 단계, 및 상기 검출된 영향으로부터 상기 심장 행동의 하나 이상의 파라미터를 재구성하는 단계를 포함한다.

[0166] 본 발명의 일부 실시예에 따르면, 상기 심장 행동은, 심장 세동(cardiac fibrillation)의 발생을 포함한다.

[0167] 본 발명의 일부 실시예에 따르면, 미리 결정된 횟수보다 많은 심장 세동의 발생을 식별할 때 부정맥 심장 상태의 존재를 결정하는 단계를 포함한다.

[0168] 본 발명의 일부 실시예에 따르면, 미리 결정된 횟수보다 적은 심장 세동의 발생을 식별할 때 부정맥 심장 상태의 존재를 배제하는 단계를 포함한다.

[0169] 본 발명의 일부 실시예에 따르면, 상기 심장 행동의 하나 이상의 파라미터는, 비 부정맥 심장 상태(non-arrhythmic cardiac condition)를 나타낸다.

[0170] 본 발명의 일부 실시예에 따르면, 상기 하나 이상의 파라미터는, 펄스 웨이브 형태(pulse wave shape)을 포함한다.

[0171] 해당 기술 분야의 통상의 기술자라면 이해할 수 있는 바와 같이, 본 발명의 일부 실시예는 시스템, 방법 또는 컴퓨터 프로그램 제품으로 구현될 수 있다. 따라서, 본 발명의 일부 실시예는 전체 하드웨어 실시예, 전체 소프트웨어 실시예 (펌웨어, 레지던트 소프트웨어, 마이크로 코드 등을 포함) 또는 일반적으로 "회로", "모듈" 또는 "시스템"으로 지칭될 수 있는 소프트웨어 및 하드웨어를 결합한 실시예의 형태를 취할 수 있다. 또한, 본 발명의 일부 실시예는 컴퓨터 판독 가능 프로그램 코드가 구현된 하나 이상의 컴퓨터 판독 가능 매체(들)에 구현된 컴퓨터 프로그램 제품의 형태를 취할 수 있다. 본 발명의 일부 실시예의 방법 및/또는 시스템의 구현은 선택된 작업을 수동, 자동 또는 이들의 조합을 수행 및/또는 완료하는 것을 포함할 수 있다. 더욱이, 본 발명의 방법 및/또는 시스템의 일부 실시예들의 실제 계측 및 장비에 따르면, 몇몇 선택된 작업들은 하드웨어, 소프트웨어 또는 펌웨어 및/또는 이들의 조합에 의해, 예를 들어 운영 체제를 사용하여 구현될 수 있다.

[0172] 예를 들어, 본 발명의 일부 실시예에 따라 선택된 작업을 수행하기 위한 하드웨어는 칩 또는 회로로서 구현될 수 있다. 소프트웨어로서, 본 발명의 일부 실시예에 따른 선택된 작업은 임의의 적합한 운영 체제를 사용하여 컴퓨터에 의해 실행되는 복수의 소프트웨어 명령으로서 구현될 수 있다. 본 발명의 예시적인 실시예에서, 본 명세서에 기술된 방법 및/또는 시스템의 일부 예시적인 실시예에 따른 하나 이상의 작업은 복수의 명령, 심지어 셀룰러 폰을 실행하기 위한 컴퓨팅 플랫폼과 같은 데이터 프로세서에 의해 수행된다. 선택적으로, 데이터 프로세서는 명령 및/또는 데이터를 저장하기 위한 휘발성 메모리 및/또는 명령 및/또는 데이터를 저장하기 위한 자기 하드 디스크 및/또는 착탈식 매체와 같은 비 휘발성 저장 장치를 포함한다. 선택적으로, 네트워크 연결도 제공된다. 키보드 및 마우스와 같은 디스플레이 및/또는 사용자 입력 장치도 선택적으로 제공된다.

[0173] 하나 이상의 컴퓨터 판독 가능 매체(들)의 임의의 조합이 본 발명의 일부 실시예에 이용될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터 판독 가능 신호 매체 또는 컴퓨터 판독 가능 저장 매체일 수 있다. 컴퓨터 판독 가능 저장 매체는 예를 들어, 전자, 자기, 광학, 전자기, 적외선 또는 반도체 시스템, 장치 또는 장치, 또는 전송한 것의 임의의 적절한 조합일 수 있지만, 이에 제한되지는 않는다. 컴퓨터 판독 가능 저장 매체의 구체적인 예 (비제한적 목록)는 하나 이상의 와이어를 갖는 전기 접속, 휴대용 컴퓨터 디스켓, 하드 디스크, 랜덤 액세스 메모리(RAM), 판독-메모리(ROM), 소거 가능한 프로그램 가능 읽기 전용 메모리(EPROM 또는 플래시 메모리), 광섬유, 휴대용 CD-ROM, 광학 저장 장치, 자기 저장 장치 또는 기타 상기의 적절한 조합을 포함한다. 이 문서의 기재에서, 컴퓨터 판독 가능 저장 매체는 명령 실행 시스템, 장치 또는 장치에 의해 또는 이와 관련하여 사용하기 위한 프로그램을 포함하거나 저장할 수 있는 임의의 유형 매체일 수 있다.

[0174] 컴퓨터 판독 가능 신호 매체는 예를 들어 기저 대역 또는 반송파의 일부로서 컴퓨터 판독 가능 프로그램 코드가 구현된 전파된 데이터 신호를 포함할 수 있다. 이러한 전파된 신호는 전자기, 광학 또는 이들의 임의의 적절한 조합을 포함하지만 이에 제한되지 않는 임의의 다양한 형태를 취할 수 있다. 컴퓨터 판독 가능 신호 매체는 컴퓨터 판독 가능 저장 매체가 아니며 명령 실행 시스템, 장치 또는 장치에 의해 또는 이와 관련하여 사용하기 위한 프로그램을 통신, 전파 또는 수송할 수 있는 임의의 컴퓨터 판독 가능 매체일 수 있다.

[0175] 컴퓨터 판독 가능 매체에 구현된 프로그램 코드 및/또는 그에 의해 사용되는 데이터는 무선, 유선, 광섬유 케이블, RF 등을 포함하지만 이에 제한되지 않는 임의의 적절한 매체를 사용하여 전송될 수 있다.

[0176] 본 발명의 일부 실시예에 대한 동작을 수행하기 위한 컴퓨터 프로그램 코드는 Java, Smalltalk, C++ 등과 같은

객체 지향 프로그래밍 언어 및 "C" 프로그래밍 언어 또는 유사한 프로그래밍 언어와 같은 종래의 절차적 프로그래밍 언어를 포함하는 하나 이상의 프로그래밍 언어의 임의의 조합으로 작성될 수 있다. 프로그램 코드는 사용자의 컴퓨터, 부분적으로 사용자의 컴퓨터, 독립형 소프트웨어 패키지, 부분적으로 사용자의 컴퓨터 및 부분적으로는 원격 컴퓨터 또는 원격 컴퓨터 또는 서버에서 완전히 실행될 수 있다. 후자의 시나리오에서, 원격 컴퓨터는 근거리 통신망(LAN) 또는 광역 통신망(WAN)을 포함하는 임의의 유형의 네트워크를 통해 사용자의 컴퓨터에 연결될 수 있거나, 외부 컴퓨터에 연결될 수 있다(예를 들어, 인터넷 서비스 제공 업체를 사용하여 인터넷을 통해).

[0177] 본 발명의 일부 실시예는 본 발명의 실시 예에 따른 방법, 장치(시스템) 및 컴퓨터 프로그램 제품의 흐름도 및/또는 블록도를 참조하여 아래에 설명될 수 있다. 흐름도 및/또는 블록도의 각 블록, 및 흐름도 및/또는 블록도의 블록의 조합은 컴퓨터 프로그램 명령에 의해 구현될 수 있다. 이들 컴퓨터 프로그램 명령은 범용 컴퓨터, 특수 목적 컴퓨터, 또는 기타 프로그램 가능 데이터 처리 장치의 프로세서에 제공됨으로써 기계를 생성하여 명령이 컴퓨터의 프로세서 또는 다른 프로그램 가능 데이터 처리 장치를 통해 실행되도록 할 수 있고, 플로우 차트 및/또는 블록 다이어그램 블록에 지정된 기능/동작을 구현하기 위한 수단을 생성할 수 있다.

[0178] 이러한 컴퓨터 프로그램 명령은 또한 컴퓨터 판독 가능 매체에 저장된 명령이 플로우차트 및/또는 블록 다이어그램 블록 또는 블록에 지정된 기능/액션을 수행하는 인스트럭션을 포함하는 제조 아티클(article of manufacture)을 생성하도록 컴퓨터, 다른 프로그램 가능 데이터 처리 장치 또는 다른 장치가 특정 방식으로 기능하도록 지시 할 수 있는 컴퓨터 판독 가능 매체에 저장 될 수 있다

[0179] 컴퓨터 프로그램 명령은 또한 컴퓨터, 다른 프로그램 가능 데이터 처리 장치 또는 다른 장치에 로딩되어 일련의 동작 단계가 컴퓨터, 다른 프로그램 가능 장치 또는 다른 장치에서 수행되어 컴퓨터로 구현되는 프로세스를 생성하게 한다. 컴퓨터 또는 다른 프로그램 가능 장치에서 실행되는 명령어는 흐름도 및/또는 블록도 블록에 지정된 기능/동작을 구현하기위한 프로세스를 제공한다.

[0180] 본 명세서에 기술 된 방법 중 일부는 일반적으로 컴퓨터에 의해서만 사용되도록 설계되며, 전문가에 의해 순전히 수동으로 수행하기에는 적합하지 않거나 실용적이지 않을 수 있다. 음향 패턴 식별 및 매칭(identifying and matching acoustic patterns)과 유사한 작업을 수동으로 수행하고자 하는 전문가는 완전히 다른 방법, 예를 들어 인간 두뇌의 전문 지식 및/또는 패턴 인식 기능을 이용하는 것과 같은, 여기에 설명 된 방법의 단계를 수동으로 수행하는 것보다 훨씬 더 효율적인 완전히 다른 방법을 사용해야 할 수 있다.

[0181] 본 명세서에서 사용되는 회로 또는 회로라는 용어는 선택된 작업을 수행하기 위한 명령을 갖는 하드웨어 및/또는 소프트웨어, 예를 들어, 특정 계산을 수행하는 메모리(소프트웨어 또는 알고리즘)를 갖는 컨트롤러, 또는 일부 신호 처리 작업을 수행하는 칩을 지칭한다(예를 들어, 음성 분석을 수행하기위한 목적으로).

도면의 간단한 설명

[0182] 본 발명의 일부 실시예는 첨부 도면을 참조하여, 단지 예로서, 설명된다. 도면을 구체적으로 참조하면, 도시된 세부 사항들은 예로써, 그리고 본 발명의 구현에 대한 예시를 위한 것임을 강조한다. 이와 관련하여, 도면과 함께 이루어진 설명은 본 발명의 실시예가 어떻게 실시 될 수 있는지 당업자에게 명백하다.

도 1a는 본 발명의 일부 실시예에 따라, 일반적인 음성 검출 및 분석 컴포넌트 및 이들 사이의 데이터 흐름을 도시하는 상위 레벨 흐름도이다.

도 1b는 본 발명의 일부 실시예에 따른 웨이블릿 분해의 단순화된 도면이다.

도 2는 본 발명의 일부 실시예에 따른 선택적 음성 분석 스테이션 사이의 데이터 흐름을 나타내는 데이터 흐름도이다.

도 3은 본 발명의 일부 실시예에 따른 음성 샘플에 기초한 심장 상태 추정의 일반적인 상위 레벨 개요이다.

도 4a 내지 4b는 예시적인 플랫폼에서의 음성 샘플링 및/또는 분석을 개략적으로 도시 한 도면으로서, 도 4a는 휴대 전화 플랫폼을 도시 한 도면이고, 도 4b는 본 발명의 일부 실시예에 따른 차량 플랫폼을 도시한다.

도 5a 내지 5c는 본 발명의 일부 실시예에 따른 예시적인 음성 처리를 도시 한 도면으로서, 도 5a는 음향 벡터를 얻기 위한 음성 샘플 처리를 도시하고, 도 5b는 음성 특징 정규화 프로세스를 도시 한 도면이고, 도 5c는 정규화 예를 나타낸다.

도 6a는 본 발명의 일부 실시예에 따라 분포 분석에 의한 음성 샘플로부터의 심장 상태의 추정을 나타내는 일

반적인 상위 레벨 개요이다.

도 6b는 본 발명의 일부 실시예에 따라, 음성 특징에 대한 예로서 피치를 도식하는 특징 분포 분석을 위한 상세한 프로세스를 나타내는 흐름도이다.

도 6c는 본 발명의 일부 실시예에 따라 심장 상태를 진단하는 것을 학습하는 방법의 단순화된 흐름도이다.

도 6d는 본 발명의 일부 실시예에 따라 심장 상태를 분류하는 방법의 단순화된 흐름도이다.

도 7a 내지 7b는 도 7a에 도식된 음성 샘플을 나타내는 것으로, 본 발명의 일부 실시예에 따라, 음성 샘플은 세그먼트를 제거하도록 분류되고 나머지 세그먼트는 도 7b에 연결되어 있다.

도 8a 내지 8b는 평탄화된 음성 샘플로부터 추출된 피치 데이터에 대한 예를 도시한 도면으로서, 도 8a는 음성의 "자연적인" 변동의 평탄화를 도시한 도면이고, 도 8b는 본 발명의 일부 실시예에 따라 평탄화된 데이터를 도시한다.

도 9a 내지 9b는 본 발명의 일부 실시예에 따라 필터링되고 평탄화된 후 음성 샘플로부터 추출된 피치 데이터에 대한 예를 도시한 도면으로서, 도 9a는 연결된 유성 세그먼트에 대한 예를 도시하고, 도 9b는 피치 추출을 위한 예를 도시한다.

도 10a 내지 10d는 일부 실시예에 따라 AF 동안의 기록된 스피치(도 10b 및 10d) 및 건강한 사람의 스피치(도 10a 및 10c) 동안의 스피치의 피치의 정규화된 표준 편차 값의 분포의 히스토그램의 예를 나타낸다.

도 11a 내지 11c는 본 발명의 일부 실시예에 따른, 펄스 간격 분석 형태의 심장 타이밍 분석을 도시한 도면으로서, 도 11a는 일반적인 흐름도를 도시하고, 도 11b는 상세한 흐름도를 도시하고, 도 11c는 본 발명의 일부 실시예에 따른 시간에 따른 예시적인 펄스 결정을 도시한 도면이다.

도 11d는 본 발명의 일부 실시예에 따른 RR 간격 추정의 예시적인 그래프를 나타낸다.

도 12a는 본 발명의 일부 실시예에 따라, MFCC 타입 음성 특징에 대한 프레임 대 프레임 거리의 예를 도시하며, 상단 그래프는 프레임 대 프레임 거리 패턴을 나타내고 하단 그래프는 기준으로서 ECG 그래프를 나타낸다.

도 12b는 본 발명의 일부 실시예에 따른 일관성 계산의 예시적인 그래프를 도시한다.

도 13은 본 발명의 일부 실시예에 따라 연속적인 펄스의 발생이 증가함에 따라 민감도 변화 대 심장 상태 결정의 특이성에 대한 예를 도시한 도면이다.

도 14a 내지 14b는 본 발명의 일부 실시예에 따라 심장 상태를 결정하기 위해 사용된 심박수에서 주기성을 결정하는 프로세스를 도시한 도면으로서, 도 14a는 상위 레벨의 개요이고, 도 14b는 상세한 흐름도이다.

도 14c는 본 발명의 일부 실시예에 따른 성문 펄스의 단순화된 도면이다.

도 15a 내지 15d는 본 발명의 일부 실시예에 따라, 음성 특징의 스펙트럼 특성을 사용하여 심박수 병리에 대한 확률의 결정을 도시한 도면으로서, 도 15a 및 15c는 건강한 패턴을 도시하고, 도 15b 및 15d는 본 발명의 일부 실시예에 따른 AF 패턴을 도시한다.

도 16은 본 발명의 일부 실시예에 따라 가능한 치료 및 모니터링의 타임 라인을 도시한 도면이다.

도 17은 본 발명의 일부 실시예에 따른 음성 샘플로부터의 심박수 추출의 예이다.

도 18a는 본 발명의 일부 실시예에 따른 스피커 독립적 검출 결과를 나타내는 그래프이고, 도 18b는 본 발명의 일부 실시예에 따른 스피커 종속적 검출 결과를 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0183] 본 출원은 2017년 2월 12일에 출원된 미국특허출원 US62/457,914의 우선권을 주장하며, 그 내용은 그 전문이 본원에 참조로 포함된다.

[0184] 본 발명은 일부 실시예에서 음성을 이용한 의료 진단에 관한 것으로, 이에 한정되지는 않지만, 보다 상세하게는 인간 음성에 기초한 심장 상태 추정에 관한 것이다.

[0185] **개관**

[0186] 본 발명의 일부 실시예의 양태는 음성 샘플(voice sample)에 대하여 혈류와 같은 심장 활동의 영향의 타이밍 및

/또는 크기의 영향을 결정하는 것에 관한 것이다.

- [0187] 심장 활동은 다음 중 하나 이상에서 음성 샘플에 영향을 미칠 수 있다:
- [0188] 음성 샘플에 영향을 미치는 혈류에 영향을 미침;
- [0189] 심장 박동(heart beat)의 영향을 폐(lungs)로 전송하는데, 이는 음성 샘플에 영향을 미침. 왼쪽 폐는 가슴의 공간을 심장과 공유함;
- [0190] 심장 박동의 영향을 기관지(bronchi)로 전송하는데, 이는 음성 샘플에 영향을 미침;
- [0191] 심장 박동의 영향을 흉막액(pleural fluid)으로 전송하는데, 이는 음성 샘플에 영향을 미침; 및
- [0192] 잠재적으로 감지 가능한 발성 소리(vocal sound)의 변조를 야기하는 동맥(arteries), 후두(larynx; “음성 상자(voice box)” 및 성도(vocal tract)에서의 심장 박동 관련 기계적 변화(heartbeat-related mechanical changes in arteries and muscles along the larynx (the “voice box” and the vocal tract which potentially cause detectable modulations in the vocal sounds));
- [0193] 일부 실시예에서, 대상의 음성 샘플이 분석되어 음성 샘플에 대한 심장 상태의 영향을 검출함으로써 부정맥 심장 상태를 결정하고,
- [0194] 일부 실시예에서, 분석하는 단계는 음성 샘플에 영향을 미치는 원인의 영향을 검출하는 단계를 포함한다.
- [0195] 이러한 원인의 예는(이에, 한정되지 않지만) 다음을 포함한다
- [0196] 혈류(flood flow);
- [0197] 심실 박동(ventricular beat); 및
- [0198] 심방 박동(atrial beat).
- [0199] 본 출원에서, 음성 샘플에 영향을 미치는 혈류의 예시적인 효과가 설명되고, 명세서 및 청구 범위에서 용어 "혈류"는 음성 샘플에 대한 심장 박동의 영향을 포함하는 의미로 의도된다.
- [0200] 일부 실시예에서, 심장 활동 영향(cardiac activity effect)은 심장 펄스의 타이밍에 의해 영향을 받는 혈류를 포함한다. 대안적으로 또는 추가적으로, 심장 활동 영향은 심장 펄스의 크기에 의해 영향을 받는 혈류를 포함한다. 일부 실시예에서, 음성 샘플에 대한 영향은 부정맥 심장 상태, 예를 들어 심방 세동(AF 또는 Afib)을 결정하는데 이용된다. 대안적으로 또는 추가적으로, 음성 샘플에 대한 영향은 비정상 심장 상태, 예를 들어 빈맥(trachycardia(VT))을 결정하는데 이용된다. 대안적으로 또는 추가적으로, 음성 샘플에 대한 영향은 건강한 심장 상태, 예를 들어 타이밍 및/또는 크기에서 충분한 주기성을 갖는 심장 상태를 결정하기 위해 이용된다.
- [0201] 일부 실시예에서, 임의로 자발적 음성 중에(optionally during spontaneous speech) 음성을 디지털 형식으로 샘플링하고, 심장 상태를 식별하기 위해 디지털 음성 샘플을 분석함으로써 대상의 음성에 대한 영향이 결정된다. 대안적으로 또는 추가적으로, 영향은 미리 정해진 유성 소리(predetermined voiced sounds)에 기초하여 결정된다. 일부 실시예에서, 대상 본인의 적극적인 행동없이 대상의 스피치가 수동적으로 샘플링된다. 일부 실시예에서, 음성 샘플은 음성 특징(voice feature)을 추출하기 위해 분석된다. 예를 들어, 음성 특징은 피치(pitch) 및/또는 MFCC(Mel Frequency Cepstral Coefficients)를 포함 할 수 있다. 대안적으로 또는 추가적으로, 음성 샘플 또는 음성 특징 자체는 RR 간격(RR interval; 예를 들어, 연속적인 심장 박동 사이의 시간 간격)과 같은 심장 데이터를 추출하기 위해 분석된다.
- [0202] 일부 실시예에서, 심박수의 변동성 또는 변화(variability or change)가 검출 및/또는 측정된다.
- [0203] 본 명세서에 사용된 바와 같이, 용어 "음성 특징(voice feature)"은 이 값의 임의의 통계적 및/또는 수학적 연산에 추가로, 미분, 표준 편차, 자기 상관, 일관성, 교차 일관성, 웨이블릿 변환, 왜곡 등(derivatives, standard deviation, auto-correlation, coherence, cross-coherence, wavelet transforms, skewedness, et c)과 같은 스피치(speech)로부터 계산될 수 있는 임의의 값을 포함한다. 또한, "음성 특징"이라는 용어는 시간에 따른 다른 음성 특징의 변화와 같이, 시간에 따른 스피치의 특성의 변화 값도 포함한다.
- [0204] 일반적으로, 스피치 생성 중에(during speech production), 성대 주름(vocal folds)이 진동하여 공기의 펄스를 생성합니다. 성대 주름의 질량 및 장력 및 폐 기압이 피치를 결정한다. 공기의 펄스는 입/목의 구조와 부피에 따라 스펙트럼 밀도(spectral density)와 구성을 변화시키고 다양한 소리를 만든다.

- [0205] 심장 박동은 여러 메커니즘에 의해 음성에 잠재적으로 영향을 미친다:
- [0206] 일시적 혈압 및/또는 혈관 부피(temporal blood pressure and/or the blood vessels volume)를 변화시킴. 성대 주름 조직(the vocal-folds tissue)에서 이러한 변화는 음성에 영향을 미치는 성대 질량(vocal cord mass)에 작은 변화를 일으킴;
- [0207] 또한, 혈압 및 심장 박동은 또한 코 및/또는 구강의 부피의 작은 변화를 야기 할 수 있으며, 이는 또한 스피치에 영향을 줄 수 있음;
- [0208] 심장 박동은 "기계적 맥박(mechanical pulse)"을 만들어, 폐, 성대(vocal cords) 및/또는 입을 진동 시켜 잠재적으로 스피치의 변조(modulation)을 유발함;
- [0209] 목을 통과하는 큰 동맥(artery)의 박동(pulsations)은 성대 주름에 기계적 진동을 일으킬 수 있음.
- [0210] 심방 세동(atrial fibrillation(AF))은 일반적으로 불규칙한 심장 리듬(heart rhythms)으로 식별되며 임상적으로 심방의 조정되지 않은 수축(uncoordinated contractions of the atria)으로 정의됩니다. AF는 증상이 없을 수 있다. AF가 있으면 뇌졸중(strokes)의 가능성이 최대 5배까지 더 높다. 현재의 의료 관행은 AF 관련 뇌졸중(strokes)의 약 80%를 예방한다. 그러므로 의학적 치료를 시작하기 위해 AF로 고통받는 대상을 조기에 식별하는 것이 잠재적인 이점이 된다
- [0211] AF를 감지하는 일부 방법은 주로 연속 ECG 기록(예를 들어, 심장 터 모니터 cardiac holter monitors), 모바일 심장 원격 측정 모니터(mobile cardiac telemetry monitors))의 사용에 의존한다. AF는 하루에 몇 분 동안만 발생할 수 있고 증상이 없으므로 연속적인 측정이 필요합니다. 그러나 ECG를 통한 지속적인 심장 모니터링은 다양한 전극의 정확한 적용, 불편한 장치, 케이블 연결, 웨어러블 센서 및 배터리 교체 또는 재충전 문제와 같은 문제를 야기할 수 있다. 수동 모니터링은 선별된 대상으로부터 능동적인 행동 단계를 유발하지 않고 일상 생활에서 벗어나지 않고도 심장 상태를 식별할 수 있는 잠재적인 이점이 있다.
- [0212] 일부 실시예에서, 음성은, 예를 들어 마이크로폰 장치(microphone device)와 같은 음성 입력 장치(voice input)에 의해 샘플링된다. 선택적으로, 음성 입력 장치는 전화(phone)의 일부를 구성한다. 대안적으로 또는 추가적으로, 음성 입력 장치는 컴퓨터(computer) 및/또는 스마트 폰(smartphone)의 일부를 구성한다. 대안적으로 또는 추가적으로, 음성 입력 장치는 차량 멀티미디어 시스템(car multimedia system) 및/또는 스피커폰(speaker phone) 및/또는 차량의 다른 관련 시스템의 일부를 구성한다. 대안적으로 또는 추가적으로, 음성 입력 장치는 음성 서비스(예를 들어, Alexa®를 포함한다. 대안적으로 또는 추가적으로, 음성 입력 장치는 스마트 워치(smart watch)의 일부 및/또는 마이크로폰을 가진 웨어러블 수단(wearable means having a microphone) 및/또는 개인용 모니터(personal monitor; 셀룰러(cellular) 및/또는 휴대용(portable))의 일부를 구성한다. 일부 실시예에서, 음성 샘플링 및/또는 분석 모듈은 마이크로폰을 갖는 패닉 알람 버튼(panic alarm button)의 일부를 구성하여, 선택적으로 개인 애플리케이션(private application)이거나 의료 응급 서비스와 관련될 수 있다.
- [0213] 일부 실시예에서, 음성 샘플링은 대상에 의해 개시된다. 대안적으로 또는 추가적으로, 음성 샘플링은, 선택적으로, 예정된 일정에 따라, 예를 들어 5분마다, 1 시간에 1회, 하루에 2번, 더 또는 덜 빈번한 계획(regimen)에 따라 대상에 대한 음성 샘플을 얻는 것과 같이, 음성 입력 장치를 갖는 샘플링 장치에 의해 개시된다. 일부 실시예에서, 복수의 음성 샘플이 획득되고 시간이 지남에 따라 분석된다. 복수의 테스트를 획득하는 것의 잠재적 이점은 테스트 결과 및 진단의 특이성(specificity) 및/또는 민감도(sensitivity)의 개선이다. 일부 실시예에서, 음성 샘플의 수 및/또는 샘플링 사이의 시간 간격은 심장 상태를 검출하기 위해 추정된 확률에 기초하여 결정된다. 예를 들어, 일주일에 한 번 발생하는 것으로 의심되는 AF 에피소드(AF episode)를 찾을 때에는, 매 시간마다 발생하는 AF 에피소드를 찾을 때보다, 아마도 더 빈번한, 다른 테스트의 계획(regimen)이 제공된다.
- [0214] 일부 실시예에서, 프라이빗 및/또는 퍼블릭 클라우드 서버(a cloud server, private and/or public)가 음성 샘플을 분석하는데 사용된다. 일부 실시예에서, 예를 들어, 대상이 심장 상태 추정 서비스를 요청하기 위해 전화할 때, 음성이 샘플링되고 전화 스위치(telephone switch) 또는 콜 센터(call center)에서 적어도 부분적으로 분석된다. 일부 실시예에서, 서버는 전화 스위치와 함께 배치되어 스위치로부터 데이터를 얻는다(a server is collocated with the telephone switch and gets the data from the switch).
- [0215] 대안적으로, 대상은 다른 서비스(difference service)를 요청하기 위해 전화하지만, 동일한 전화에 의해 심장 상태 추정 서비스를 제공 받는다(a subject calls to seek a difference service, but receives cardiac condition estimation service by the same call). 선택적으로, 분석 결과가 대상에게 전송된다. 대안적으로

또는 추가적으로, 이들은 간병인에게 보내진다. 대안적으로 또는 추가적으로, 이들은 요청되거나 비용이 지불 될 때까지 데이터베이스에 저장된다. 일부 실시예에서, 데이터베이스 및/또는 영구 메모리(permanent memory)는 분석된 음성 샘플(voiced samples)에 대한 참조(reference)로 사용된다.

[0216] 일부 실시예에서, 분석은 스마트 폰, 및/또는 CPU 기반 장치 및/또는 서버의 회로 및/또는 소프트웨어에 (circuitry and/or software in a smartphone, and/or CPU-based device, and/or server) 의해 수행된다. 일부 실시예에서, 회로는 음성 샘플을 처리하고 스피커의 심장 상태를 추정하기 위한 인스트럭션을 갖는다. 기존 인프라(Infrastructure)를 수동적으로 사용하는 것의 잠재적 이점은 심장 질환에 대한 수동 모니터링을 수행할 수 있다는 것으로, 잠재적으로 샘플 대상의 적극적인 조치가 취해지지 않은 경우에도, 선택적으로 잠재적인 우려가 있는 경우 추가로 검진 대상에 대해 경고 표시 및/또는 권장 사항을 제공 할 수 있다. 또 다른 잠재적인 이점은 일상 생활에서 테스트의 접근성으로 인해 조기 탐지 가능성이 높다는 것이다. 조기 발견은 뇌졸중 및 기타 합병증을 예방하여 재 입원(re-hospitalizations) 및/또는 사망을 예방할 수 있는 조기 개입을 제공 할 수 있습니다.

[0217] 일부 실시예에서, 전화 대화(telephone conversation)는 제 1 스테이션에서 암호화되고 제 2 스테이션에서 해독 된다. 이러한 실시예에서, 스피치의 분석은 선택적으로 제 1 스테이션 또는 제 2 스테이션에서 회로 및/또는 소프트웨어에 의해 수행되고, 선택적으로 암호화되지 않은 스피치를 분석한다.

[0218] 일부 실시예에서, 영향의 특성화 파라미터(a characterizing parameter of an effect)는 임계 값 또는 기준 (reference)과 비교된다. 일부 실시예에서, 기준 값(reference value)은 동일한 대상으로부터 얻어지지만 건강 및/또는 병적인(healthy and/or pathologic) 것으로 알려진 샘플에서 얻어진다. 영향을 자기 기준(self-reference)과 비교할 때 얻을 수 있는 잠재적인 이점은 샘플링된 개체(sampled individual)에서 발생하는 변화를 식별하고 선택적으로 이러한 변화에 따라 심장 상태를 결정할 수 있다는 것이다. 대안적으로 또는 추가적으로, 기준은 건강 및/또는 심장 병을 가진(healthy and/or as having a cardiac pathology) 것과 관련된 복수의 기준을 갖는 데이터베이스로부터 제공된다. 대안적으로 또는 추가적으로, 특성화 파라미터(a characterizing parameter)는 심장 상태를 결정하기 위해 컷오프(cut-off)로 미리 결정된 임계 값과 비교되는데, 예를 들어 임계 값보다 낮은 것이 건강 상태로 결정될 것이고, 더 높은 것이 병적 상태(pathologic condition)로 결정될 것이다. 일부 실시예에서, 컷오프 값은 대상체가 병리를 가질 확률을 계산하기 위해 사용될 수 있다(a cutoff value can be used to calculate the probability of the subject to have pathology).

[0219] 일부 실시예에서, 임계 값 또는 다차원 임계 값(a multi-dimensional threshold)은 선택적으로 음성 샘플로부터 도출된 다른 데이터를 사용하여 기계 학습에 기초하여 동적으로 결정된다. 일부 실시예에서, 트레이닝 세트(training set)는 병적 및 비 병적 환자로부터 계산된 음성 특징(voice features calculated from other patients with the pathology and without)을 포함한다. 일부 실시예에서, 분류기(a classifier)는 트레이닝 세트의 최적 분리 및 별도의 테스트 세트에서의 높은 분리를 위해 선택된다(a classifier is chosen for optimal separation of the training set and high separation on a separate test set). 예를 들어, 이러한 방법에는 SVM, 회귀, PCA KNN 등(SVM, regression, PCA KNN and more)이 포함될 수 있다.

[0220] 일부 실시예에서, 여러 음성 샘플 분석 방법의 통합이 수행된다. 일부 실시예에서, 선택적으로 분석 결과의 특이성 및/또는 민감도에 기초하여, 상이한 가중치(weight)가 각각의 분석 방법에 할당된다. 일부 실시예에서, 기계 학습은 각 분석의 가중치를 동적으로 결정하기 위해 사용된다. 선택적으로, 대상의 병력 및/또는 가족 병력 및/또는 생리학 적 검사 결과와 같은 다른 데이터가 통합 과정에 이용된다.

[0221] 일부 실시예에서, 선택적으로, 통합된 결과에 기초하여 부정맥 심장 상태의 심각도(a severity of an arrhythmic cardiac condition)이 결정된다. 선택적으로, 부정맥의 심각도는 식별된 에피소드 수, 예를 들어 식별 된 AF 에피소드 수에 따라 결정된다. 대안적으로 또는 추가적으로, 심각도는 식별된 에피소드 사이의 시간 간격 및/또는 그들의 지속 기간에 기초하여 결정된다. 대안적으로 또는 추가적으로, 심각도는 음성 샘플에 대해 결정된 영향의 정도(the extent of the determined effect over the voice sample)에 기초하여 결정되며, 예를 들어 음성 샘플이 불규칙할수록, 병리학 적 상태가 더 심각하게 결정된다(the more irregular a voice sample is, the more severe the pathologic condition which is determined).

[0222] 일부 실시예에서, 심장 상태를 결정한 후, 검증 테스트(validation test)가 수행될 수 있는데, 이는 선택적으로, 제 2 음성 샘플의 형태로 수행된다(a validation test is performed, optionally in the form of a second voice sample). 일부 실시예에서, 제 2 음성 샘플은 미리 결정된 단어 또는 모음(predetermined words or vowels)을 포함한다. 대안적으로 또는 추가적으로, 검증 테스트는 심전도 검사

(electrocardiogram(ECG) test) 및/또는 광혈류측정 검사(photoplethysmography(PPG) test) 및/또는 ECG 및/또는 PPG 테스트를 수행하기 위한 지정된 모바일 애플리케이션(designated mobile applications for performing ECG and/or PPG tests)을 포함한다. 보완 검사(complement testing)의 잠재적인 인들은 진단의 특이성과 민감도가 높아지는 것이다.

[0223] 일부 실시예에서, 음성을 통한 혈류의 영향은 심박수 및/또는 심박수의 변동 또는 변화(heart rate and/or variability or change in the heart rate)를 식별하는데 이용되며, 이는 예를 들어 스트레스, 분노, 신경질, 흥분 등을 식별하기 위해 이용된다. 대안적으로 또는 추가적으로, 음성 분석은 샘플링된 대상의 체력(the fitness level of the sampled subject) 수준을 결정하기 위해 사용된다. 대안적으로 또는 추가적으로, 음성 분석은 샘플 대상의 일반적인 건강 및/또는 임상 상태(a general health and/or a clinical state of the sampled subject)를 결정하기 위해 사용된다.

[0224] **음성에 대한 심장 활동의 타이밍의 영향 추출(Extracting effects of the timing of the cardiac activity on the voice)**

[0225] 본 발명의 일부 실시의 양태는, 심장 활동의 타이밍이 음성에 미치는 영향(the effect the timing of the cardiac activity has on the voice)을 검출하는 것, 예를 들어, 음성에 영향을 미치는 혈류에 대한 심장 펄스 타이밍의 영향을 감지하는 것(by detecting the effect of the timing of the heart pulses on the blood flow which affects the voice)에 기초하여 심장 상태를 결정하는 것에 관한 것이다. 불규칙한 심장 박동수(irregular heart pulse rate)는 심장 펄스 영향의 불규칙한 발생(irregular occurrences of heart pulse effects)을 가지고 있어서 음성에 나타나는 혈류에 영향을 미칠 것으로 추정된다.

[0226] 일부 실시예에서, 시간별 심장 데이터(timed cardiac data)는 음성 샘플의 유성 세그먼트(voiced segments)로부터 추출된다. 예를 들어, 시간별 심장 데이터는 적어도 하나의 RR 간격(RR interval)을 포함한다. 선택적으로, R 정점(R apexe)는 연속적이지 않다(non-successive). 본원에 사용된 용어 "R"은 심장 맥박 동안 혈압 피크의 타이밍(the timing of the peak of the blood pressure during heart pulse)을 지칭한다. RR은 두 개의 연속 심장 펄스 사이의 시간 거리(the distance in time between two successive heart pulses)이다.

[0227] 일부 실시예에서, 심장 파라미터(cardiac parameters)는 자발적 스피치(spontaneous speech)로부터 추출된다. 자발적 스피치는 일반적으로 유성음, 무성음 및 무음(voiced sounds, unvoiced sounds and silence)으로 조합된다. 본 명세서에서 사용되는 바와 같이, 유성음(voiced sounds)은 "cat"에서와 같이 #a# 음소(phoneme)와 같이 성대가 닫히고 진동하는 음성 부분이고, 무성음(unvoiced sounds)은 "shame"의 #sh# 음소와 같이 성대가 열린 음성의 부분이고, 무음 기간(silence period)은 말이 없는 부분 (일시 정지 등)이다. 심장 활동은 주로 유성음에 영향을 미치는 것으로 추정되며, 무성음은 그 영향을 덜 받고, 무음 기간은 전혀 영향을 받지 않는다. 자발적 스피치 중 유성 세그먼트는 비교적 짧으며, 대개 3초 미만 지속된다. 일부 실시예에서, 심장 상태는 비연속적이고, 단편화된 스피치 데이터(non-continuous, fragmented speech data)로부터 추정된다. 선택적으로, 단편화된 스피치 데이터는 자발적 스피치의 유성 세그먼트로부터 추출된다. 일부 실시예에서, 음성 샘플은 유성 세그먼트를 식별하도록 분류되고, 선택적으로, 유성 세그먼트는 적어도 일부 세그먼트가 제거되고 나머지 세그먼트 엔드(segment ends)가 평탄화되도록(smoothed) 연결된다(the voiced segments are concatenated such that at least some segments are removed and remaining segment ends are smoothed).

[0228] 일부 실시예에서, 유성음은, 예를 들어, 본 명세서에 그 전체가 레퍼런스로 포함된 Yingyong Qi and Bobby R. Hunt, Voiced-Unvoiced-Silence Classifications of Speech Using Hybrid Features and a Network Classifier, Yingyong Qi and Bobby R. Hunt, 1993에 언급된 것과 같은 표준 음성 분류 방법(standard speech classification methods)에 의해 분류된다.

[0229] 일부 실시예에서, 심장 펄스 데이터(heart pulse data)는 간헐적 인 유성 세그먼트로(intermittent voiced segments)부터 도출된다. 일부 실시예에서, 심장 펄스 데이터 및/또는 혈류 행동(blood flow behavior)은, 선택적으로, 펄스 발생에 대한 암시적 경향(suggestive trends for the occurrence of a pulse)을 갖는 두 개의 유성 세그먼트 사이에서 발견되는 유성이 아닌 세그먼트(segments which are not voiced)로부터 외삽된다(extrapolated). 일부 실시예에서, 두 개의 R 펄스 발생 사이의 시간 간격(a time interval between two occurrences of R pulses)은 심장 상태를 결정하기에 충분하다(예를 들어, 심실 빈맥과 같은 비정상적으로 높은 심박수 조건을 특징 짓는 시간 간격(a time interval characterizing an abnormally high heart rate condition)을 식별할 때). 대안적으로 또는 추가적으로, 적어도 세 개의 R 펄스 발생은 AF와 같은 부정맥을 결정하기 위해 충분하다. 일부 실시예에서, 부정맥 심장 상태는 심장 펄스의 세 번의 연속적인 발생 사이의 시간

간격을 계산함으로써 결정된다. 일부 실시예에서, 더 많은 수의 연속적인 R 펄스 발생은 테스트의 더 높은 민감도 및/또는 특이성을 제공하기 위해 필요하다(예를 들어, 4 초과, 및/또는 5 초과, 및/또는 6 초과, 및/또는 7 초과, 및/또는 8 초과 및/또는 더 많은 수의 연속 발생). 일부 실시 양태에서, 식별된 심장 펄스 사이의 시간 간격은 건강 또는 병적 및/또는 부정맥인 것으로 알려진 기준 심장 상태(reference heart condition)로부터 획득된 기준 시간 간격과 매칭된다. 선택적으로, 매칭 확률이 계산되고, 심장 상태를 결정하기 위한 임계 값과 비교된다.

[0230] 본 발명의 일부 실시예의 한 측면은 시간이 지남에 따라 음성으로부터 관련 정보를 전달하는 특정 파라미터의 변형(variations of specific parameters that carry relevant information from the voice over time)을 찾아서 심장 상태를 추정하는 것에 관한 것이다. 예를 들어, 시간이 지남에 따라 음성 특징을 분석하고 음성 특징의 값의 주기성(periodicity of the values of the voice features)을 계산한다. 일부 실시예에서, 음성 특징은 유성 샘플(voiced sample), 선택적으로, 자발적 스피치로부터 추출된다. 일부 실시예에서, 음성 특징들은, 예를 들어 가중 스펙트럼(weighted spectrum), 및/또는 선형 예측 계수(Linear Predictive Coefficient (LPC)) 및/또는 LPC 기반 스펙트럼(LPC based spectrum), 및/또는 멜 주파수 Cepstral 계수(Mel Frequency Cepstral Coefficients(MFCC)), 및/또는 기본 주파수(fundamental frequency(pitch))를 및/또는 에너지(energy), 및/또는 제로 크로싱(zero crossing) 및/또는 포먼트(formants), 및/또는 성문 펄스(glottal pulse(vocal cord pulse)), 및/또는 지터(jitter) 및/또는 쉬머(shimmer), 및/또는 프랙탈 치수(fractal dimension), 및/또는 일관성(coherence), 및/또는 웨이블릿 분석(wavelet analysis), 또는 스피치 샘플의 다른 수학적/통계적 표현을 포함한다.

[0231] 일부 실시예에서, 선택적으로, 그의 음성 샘플을 분석함으로써, 대상의 심박수가 추정된다. 일부 실시예에서, 음성 특징의 불균일성(a non-uniformity of the voice feature)은, 예를 들어 심장 박동수의 주파수 주위의 미리 결정된 범위의 주파수에서 주기성(periodicity at frequencies at a predetermined range around the frequency of the heart rate)을 식별함으로써, 심장 활동의 타이밍에서 불규칙성(irregularities in the timing of the cardiac activity)을 식별하는데 이용된다. 일부 실시예에서, 스펙트럼 분석(spectral analysis) 및/또는 자기 상관(autocorrelation)은 음성 샘플에서 주기적(periodic) 및/또는 반주기적(semi-periodic) 변화를 식별하기 위해 이용된다. 일부 실시예에서, 주기성은 음성 특징의 미리 결정된 심박수 범위에서 스펙트럼 피크의 대역폭(a band width of a spectral peak at the predetermined range of the heart rate)에서 계산된다. 일반적으로 대역폭이 넓을수록 주기성이 낮으므로, 부정맥에 대한 확률이 높다. 일부 실시예에서, 심장 상태를 결정하기 위해, 대역폭은 미리 결정된 임계 값과 비교된다.

[0232] 일부 실시예에서, 주기성의 특성화 파라미터는 심장 상태를 결정하기 위해 임계 값과 비교된다. 예를 들어, 심박수의 주파수 주위의 (피치와 같은 음성 특징의) 자기 상관 함수의 피크(a peak of an autocorrelation function)는 그 대역 폭에 의해 특징지어질 수 있고, 미리 정해진 임계 값보다 높은 값을 갖는 자기 상관 함수의 대역폭은 부정맥 심장 질환의 가능성이 높다.

[0233] 일부 실시예에서, 스피치의 스펙트럼 상호 간섭(spectral cross-coherence of the speech)은 스피치의 세그먼트들 사이에서, 선택적으로 피치 및/또는 포먼트 주파수 주위 및/또는 심장 펄스에 의해 잠재적으로 영향을 받는 임의의 주파수 주위에서 계산된다. 짧은 시간 동안 더 낮은 값에 도달하는 일관성(Coherence reaching lower values for a short period of time)은 심장 펄스의 표시 일 수 있다. 이러한 방식으로, 심장 펄스는 스피치 타임 라인(the speech time line) 상에 위치 될 수 있다.

[0234] **음성에 대한 심장 활동의 크기의 영향 추출(Extracting effects of the magnitude of the cardiac activity on the voice)**

[0235] 본 발명의 일부 실시예의 양태는 음성에 대한 심장 활동의 크기의 영향(the effect of the magnitude of the cardiac activity on the voice)을 추출하는 것에 기초하여 심장 상태를 결정하는 것에 관한 것이다. 본원에 사용된 용어 "크기(magnitude)"는 혈압(blood pressure) 및/또는 혈관 부피(blood vessels volume)와 및/또는 심실 및/또는 심방 박동의 영향과 같은 혈액의 기계적 효과의 정도(extent of the mechanical effect of the blood)을 지칭한다. 불규칙한 심장 박동수는 보컬 영역(vocal region)으로의 혈류 크기의 변동을 초래할 수 있으며, 따라서 음성에 미치는 영향의 크기의 변동을 초래할 것으로 추정된다. 심장 박동(heartbeat)은 음성을 변조하여 음향 신호(acoustic signal)의 변화를 초래하는 것으로 추정된다. 따라서, 예를 들어 동리듬(sinus rhythm)을 특징으로 하는 규칙적인 심장 펄스 웨이브(regular heart pulse wave)는 주기적 변조(periodic modulation)을 야기하며, 아마도 음성 기관(speech organs)을 통과하는 혈류의 주기적인 변화에 기

인 할 것으로 추정된다. 반면, 불규칙한 심장 박동, 예를 들어 AF는 혈류의 혼란스러운 변화를 일으키고 음향 신호의 큰 변화를 유발한다.

[0236] 일부 실시예에서, 음성 특징의 값의 분포(the distribution of the values of the voice feature), 예를 들어 표준 편차(the standard deviation)가 결정된다. 일부 실시예에서, 분포의 형태(the shape of the distribution)의 특성화 파라미터(a characterizing parameter)는 심장 상태를 결정하기 위해 임계 값과 비교된다. 예를 들어, 분포 형태 및/또는 스펙트럼 피크 값(the spectral peak values)의 큰 폭은 미리 결정된 임계 값과 비교 될 수 있고, 이는 부정맥 심장 상태에 대한 높은 확률과 연관이 있다.

[0237] 일부 실시예에서, (여러 특징들을 결합하는) 다중 특징 분류기(a multi-feature classifier)가 선택적으로 이용되고, 음성 특징들의 값의 다차원 분포(multi-dimensional distribution)에 대한 선택적 다차원 임계 값(optionally multi-dimensional threshold)은, 예를 들어, Support Vector Machine (SVM) 방법, 및/또는 K-MEANS clustering analysis와 같은 Vector Quantization methods을 이용하여 결정된다.

[0238] 일부 실시예에서, 다차원 분포의 형태의 특성화 파라미터는 다차원 임계 값과 비교되어 심장 상태를 결정한다.

[0239] 심방 세동 검출(Detecting Atrial Fibrillation)

[0240] 본 발명의 일부 실시예의 한 측면은 환자의 음성에 대한 영향에 의해 심방 세동을 검출하는 것에 관한 것이다.

[0241] 일부 실시예에서, 만성 병리적 상태(chronic pathological conditions)는 음성 파라미터의 변화(changes in voice parameters)에 의해 잠재적으로 검출된다. 어떤 경우에는 심방 세동(AF)이 좌심방(Left Atrium)의 확장을 유발한다(대부분의 병리적 AF 사례에서 발생). 팽창은 반회 신경(또는 후두신경; recurrent laryngeal nerve)에 영향을 미쳐, 지속적으로 음성을 변화시킨다. 변화는 피치 신호의 파라미터(the parameters of the pitch signals)에서 잠재적으로 나타나며, 분석 및 검출 될 수 있다.

[0242] 이를 통해 심장이 정상 속도로 뛰는 경우에도 만성 질환이 있는 AF 환자를 검출할 수 있다.

[0243] 본 발명의 적어도 하나의 실시예를 상세하게 설명하기 전에, 본 발명은 명세서의 세부사항이나 이하의 설명 및/또는 도면 및/또는 예제에 설명된 구성 요소 및/또는 방법의 배치에 국한되지 않는다는 것을 이해해야 한다. 본 발명은 다른 실시예가 있을 수 있으며, 다양한 방식으로 실시 또는 수행 될 수 있다.

[0244] 예시적인 음성 샘플링 및 분석(Exemplary voice sampling and analysis)

[0245] 이제 도면을 참조하면, 도 1a는 본 발명의 일부 실시예에 따라, 사람의 음성으로부터 심장 상태를 결정하는 예시를 나타내는 상위 레벨 블록도 및 흐름도이며, 구성 요소 및 그들 사이의 데이터 흐름을 도시한다.

[0246] 일부 실시예에서, 아날로그 음성 샘플(analog voice sample; 1)은 하나 이상의 말하는 대상(speaking subject)를 포함 할 수 있는 음성 출력(voice output; 102)에 의해 생성된다. 일부 실시예에서, 아날로그 음성 샘플(1)은 음성 입력 장치(voice input; 104), 예를 들어 마이크를 갖는 전자 장치에 의해 샘플링되어 디지털 음성 샘플(digital voice sample; 3)으로 변환된다. 일부 실시예에서, 디지털 음성 샘플(3)은 음성 분석기(voice analyzer; 106)에 의해 처리된다. 실시예에서, 분석기(106)에 의한 처리는 음성 특징(voice features; 5a) 및/또는 심장 특징(cardia features; 5b)을 식별한다. 예를 들어, 심장 특징은 RR 간격(RR intervals) 및/또는 심박수(heart rate) 및/또는 심박수 변동성(heart rate variability)을 포함할 수 있다. 음성 특징들에 대한 예는 피치(pitch), 및/또는 포먼트(formantss), 및/또는 MFCC (Mel Frequency Cepstral Coefficients) 및/또는 프레임 사이 교차 일관성 값(cross coherence values between frames) 및/또는 웨이블릿 기반 특징(wavelet based features), 및/또는 상기 파라미터 중 하나 이상의 스펙트럼 대역 에너지(spectral band energies) 및 대역폭(band width), 및/또는 발화 중 변동(fluctuation during utterance) 및/또는 그러한 특징의 미분(derivatives)을 포함 할 수 있다.

[0247] 스피치 특징(speech feature)으로서 교차 일관성(Cross Coherence): 시간 (t)에서 음성 샘플의 특정 프레임에 대해 - 교차 일관성은 시간의 프레임 (t + Δn)으로 계산된다. 이 값은 선택적으로 N개의 값을 포함하는 벡터에 저장된다.

[0248]
$$\text{CohVec}(t) = [\text{CrossCoh}(\text{frame}(t), \text{frame}(t + \Delta)),$$

[0249]
$$\text{CrossCoh}(\text{fame}(t), \text{frame}(t + \Delta)),$$

[0250] ...

- [0251] $CrossCoh(fame(t), frame(t+\Delta))$]
- [0252] 특징으로서 웨이블릿(Wavelet):
- [0253] 웨이블릿 변환은 여러 가지 방법으로 스피치 특징으로 선택적으로 이용된다.
- [0254] 제1 방식으로, 일부 실시예에서, 선택적으로, 고역 통과 필터(high pass filter(HPF)) 및/또는 저역 통과 필터(low pass filter(LPF)) 분해를 이용하여 웨이블릿 분해가 이용된다.
- [0255] 이제 도 1을 참조한다. 도 1b는 본 발명의 일부 실시예에 따른 이산 웨이블릿 변환(Discrete Wavelet Transform(DWT))에 의한 웨이블릿 분해의 단순화 된 예시이다.
- [0256] 도 1b는 레벨 1 DWT 계수(Level 1 DWT coefficients) $h(n)$ (124) 및 $g(n)$ (122)를 생성하는 신호 $x(n)$ (120)을 도시한다.
- [0257] 도 1b는 또한, 선택적으로, 레벨 2 DWT 계수 $h(n)$ (126) 및 $g(n)$ (128)을 생성하는 레벨 1 DWT 계수 $g(n)$ (122)를 도시한다.
- [0258] 도 1b는 각 필터의 출력 에너지(filter's output energy)가 출력 특징 벡터의 구성 요소(component of the output feature vector)로서 사용되는 예시적인 실시 예를 도시한다.
- [0259] 제2 방식으로, 일부 실시예에서, 스펙트럼을 사용하는 에너지 및 주파수와 유사하게 각 프레임에 대한 "스케일(Scale)" 및 "계수(coefficient)"값을 계산하며, 이 값을 특징(features)로 이용한다(In a second way, in some embodiments, "Scale" and "coefficient" values for each frame are optionally calculated, similarly to energy and frequency using spectrum, and using the values as features.).
- [0260] 스펙트럼 에너지(Spectral energy) & 변동 특징의 BW(BW of fluctuation feature):
- [0261] 일부 실시예에서, 다음 두 단계 계산이 수행된다:
- [0262] 각 프레임마다, 발화의 특징 벡터(feature vector of the speech utterance)를 계산하는 단계; 및
- [0263] 스펙트럼 추정 방법(spectral estimation methods)을 이용하여 특징 벡터의 스펙트럼을 계산하는 단계.
- [0264] 이에 한정되는 것은 아니지만, 일부 예에서, 특징 벡터가 피치 주파수(프레임 당) 인 경우:
- [0265] $FtrVec(n)=Pitch(n)$, $n=1,2,\dots$ where N is a frame number,
- [0266] The spectrum is:
- [0267] $SpecVec= Spectrum_Estimation(FtrVec(1)\dots FtrVec(n))$.
- [0268] 일부 실시예에서, SpecVec은 특징 벡터가 시간에 따라 변하는 빈도를 나타 내기 위해 선택적으로 사용된다. 심장 박동이 음성 신호(the voice signal)의 일부 기능을 변조하므로 SpecVec은 잠재적으로 심장 박동수 주파수에서 주파수 피크(frequency peak at a heart rate frequency)를 보여준다.
- [0269] 일부 실시예에서, 스펙트럼 추정(spectrum estimation)은 선택적으로 다음과 같은 방법에 의해 수행된다:
- [0270] 푸리에 변환(a Fourier transform);
- [0271] 공분산 법(a covariance method);
- [0272] 주기도 법(a periodogram method);
- [0273] 율-워커 방법(a Yule-Walker method);
- [0274] 다중 시그널 분류 방법(a Multiple Signal Classification(MUSIC) method)); 및
- [0275] 고유 벡터 방법(an eigenvector method).
- [0276] 일부 실시예에서, 스펙트럼 추정 방법은 선택적으로 발화당 다음 몇 개의 값(produces several values per utterance)을 생성한다:
- [0277] 스펙트럼 피크의 위치(Location(s) of spectral peak(s)) [Hz];
- [0278] 스펙트럼 피크의 대역폭(Band-width(s) of spectral peak(s)) [Hz]; 및

- [0279] 스펙트럼 피크 에너지, 선택적으로 대역폭 * 높이(Energy of spectral peak(s), optionally Band-width * height) [Watt].
- [0280] 상기 값은 선택적으로 심박수 상태 표시(heart rate status)로 사용된다.
- [0281] 일부 실시예에서, 심장/음성 특징은 심장 상태(cardiac condition; 7)를 추정하기 위해 심장 상태 분류기(cardiac condition classifier; 108)에 의해 분석된다. 예를 들어, 심장 상태는 건강한 심장 상태(healthy cardiac condition), 또는 병적 심장 상태(pathologic cardiac condition, 예를 들어 부정맥 및/또는 AF로 식별된)일 수 있다. 선택적으로, 심장 상태(7)는 예를 들어, 스크린(screen)과 같은 심장 상태 출력(cardiac condition output; 110)에 의해 제시된다.
- [0282] 일부 실시예에서, 음성 출력(102)은 아날로그 음성 샘플(1)을 생성하는 인간 대상(human subject)을 포함한다. 인간 대상은 음성을 이용한 심장 상태 진단의 후보인 임의의 사람 일 수 있다. 선택적으로, 음성 출력(102)은 하나 이상의 인간 대상의 음성 샘플(1)을 생성하는 것(예를 들어, 음성 샘플(1)이 둘 이상의 대상들 사이의 대화를 포함할 때)를 포함한다. 일부 실시예에서, 식별된 화자(speaker)만이 분석된다. 일부 실시예에서, 음성 샘플(1)은 자발적 스피치(spontaneous speech), 예를 들어, 사용자 스피킹 세그먼트(segments of user speaking), 및/또는 대화하는 사용자(users conversing) 및/또는 침묵 기간(silence periods), 및/또는 배경 잡음(background noise) 및/또는 배경 화자(background speakers)를 포함 할 수 있는 시퀀스(sequence)를 포함한다. 대안적으로 또는 추가적으로, 음성 샘플(1)은 선택적으로, 미리 결정되거나 룩업 테이블(look-up table)로부터 선택된 특정 단어 및/또는 모음(vowels)과 같은 제한된 음성을 포함한다.
- [0283] 일부 실시예에서, 음성 입력 장치(104)은 아날로그 음성 샘플(1)을 디지털 음성 샘플(3)로 변환함으로써 아날로그 음성 샘플(1)을 레코드(record)한다. 일부 실시예에서, 음성 입력 장치(104)는 마이크로폰을 갖는 장치이다. 예를 들어, 음성 입력 장치(104)는 유선 전화(land-line phone), 및/또는 휴대 전화(cellular phone), 및/또는 컴퓨터(computer), 및/또는 스피커폰(speaker-phone) 및/또는 음성 녹음기(voice recorder)일 수 있다. 대안적으로 또는 추가적으로, 음성 입력 장치(104)는 디지털 음성 샘플(3)을 레코드하는 서버(선택적으로, 원격 서버(remote server), 예를 들어 콜 센터(call center)에 위치된 서버)이다.
- [0284] 선택적으로, 음성 입력 장치(104)는 디지털 음성 샘플(3)을 저장하기 위한 데이터베이스를 포함한다. 대안적으로 또는 추가로, 음성 입력 장치(104)는 사운드를 장기적으로 저장하지 않고 단지 사운드를 변환하여 전송한다. 일부 실시예에서, 디지털 음성 샘플(3)은 음성 데이터 포맷, 예를 들어 PCM, 및/또는 WAV, 및/또는 ADPCM으로, 저장된다. 대안적으로 또는 추가로, 아날로그 사운드는 디지털 포맷으로 변환되지 않지만 아날로그 분석 회로(analog analysis circuitry)에 의해 분석 된 후 이를 디지털 포맷으로 변환한다.
- [0285] 일부 실시예에서, 음성 입력 장치(104)에 의한 샘플링은 자동으로, 즉 샘플링과 관련하여 대상이 적극적으로 수행한 활동없이 자동으로 제공된다. 대안 적으로 또는 추가적으로, 샘플링은 수동적으로 제공된다. 예를 들어, 전용 콜 센터에 능동적으로 전화를 걸거나, 또는 장치상의 기록 버튼을 누름으로써, 대상의 능동적 행동에 의해 작동함으로써 샘플링이 수동으로 제공된다. 대안적으로 또는 추가로, 샘플링은 사전에 예약된 구성(pre-scheduled configuration)으로 제공되며, 선택적으로, 사용자는 특정 단어 또는 소리를 말하거나/받음하도록 요청 받는다. 예를 들어, 서비스는 미리 예약 된 시간에 사용자에게 전화를 걸고 대화를 샘플링 할 수 있다. 대안적으로, 대상은 자유 시간에 디지털 음성 샘플(3)을 레코드한 다음 이 음성 샘플을 예를 들어 서버에 업로드 함으로써 서비스로 전송할 수 있다.
- [0286] 일부 실시예에서, 사전 스케줄링된 동작(pre-scheduled operation)은 가능한 위험(possible risk)이 검출된 후 모니터링 인프라(monitors infrastructure) 및/또는 모니터링 애플리케이션(monitors application), 및/또는 의료 지원(medical assistance) 및/또는 소프트웨어 애플리케이션(software application)으로부터 초기화 될 수 있다. 일부 실시예에서, 사전 스케줄링된 동작은 주기적 검사(periodic screening)를 포함한다. 주기적 인 검사의 잠재적 이점은 대상의 상태를 지속적으로 모니터링하는 것이다. 일부 실시예에서, 미리 스케줄링된 동작은, 선택적으로, 통지를 전송함으로써 특정 음성을 발음하도록 사람을 유도하는 단계를 포함한다. 일부 실시예에서, 통지는 구두 및/또는 문자 경고, 및/또는 비프 음, 및/또는 스마트 폰으로의 문자 메시지 및/또는 음성 지시를 포함한다. 일부 실시예에서, 사용자는 시스템의 수동 활성화를 사용하여 본인 및/또는 그의 친척들(relatives)에 대한 테스트를 개시한다. 선택적으로, 사용자의 동의에 따라, 일부 실시예에서, 장치는 마이크로폰을 자동으로 활성화 시키도록 프로그램될 수 있다. 일부 실시예에서, 장치는 예를 들어, 건강 관리 서비스를 통해 원격 위치로부터 활성화 될 수 있다.

- [0287] 일부 실시예에서, 디지털 샘플(3)은 음성 분석기(106)에서 분석하기 전에 분류된다. 예를 들어, 스위치/서버 애플리케이션(switch/server applications, 매스 센트럴 스크리닝(mass central screening))에 대해 선택된 콜(call)만이 분석될 수 있다. 선택적으로, 샘플 특성(sample characteristics, 예를 들어, 통화 지속 시간 및/또는 잡음 레벨, 및/또는 특정 화자의 음성 세그먼트의 지속 시간, 및/또는 음성 세그먼트의 지속 기간, 및/또는 화자 식별 데이터 및/또는 진단의 민감성 및/또는 특이성에 영향을 줄 수 있는 다른 기준)에 기초하여 선택이 이루어진다.
- [0288] 일부 실시예에서, 음성 분석기(106)는 (speaker) 식별을 수행한다. 화자 식별은 여러 가지 이유(분석 결과를 누구에게 보내야 하는지를 알기 위해와 같은 이유)로 중요 할 수 있다. 대안적으로 또는 추가로, 화자 식별은 화자의 개인 정보와 연관 될 수 있으며, 이는 임상 진단을 보조하기 위해 선택적으로 사용된다. 예를 들어, 개인 정보는 체중 및/또는 흡연 습관, 및/또는 이전에 알려진 병리 및/또는 이전의 의료 진단을 포함 할 수 있다. 일부 실시예에서, 사용자 식별은 복수의 화자로부터 및/또는 음성 샘플이 참가자들 사이에서 번갈아 나타나는 대화로부터 단일 화자를 식별하는 것을 돕는다. 선택적으로, 식별 후, 음성 분석기(106)는 대상의 음성 세그먼트(voice segment)를 다른 화자의 음성 세그먼트로부터 분리한다.
- [0289] 일부 실시예에서, 식별된 화자만이 분석된다. 선택적으로, 화자는 요청에 따라 분석된다. 대안적으로, 서비스에 대한 비용이 지불되면 화자가 분석된다. 대안적으로, 화자가 그룹의 구성원(예를 들어, 보험 또는 건강 유지 조직의 구성원)으로 식별되는 경우 분석된다.
- [0290] 선택적으로, 대상의 음성은 사전 식별된 화자가 있는 데이터베이스를 사용하여 식별된다. 일부 실시예에서, 예를 들어, 몇몇 사용자가 동일한 분석 장치를 사용하는 경우, 음성 서명을 구축 및/또는 스피치와 연관된 다른 디테일을 이용 및/또는 발신자 식별 번호(caller identification number)와 같은 다른 파라미터와 이용함으로써 사용자간 분리가 이루어질 수 있다. 대안적으로 또는 추가적으로, 사용자는 미리 결정된 단어 또는 단어 목록을 발음함으로써 식별된다.
- [0291] 일부 실시예에서, 화자 식별이 알려지지 않았지만 시스템이 심장 병리의 가능성을 발견한 경우, (통신 네트워크에서의 모니터링과 같은) 일부 애플리케이션에 대해, 선택적으로, 시스템 데이터베이스에 분석 결과가 저장된다. 그러한 경우에, 일부 실시예에서, 결과는 예를 들어, 전화 번호 및/또는 호출자 ID, 및/또는 국제 이동 장비 신원(International Mobile Equipment Identity) 등과 같은 대안적인 분류를 사용하여 저장된다. 예를 들어, 콜 센터(call center)에서 대량 스크리닝 동안 진단 결과가 시스템 데이터베이스에 저장 될 때 화자가 등록 된 경우 화자 개인 파일에 따라 저장되고, 화자를 알 수 없는 경우 발신자 정보에 따라 저장된다.
- [0292] 일부 실시예에서, 디지털 음성 샘플(3)은 화자의 심장 상태의 추정에 사용되는 것 이외로, 미래의 사용을 위해 저장된다. 선택적으로, 음성 샘플 및/또는 분석된 음성 샘플(analyzed voice samples) 및/또는 진단 결과(diagnostic results)를 포함하는 데이터베이스는 다른 음성 샘플의 향후 분석에 이용된다. 대안적으로 또는 추가적으로, 이러한 데이터베이스는, 예를 들어, 모니터링된 병리에 관한 통계적 지식을 얻기 위해 역학 연구(epidemiological research)에 사용된다.
- [0293] 일부 실시예에서, 음성 샘플(3) 처리는 음성 분석기(106)에 의해 수행된다. 예를 들어, 음성 분석기(106)은 스마트폰 회로 및/또는 원격 서버(클라우드 컴퓨팅) 및/또는 독립형 장치(standalone device, 예컨대, PC 및/또는 태블릿), 및/또는 콜 센터의 서버(예를 들어, 사람이 전화를 걸 수 있고 그의 음성이 기록되고 분석되는 경우)를 포함할 수 있다.
- [0294] 일부 실시예에서, 음성 분석기(106)는 음성 입력 장치(104)와 동일한 장치에 내장된다. 예를 들어, 스마트폰은 아날로그 음성 샘플(1)을 샘플링 하고, 이를 디지털 음성 샘플(3)로 변환하기 위한 마이크로폰(104)을 갖고, 음성을 분석하기 위한, 선택적으로, 음성 특징(5a) 및/또는 심장 특징(5b)를 도출하기 위한, 음성 분석기(106; 예를 들어, 회로 및/또는 신호 처리를 수행하는 소프트웨어)를 가질 수 있다.
- [0295] 대안적으로 또는 추가적으로, 음성 분석기(106)는 음성 입력 장치(104)와 다른 장치 및/또는 다른 지리적 위치에 있다. 예를 들어, 스마트폰은 (마이크와 같은) 음성 입력 장치(104)을 가질 수 있지만, 디지털 음성 샘플(3)은 음성 분석기(106) 회로를 갖는 원격 서버로 전송 될 수 있다.
- [0296] 대안적으로 또는 추가적으로, 음성 입력 장치(104) 및 음성 분석기(106)는 음성 출력(102)으로 부터 멀리 떨어져 있다. 예를 들어, 대상이 콜 센터에 전화를 걸고 그의 음성이 통신 네트워크에서 발견된 서버에 기록되고 그 서버에 의해 분석된다.

- [0297] 일부 실시예에서, 음성 분석기(106)는 디지털 음성 샘플(3)로부터 음성 특징 (5a)을 추출한다. 예를 들어, 음성 특징(5a)은 가중 스펙트럼(weighted spectrum) 및/또는 LPC(Linear Predictive Coefficient) 및/또는 LPC 기반 스펙트럼 및/또는 Mel Frequency Cepstral Coefficients (MFCC) 및/또는 기본 주파수(pitch) 및/또는 에너지(energy), 및/또는 제로 크로싱(zero crossing) 및/또는 포먼트(formants), 및/또는 성문 펄스(glottal pulse(vocal cord pulse)), 및/또는 지터(jitter) 및/또는 쉬머(shimmer), 및/또는 프랙탈 치수(fractal dimension), 및/또는 일관성(coherence), 및/또는 웨이블릿 분석(wavelet analysis) 또는 스피치 샘플에서 관련 정보를 추출하는 다른 특징(예를 들어, 에너지 및/또는 평균 파워(mean power) 및/또는 엔트로피 등)을 포함할 수 있다.
- [0298] 대안적으로 또는 추가적으로, 음성 분석기(106)는 디지털 음성 샘플(3) 및/또는 음성 특징(5a)으로부터 심장 특징(5b)을 추출한다. 일부 실시예에서, 심장 특징(5b)는 펄스 데이터(pulse data), 예를 들어 RR 간격(RR intervals)을 포함한다. 대안적으로 또는 추가적으로, 심장 특징은 심박수 및/또는 심박수 변동성을 포함한다.
- [0299] 일부 실시예에서, 심장 상태 분류기(108)는 음성 특징(5a) 및/또는 심장 특징(5b)을 처리하고 심장 상태(7)를 추정한다. 일부 실시예에서, 심장 상태 분류기(108)는 대상이 부정맥을 가질 확률을 결정한다. 대안적으로 또는 추가적으로, 심장 상태 분류기(108)는 대상이 심방 세동을 가질 가능성을 결정한다. 대안적으로 또는 추가적으로, 다른 심장 상태, 예를 들어 급성 관상 동맥 증후군(acute coronary syndrome) 및/또는 색 전성 뇌졸중의 경고 징후(warning signs of embolic stroke), 및/또는 심근 경색(myocardial infarction) 및/또는 갑작스런 심장 마비(sudden cardiac arrest), 및/또는 심실 플러터(ventricular flutter) 및/또는 심방 플러터(atrial flutter), 및/또는 심방 빈맥(atrial tachycardia) 및/또는 심실 빈맥(ventricular tachycardia), 및/또는 서맥(bradycardia), 및/또는 호흡 곤란(dyspnea) 및/또는 흉통(chest pain)을 결정한다.
- [0300] 선택적으로, 추정된 심장 상태(7) 및/또는 음성 특징(5a) 및/또는 심장 특징 (5b)은 심장 상태 출력(110)으로 제시된다. 일부 실시예에서, 심장 상태 출력 (110)은 스크린, 예를 들어 모니터 및/또는 장치의 디스플레이를 포함한다. 일부 실시예에서, 심장 상태 출력 (110)은 화자 및/또는 화자의 간병인(caregiver)에게 알림(notification)을 포함한다. 예를 들어, 알림은 문자 메시지 및/또는 이메일, 및/또는 전화, 및/또는 메시징 애플리케이션(예를 들어, WhatsApp, Facebook 및 사용자에게 의해 선택된 다른 수단)을 포함할 수 있다. 대안적으로 또는 추가적으로, 알림은 그래픽 표현(graphical presentation)일 수 있다. 대안적으로 또는 추가로, 알림은 시각적 및/또는 오디오 통지의 형태일 수 있다. 예를 들어, 본 발명의 일부 실시예에 따라, 차량에 내장되어 스피커 분석을 제공하는 애플리케이션이 있는 스피커폰은 깜박이는 빛, 또는 차량의 스피커 시스템을 통한 음성 알림과 같은 차량의 알림 인프라에 연결될 수 있다.
- [0301] **예시적인 데이터 흐름(Exemplary data flow)**
- [0302] 이제 도 2를 참조한다. 도 2는 본 발명의 일부 실시예에 따른 선택적 음성 분석 스테이션 사이의 일부 데이터 전송 옵션의 데이터 흐름도를 도시한다.
- [0303] 일부 실시예에서, 적어도 하나의 대상(subject; 220)은 음성 입력 장치(예를 들어, 마이크로폰을 가진 장치, 음성 레코더 및/또는 피쳐 폰(feature phone); 도 2에는 피쳐 폰(240)을 예시로 도시함)의 범위 내에서 아날로그 음성 샘플(1)을 출력한다. 대안적으로 또는 추가적으로, 음성 입력 장치는 스마트 장치, 예를 들어 퍼스널 컴퓨터, 및/또는 스마트 폰 및/또는 다른 임의의 폰 디바이스(예를 들어, 무선 전화 및/또는 스피커폰) 및/또는 태블릿의 모듈(module)일 수 있다(도 2에는 태블릿(260)을 예시로 도시함). 일부 실시예에서, (240) 및/또는 (260) 내의 음성 입력 장치는 아날로그 음성 샘플(1)을 디지털 음성 샘플(3)로 변환한다.
- [0304] 일부 실시예에서, 음성 입력 장치는 선택적으로 웨어러블 마이크로폰 (wearable microphone; 미도시)에, 선택적으로 분석을 위한 컴퓨팅 유닛(computing unit) 또는 분석을 위해 컴퓨팅 유닛으로 전송하기 위한 전송 유닛에 유선 또는 무선으로 연결된다.
- [0305] 일부 실시예에서, 디지털 샘플(3)은 원격 서버 또는 클라우드 서버일 수 있는 서버(262)와 같은 분석기로 전송된다. 대안적으로 또는 추가로, 디지털 샘플(3)은 콜 센터로 전송되며, 여기서 선택적으로, 기록되고 분석된다. 대안적으로 또는 추가적으로, 디지털 샘플(3)은 전화 스위치(telephone switch; 265)로 전송된다. 예를 들어, 디지털 샘플(3)은 콜 센터(290)로 향하기 전에 먼저 전화 스위치(265)를 통과하도록 구성된다. 예를 들어, 대상은 (290)에 위치한 정부 또는 민간 부문 고객 서비스와 통화하기 위해 전화를 걸 수 있지만, 동의하면 그의 음성은 스위치(265)를 통해 전달되어 심장 상태를 추정하기 위해 분석된다. 일부 실시예에서, 콜 센터는 분석하고자 하는 사람들을 위한 서비스를 포함한다. 대안적으로 또는 추가로, 콜 센터는 심장 상태 진단

이외의 다른 목적을 위한 서비스를 포함한다. 예를 들어, 정부 서비스를 위한 콜 센터는 진단 목적의 부가 서비스를 위한 음성 샘플링을 포함 할 수 있다. 통신 서비스의 다른 예는 건강 통신 서비스 및/또는 응급 콜 센터, 및/또는 의료 서비스, 및/또는 보험 서비스 및/또는 퍼스널 가상 어시스턴트(personal virtual assistant s)를 포함한다.

[0306] 대안적으로 또는 추가적으로, 스마트 디바이스(260)는 또한 디지털 샘플(3)을 분석하고 선택적으로, 예를 들어, 심장 및/또는 음성 특징(5) 및/또는 추정된 심장 상태(7)의 형태로 출력을 전송하기 위한 명령(instructions)을 갖는다. 일부 실시예에서, 출력은 간병인(caregiver; 280)에게 직접 전송된다. 대안적으로 또는 추가로, 추가 분석 및/또는 저장을 위해 서버(262)로 전송된다. 대안적으로 또는 추가적으로, 출력은 스위치(265)를 통과하거나 또는 통과하지 않고 콜 센터(290)로 전송되고, 선택적으로, 그로부터 간병인(280)으로 전송된다. 일부 실시예에서, 출력은 분석된 대상(220)에게 다시 제공된다.

[0307] 일부 실시예에서, 회로는 심방 세동의 현재 상태를 사용자에게 보고하도록 구성된다. 이러한 회로는 스마트 폰(240) 및/또는 클라우드 서버(262), 개인(private) 및/또는 공공(public)에서 구현 될 수 있다. 예를 들어, 기록 및/또는 분석 및/또는 전송을 위한 외부 무선 장치와 달리, 기존 인프라를 사용하는 것의 잠재적 이점은 감지(sensing) 및/또는 데이터 전송을 위한 추가 레디에이션(radiation)이 없다는 것이다.

[0308] 일부 실시예에서, 로컬 평가(local assessment)가 제공된다. 예를 들어, 동일한 장치가 음성을 샘플링하고 분석하는 경우(예를 들어, 음성 샘플링을 수행하고, 음성 특징 및/또는 심장 상태 추출을 위해 음성을 분석하는 스마트 폰 등) 로컬 평가가 제공된다. 선택적으로, 로컬 평가는 수동 조작(manual operation)으로 제공됩니다. 대안적으로 또는 추가적으로, 로컬 평가는 예를 들어 대상이 자신의 전화를 사용할 때마다 작동, 및/또는 장치에 의해 주기적으로 음성을 샘플링 함으로써 자동으로 동작한다.

[0309] 대안적으로 또는 추가적으로, 원격 평가(remote assessment)가 제공된다. 예를 들어, 음성 샘플링이 하나의 지리적 위치에서 발생하는 동안, 음성 샘플의 분석의 적어도 일부가 다른 지리적 위치에 제공될 때 원격 평가가 제공된다. 예를 들어, 대상은 전화로 샘플링 될 수 있지만, 샘플링된 음성은 원격 서버(클라우드, 셀룰러, 유선)에 의해 분석되도록 전송될 수 있다. 일부 실시예에서, BTS 및/또는 MSC 또는 서버/클라우드 및/또는 스위치와 같은 통신 네트워크 구성 요소는 분석 회로를 구비한다. 일부 실시예에서, 분석 회로는 Skype, WhatsApp, Viber, 전화 회의, 비디오 메시지 및/또는 레코딩(recordings) 등과 같은 VoIP 서버에 제공된다. 암호화된 통신의 경우, 일부 실시예에서, 심장 정보를 포함하는 관련 음성 특징은 암호화 전에 장치에서 계산되고 선택적으로, 로컬(local)에서 분석되거나, 대안 적으로 또는 추가적으로 서버에서 원격 분석(remote analysis)으로 전송된다.

[0310] 일부 실시예에서, 화자 식별은 그의 음성을 통해 대상을 식별하기 위해 사용된다. 예를 들어, 사람이 서비스를 받기 위해 콜 센터(290)에 전화 할 때, 음성 분석에 적합한 회로는 또한 대상을 식별하도록 구성될 수 있으며, 예를 들어, 선택적으로, 그를 추적하고, 검출된 상태를 통지하기 위해 사용자의 전화 번호가 이용될 수도 있다. 일부 실시예에서, 적어도 하나의 대상 또는 복수의 대상은, 예를 들어 대화 및/또는 회의 호출을 샘플링 할 때와 같이 동시에 샘플링된 복수의 음성으로부터 검출 될 수 있다.

[0311] 예시적인 상위 레벨 개요(Exemplary high level overview)

[0312] 이제도 3을 참조한다. 도 3은 본 발명의 일부 실시예에 따라 음성 샘플에 기초한 심장 상태 추정을 위한 일반적인 상위 레벨 흐름도를 도시한다.

[0313] 일부 실시예에서, 분석은, 예를 들어 5a에 도시 된 바와 같이 선택적으로, 단편화(fragmented) 및/또는 분류된(classified) 음성 샘플을 획득함으로써 단계(302)에서 시작한다. 일부 실시예에서, 음성 특징들은 단계(304)에서 선택적으로, 음성 샘플의 각 프래그먼트(fragment)마다 음성 특징 벡터(voice features vector)의 형태로 추출된다.

[0314] 일부 실시예에서, 음성 샘플 및/또는 음성 특징은 음성 특징 분포(voice feature distribution)를 계산 및/또는 임의의 수학적 연산을 계산하여 단계(306)에서 관련 정보를 드러내기 위해 사용된다. 예를 들어, 복수의 음성 특징의 값의 변동성(the variability of a plurality of values of a voice feature)과 같은 분포가 계산된다. 음성 특징은 보컬 기관(vocal organ)으로의 혈류에 의해 영향을 받는 것으로 추정되는데, 이는 펄스 타이밍(pulse timing) 및 힘(power/strength)에 의해 영향을 받는다. 따라서 부정맥에서 발견되는 것과 같은 불규칙한 펄스는 음성 특징에 대해 불규칙한 영향을 초래하여 음성 특징 값의 높은 변동성(variability)으로 나타날 것으로 추정된다. 음성 특징 분포는 도 6a 내지 도 6b, 도 7a 내지 도 7b, 도 8a 내지 도 8b, 도 9a 내지 도

9b 및 도 10a 내지 도 10b에 더 상세하게 예시된다.

- [0315] 대안적으로 또는 추가적으로, 음성 샘플 및/또는 음성 특징은 단계(308)에서 심장 펄스 발생 사이의 시간 간격(the time interval between heart pulse occurrences)을 계산하기 위해 이용된다.
- [0316] 일부 실시예에서, 음성 특징 값(voice feature values)의 일시적 변화(temporal changes)는 심장 펄스의 표시로서 이용된다. 시간 척도(time-scale)에서 일시적 변화를 찾으면 심장 맥박 위치가 근사치가 된다(Locating the temporal changes on a time-scale gives approximation of heart pulse locations).
- [0317] 일부 실시예에서, 음성 샘플은 펄스 발생을 식별하고 연속적인 발생 사이의 시간 간격(time interval between consecutive occurrences)을 계산하기 위해 사용된다. 일부 실시예에서, 복수의 시간 간격이 시퀀스(sequence)로서 합쳐진다. 일부 실시예에서, 심장 상태는 시퀀스에 기초하여 추정된다. 펄스 추출은 도 11a 내지 도 11d, 도 12a 내지 도 12b, 및 도 13에 더 상세하게 예시된다.
- [0318] 대안적으로 또는 추가적으로, 음성 샘플 및/또는 음성 특징은 단계(310)에서 심박수 주위의 음성 특징 주기성(voice feature periodicity around the heart rate)을 계산하기 위해 사용된다. 예를 들어, 음성 특징은 심박수 값에서 자기 상관 계산의 대상이 될 수 있다(For example, a voice feature can be subjected to autocorrelation calculation at the heart rate value). 음성 특징에 대한 펄스의 영향은 심장 펄스가 건강한 상태 일 때 주기적이지만, 심장 펄스가 부정맥인 경우 불규칙적 일 것으로 추정된다. 자기 상관과 같은 주기성 계산을 통해 음성 특징이 얼마나 정기적으로 영향을 받는지 알 수 있다. 추가 세부 사항 및 예시는 도 14a 내지 도 14b 및 도 15a 내지 도 15d에서 계속 된다.
- [0319] 일부 실시예에서, 분석된 패턴의 결과(results of the analyzed patterns)는 단계(362)에서 통합되어 심장 상태의 전체 확률(overall probability of a cardiac condition)을 제공하고, 단계(364)에서 심장 상태 평가(cardiac condition assessment)에 도달한다.
- [0320] 일부 실시예에서, 분석 모듈(362)은 스피치 분석으로부터 획득된 값을 여기에 나열된 세 가지 방법 각각에 지속적으로 저장한다(In some embodiments, the analysis module 362 stores the values obtained from the speech analysis on an ongoing basis in each of the three methods listed herein). 선택적으로, 시스템은 이러한 값을 학습한다; 이들은 특정 모니터링 대상자의 건강한 상태를 정의하는 "정상 값(normal values)"으로 정의된다. 일부 실시예에서, 특정 개인에 대한 지속적인 모니터링 동안, 시스템은 저장된 (정상) 값들을 현재 값들과 비교하고, 상대 변화 레벨(relative change level)을 측정한다. 일부 실시예에서, "정상 값"과 비교한 값의 변화량은 병리를 가질 확률에 비례한다.
- [0321] 선택적으로, 심장 상태 평가는, 또한 단계(370)에서 검사된 대상의 의학적 및/또는 개인 이력 데이터(medical and/or personal history data)를 고려한다. 대안 적으로 또는 추가적으로, 바이오 마커 테스트(biomarker tests)는 단계(380)에서 고려된다.
- [0322] 바이오 마커의 예에는 체중, 성별, 직업, 과거 진단 질환 및 AF 에피소드(AF episodes), 흡연 습관, CHADS2(울혈 성 심부전 병력(congestive heart failure history)), 고혈압 병력(hypertension history), 연령, 당뇨병 병력(diabetes mellitus history), 뇌졸중(stroke) 및 TIA (일시 허혈성 발작(Transient Ischemic Attack)) 증상 중 하나 또는 조합이 포함된다.
- [0323] 일부 실시예에서, 심장 상태 평가는 출력(390), 예를 들어 디스플레이를 통해 사용자에게 제시된다.
- [0324] 선택적으로, 획득된 데이터는 단계(395)에서 기계 학습에 이용된다. 예를 들어, 음성 샘플, 처리된 음성 데이터, 음성 특징, 이력 데이터(historical data), 바이오 마커 데이터(biomarker data) 및/또는 심장 상태는 향후 프로세스에서 보다 정확한 진단을 위한 규칙 및 통계적 특징(rules and statistical characterizations)을 생성하는데 사용될 수 있고, 단계(362)에서 수행되는 통합(integration)을 지원할 수 있다.
- [0325] 일부 실시예에서, 기계 학습은 훈련 단계(training stage)를 포함한다. 예를 들어, 훈련 단계에서, 음성 샘플 및/또는 음성 특징 및/또는 심장 특징은 심장 병리를 갖는 것으로 알려진 건강한 대상 및 병적 심장을 가진 것으로 알려진 대상으로부터 획득될 수 있고, 레퍼런스로 이용될 수 있다. 예를 들어, 이 데이터에 기초하여, 훈련 단계는 건강한 스피치와 병적 스피치를 분리하는 가장 적절한 알고리즘을 제공할 수 있다. 작동 단계(operation stage)에서 기계 학습 모델은 사용자 음성(및/또는 특징)으로부터 병리의 확률을 계산하고 의사 결정(decision making)을 지원한다.
- [0326] 추가적으로 또는 선택적으로, 음성 특징들은 음성 샘플의 선택된 프레임들에 대해 계산될 수 있고, 스펙트럼 포

현(spectral representation)은 각각의 특징에 대해 계산 될 수 있다. 각 모집단의 스펙트럼 분석(spectral analysis of each population)에 기초하여, 건강한 프로파일(healthy profiles)을 특징 짓기 위해 훈련 통계 모델(training statistical models)이 형성 될 수 있고, 비건강한 프로파일(non-healthy profiles)을 특징 짓기 위해 훈련 통계 모델이 형성 될 수 있다.

- [0327] 선택적으로, 분류기(classifier)는 두 가지 훈련 통계 모델을 구분하기 위해 제공된다. 대안적으로 또는 추가적으로, 확률이 비교되고, 그에 따라 결정이 이루어진다.
- [0328] 일부 실시예에서, 도 3의 흐름도에 개시된 프로세스는 부정맥이든 그렇지 않은 다양한 심장 상태를 식별하는데 사용될 수 있다. 예를 들어, 심방 플러터(atrial flutter) 및/또는 심실 빈맥(ventricle tachycardia) 및/또는 조기 심실 수축(premature ventricular(PVC)) 및/또는 비정상 심박동(abnormal cardiac rhythm) 및/또는 이소성 박동(ectopic beats)의 진단은 앞서 묘사된 것과 유사한 평가에 기초하는데, 선택적으로, 단계(362)의 통합 과정 및/또는 단계(364)의 평가 과정을 특정 비정상 심장 리듬을 결정할 수 있도록 조정한다(optionally with adjustment of the integration process at 362 and/or the assessment at 364 to the specific abnormal cardiac rhythm determined). 대안적으로 또는 추가적으로, 조정은 단계(306)에서의 특징 분포 평가 및/또는 단계(308)에서의 펄스 간격 및/또는 단계(310)에서의 주기성 평가에 이용되는 임계 값 및/또는 레퍼런스에 대해 이루어진다.
- [0329] 심실 빈맥 또는 서맥(ventricle tachycardia or bradycardia)과 같은 비정상적인 심장 리듬을 검출할 때, 일부 실시예에서, 프로세스는 위에서 묘사된 것과 유사하지만 약간의 변화가 있다. 예를 들어, 단계(308)에서 - 일부 실시예에서는, 매우 높거나 낮은 펄스 간격을 찾는다(예를 들어, 200 BPM보다 높거나(빈맥) 60 BPM보다 낮은(수맥)). 일부 실시예에서는, 단계(306)에서 - 분포를 특징 짓는 다른 파라미터가 계산된다. 일부 실시예에서는, 단계(380)에서 중요한 입력은 모니터링 된 사람 연령(60 세 초과)이다.
- [0330] 일부 실시예에서, 심방 플러터(atrial flutter) 및/또는 PVC, 및/또는 임의의 다른 유형의 부정맥을 검출하기 위해, 도 3에 도시된 과정이 뒤따르지만, 상이한 임계 값 및 파라미터 값을 갖는다.
- [0331] 일부 실시예에서, 호흡 곤란(Dyspnea) 및/또는 흉통(chest pain)을 검출하기 위해, 동일한 프로세서가 이용되지만, 다른 음성 특징, 및/또는 임계 값 및/또는 파라미터 값이 이용된다.
- [0332] 예를 들어, 위의 모든 병리에 대한 검출은 다음을 기반으로 할 수 있다:
- [0333] (1) 심장 리듬을 검출하기 위한 심장 맥박 타이밍(RR 간격) 추정하고, 병적/건강한 것으로 알려진 리듬과 비교한다.
- [0334] (2) 스피치 특징 (자기 상관(autocorrelation) 및/또는 스펙트럼 도메인(spectral domain))의 주기성을 계산하고, 이를 병적/건강한 것으로 알려진 주기성과 비교한다.
- [0335] (3) 기계 학습 전략(machine learning strategy)을 사용하여 건강한 화자 및/또는 병리가 있는 화자의 음성 특징으로부터 통계 모델을 학습한다.
- [0336] 일부 실시예에서, 상기 방법은, 예를 들어 동리듬(sinus rhythm)을 식별함으로써 건강한 심장 상태를 확인하는데 사용된다.
- [0337] 일부 실시예에서, 알려지지 않은 화자에 대해 - 병리는, 예를 들어 확률 계산에 의해, 음향 특징을 각각의 병적 및/또는 건강한 모델(예를 들어, 훈련 단계에서 계산된 모델)과 비교하고, 결과로서 가장 높은 확률 모델 취함으로써 검출될 것이다.
- [0338] **음성 샘플링 및/또는 분석을 위한 예시적인 장치(Exemplary apparatus for voice sampling and/or analysis)**
- [0339] 이제, 본 발명의 일부 실시예에 따라, 도 4a의 이동 전화 및 도 4b의 차량과 같은 예시적인 플랫폼에서의 음성 샘플링 및/또는 분석을 개략적으로 도시한 도 4a 내지 4 b를 참조한다. 선택적으로, 시스템은 모든 음성 플랫폼 및/또는 음성 지원 플랫폼에서 구현될 수 있다.
- [0340] 일부 실시예에서, 차량은 샘플링 및/또는 분석 회로를 내장시키기 위해 사용된다. 예를 들어, 음성 입력 장치 및/또는 분석기는 차량용 멀티미디어 시스템 및/또는 스피커폰 시스템에 내장될 수 있다(예를 들어, 대상이 차량의 전화/스피커에 말할 때 샘플링 및/또는 음성 명령(voice commands)을 출력 및/또는 차량 내부에서 자발적으로 대화함으로써).

- [0341] 일부 실시예에서, 전화 및/또는 스마트 폰은 음성 입력 장치(422), 예를 들어 마이크로폰을 사용하여 음성 샘플을 샘플링하기 위해 이용된다. 선택적으로, 음성 샘플의 분석은 전화 내에서, 그러한 분석에 대한 인스트럭션(instructions)을 갖는 회로에 의해 수행된다. 일부 실시예에서, 음성 프로세서(420)는 도 5a 내지 도 5c에 추가로 도시된 바와 같이 음성 샘플을 처리하기 위해 사용된다. 일부 실시예에서, 음성 프로세서(420)는, 예를 들어 프래그먼트(fragments)로 필터링(filtering) 및/또는 분류(classifying) 및/또는 분할(segmenting)으로써 음성 필터링 유닛(424)에서 음성 샘플을 세정(clean)한다. 예를 들어, 필터링 유닛(424)은 샘플링된 음성의 신호의 잡음을 제거하고 "세정(clean)"할 수 있다. 일부 실시예에서, 필터링된 또는 필터링되지 않은 음성은 추가로 분류의 대상이 된다. 선택적으로, 음성 샘플은 스피치 및/또는 논스피치(non-speech) 부분 및/또는 무음 부분(silence parts) 및/또는 전이 부분(transition parts)으로 분류된다. 일부 실시예에서, 필터링 유닛(424)은, 선택적으로 사운드 기준(sound criterion)을 이용함으로써, 음성 스트림(voice stream)으로부터 관련 스피치 데이터(relevant speech data)의 위치를 *?*검출하고 표시하기 위해 이용된다. 예를 들어, 사운드 기준은 에너지 및/또는 제로 크로싱(zero crossing) 및/또는 포먼트(formants) 및/또는 피치(pitch)일 수 있다. 예를 들어, 유성음 부분(voiced parts)에서, 자기 상관 분석은 피치 주파수(pitch frequency)에서 상당한 피크를 나타낸다. 이는 선택적으로 유성음(voiced)/무성음(unvoiced) 분류에 사용된다.
- [0342] 일부 실시예에서, 특징 추출기(426)는 음성 샘플로부터 데이터, 예를 들어 음성 특징 및/또는 음향 벡터 및/또는 심장 특징을 추출한다. 일부 실시예에서, 음성 특징의 적어도 하나의 값은 프레임마다 계산된다. 예를 들어, 음성 특징은 가중 스펙트럼(weighted spectrum) 및/또는 LPC(Linear Predictive Coefficient) 및/또는 LPC 기반 스펙트럼 및/또는 Mel Frequency Cepstral Coefficients (MFCC) 및/또는 기본 주파수(pitch) 및/또는 에너지(energy), 및/또는 제로 크로싱(zero crossing) 및/또는 포먼트(formants), 및/또는 성문 펄스(glottal pulse(vocal cord pulse)), 및/또는 지터(jitter) 및/또는 쉬머(shimmer), 및/또는 프랙탈 치수(fractal dimension), 및/또는 일관성(coherence), 및/또는 웨이블릿 분석(wavelet analysis) 또는 스피치 샘플에서 관련 정보를 추출하는 다른 특징을 포함할 수 있다. 대안적으로 또는 추가적으로, 일부 또는 모든 음성 특징의 조합은 다수의 특징을 포함하는 단일 고차원 벡터(a single high dimension vector containing multiple features)로 이용된다. 선택적으로, 이는 도 5a에서 추가로 설명되는 바와 같이, 각각의 파라미터에 대한 통계적 가중치(statistical weights for each parameter)를 포함한다.
- [0343] 일부 실시예에서, 음성 특징 및/또는 음향 벡터 및/또는 심장 특징은 특징 프로세서(440)에서 분석된다. 일부 실시예에서, 펄스 인터벌 계산기(442)는 심장 펄스 사이의 시간 간격을 계산하기 위해 사용된다. 대안적으로 또는 추가적으로, 음성 특징 분포 계산기(444)는 적어도 하나의 음성 특징의 분포 및/또는 변동성을 특성화하기 위해 사용된다. 대안적으로 또는 추가적으로, 음성 특징 주기성 계산기(446)는 적어도 하나의 음성 특징의 시간에 따른 변화의 규칙성을 식별하기 위해 사용된다.
- [0344] 일부 실시예에서, 특징 프로세서(440)의 결과는 다양한 통계적 방법을 사용하여 심장 병리의 확률을 추정하기 위해 심장 상태 분류기(460)에 의해 분석된다. 일부 실시예에서, 확률 통합 모듈(462)은 계산기(442, 444 및 446)의 결과의 임의의 조합일 수 있는 특징 프로세서(440)의 다양한 결과를 웨이트(weight)하기 위해 사용된다. 일부 실시예에서, 심장 상태 평가 모듈(464)은 심장 상태 추정을 유도하기 위해 확률 통합 모듈(462)의 웨이트 결과(weighting result)를 이용한다. 일부 실시예에서, 확률 통합 모듈(462) 및/또는 심장 상태 평가 모듈(464)은 기계 학습 모듈(466)과 데이터를 교환, 송신 및/또는 수신한다.
- [0345] 선택적으로, 도 4b에 예시된 바와 같이, 심장 상태 결정을 위한 시스템(400)은, 예를 들어 차량(402)과 같은 대상(450)을 수용하기 위한 하우징에 내장된다. 일부 실시예에서, 시스템(400)은 차량 및/또는 기타 전산 하드웨어(computational hardware)의 멀티미디어 시스템과 통합된다. 일부 실시예에서, 시스템(400)은 대상(450)으로부터 음성 샘플(420)이 입력되도록 구성된다. 일부 실시예에서, 시스템(400)은 심장 병리의 분석 중 적어도 일부를 수행하기 위한 회로를 포함한다. 대안적으로 또는 추가적으로, 음성 샘플 및/또는 적어도 부분적으로 분석된 사운드 샘플은, 예를 들어 연결된 차량의 멀티미디어 시스템을 통해, 추가 분석을 위해 서버로 전송된다. 일부 실시예에서, 결과는, 예를 들어 스마트 폰 및/또는 이메일을 통해, 대상(450)에게 직접 전송된다. 대안적으로 또는 추가적으로, 출력은, 예를 들어 멀티미디어 시스템을 통한 차량 인터페이스 및/또는 스피커, 및/또는 대시 보드 인터페이스(예를 들어 경고등의 형태)와 같은 차량 인터페이스를 통해 사용자에게 제시된다. 대안적으로 또는 추가적으로, 결과는 간병인에게 보내진다.
- [0346] **예시적인 음성 샘플 프로세싱(Exemplary voice sample processing)**
- [0347] 본 발명의 일부 실시예에 따른 예시적인 음성 처리를 도시한 도 5a를 참조한다. 일부 실시예에서, 획득된 음성

샘플을 잠재적으로 심장 및/또는 음성 특징 정보를 갖는 프레임으로 분할하고, 그러한 정보를 제공할 가능성이 적은 프레임을 제거하는 것이 잠재적으로 유리하다.

[0348] 일부 실시예에서, 예를 들어 잡음 및/또는 왜곡을 제거하기 위해 음성 샘플이 단계(502)에서 필터링된다. 선택적으로, 음성 필터링은 소음 감소 및/또는 능동 소음 제거 기술(active noise cancellation techniques)을 포함한다. 예를 들어, 배경 잡음 및/또는 음성 왜곡을 감소시킴으로써 필터링의 잠재적 이점은 향후 프로세싱 효율을 잠재적으로 증가시키는 것이다. 일부 실시예에서, 필터링은 능동 잡음 제거에 의해 수행되는데, 여기서 잡음은 스피치로부터 차감된다. 대안적으로 또는 추가적으로, 필터링은 시간 영역 또는 주파수 영역에서 잡음 필터링에 의해 수행된다.

[0349] 일부 실시예에서, 음성 샘플은 단계(504)에서 분류된다. 심장 박동에 따른 혈류는 음성 신호의 변조를 야기하는 것으로 추정된다. 일부 실시예에서, 스피치 신호로부터 심장 및/또는 음향 정보를 추출하기 위해, 스피치 신호는, 선택적으로 네 가지 클래스로 분류된다: (i) 침묵/배경 잡음 프래그먼트 - 일반적으로 관련 정보가 없는, (ii) 무성음 프래그먼트 - 일반적으로 사소한 정보를 갖는, (iii) 유성음 프래그먼트 - 관련 정보가 있는 것으로 추정되는, 및 (iv) 특정 소리 및 다른 소리 사이 성대의 구조가 변화하는 스피치의 부분으로 정의된 전이 부분(transition parts)(예를 들어, 구강 캐비티 방향 및/또는 혀 위치 및/또는 입술 위치 및/또는 성대 구성)(transition parts defined as parts of the speech where the vocal organs change their structure (e.g. the mouth cavity orientation, and/or the tongue position, and/or the lips position, and/or the vocal cords configuration) between one sound to another)). 선택적으로, 분류 후, 논스피치(non-speech) (및/또는 잡음) 세그먼트가 제거된다. 대안적으로 또는 추가적으로, 무성 세그먼트(unvoiced segments)는 제거된다. 일부 실시예에서, 특정 모음, 예를 들어, /ah/ 모음만, 또는 /eh/ 모음만, 또는 /oh/ 모음만, 또는 /uh/ 모음만, 또는 /eeh/ 모음만 포함하거나 제외하도록 유성 세그먼트(voiced segments)에 분류가 이루어진다.

[0350] 일부 실시예에서, 스피치 샘플은 단계(506)에서, 예를 들어 약 0.1msec 내지 약 400msec의 길이를 갖는 {m} 프레임으로 분할된다. 대안적으로 또는 추가로, 프레임은 약 20msec 내지 약 40msec의 길이를 갖는다. 선택적으로, 프레임은 중첩(overlapping)되는데, 예를 들어 10 % 및/또는 20 %, 및/또는 30 %, 및/또는 40 % 및/또는 50 %, 및/또는 60 % 또는 1 % ~ 99 % 범위로 중첩된다.

[0351] 일부 실시예에서, 각각의 프레임 {m}, {j}에 대해 음성 특징은 단계(508)에서 계산되어, 예를 들어:

[0352] 음성 특징 벡터 $\{j,m\}=F_j\{\text{Speech_frame}(m)\}$ 는 프레임 {m}에서 스피치 샘플로부터 계산됨

[0353] $\text{Speech_frame}(m)=\{\text{Speech}(t+Ts*m),\text{Speech}((t+1)+Ts*m)\cdots\text{Speech}((t+Tw)+Ts*m)\}$

[0354] Where:

[0355] (m) - 프레임 번호(frame number)

[0356] (t) - 샘플의 타임 인덱스(samples time index) ($t=1/F_s$ where F_s =샘플링 주파수, 일반적으로 6Khz 에서 100Khz 사이(sampling frequency usually between 6Khz to 100Khz))

[0357] (T_s) - 프레임 스텝(frame step) (중첩이 없는 경우 $T_s=T_w$, 50% 중첩된 경우 $T_s=0.5*T_w$ is 50%)

[0358] (T_w) - 프레임 크기(Frame size (duration))

[0359] (Speech) - 스피치의 샘플(samples of speech)

[0360] $F\{\}$ 는 스피치 샘플의 함수 = 스피치의 정보의 대체 표현인 음성 특징. 예를들어: 스펙트럼 및/또는 MFCC 및/또는 LPC 및/또는 자기 상관, 웨이블릿 기능 등.

[0361] 일부 실시예에서, (임의의 순서의) 시간 미분(time derivatives)는 또한 단계(510)에서 각각의 특징에 대해 계산된다:

[0362] **First derivative:** $\text{Diff1}(j,m) = \frac{d}{dm} \text{Acoustic_parameter}(j,m)$

[0363] **Second derivative:** $\text{Diff2}(j,m) = \frac{d}{dm} \text{diff1}(j,m)$

[0364] 일부 실시예에서, 특징 값들 간의 차이는 스피치 신호를 따른 다양한 포인트들 사이에서 계산된다. 비 제한적

인 예로서, 심장 박동은 잠재적으로 RR 펄스를 찾는 데 이용될 수 있는 패턴에서 특징 값을 변경한다.

[0365] 비 제한적인 예로서, 심장 박동은 피치 및 포먼트 주파수에서 스펙트럼 교차 일관성(spectral cross coherence at the pitch and formant frequencies)을 잠재적으로 변화 시켜서 일관성이 최소인 스피치의 두 부분 사이의 시간차가 RR 간격을 추정하는데 사용될 수 있다.

[0366] 일부 실시 예에서, 일부 특징은 전체 발화에 대해 평균화된다(some features are averaged on the whole utterance).

[0367] 일부 실시예에서, 일부 특징의 일부 통계 값은 전체 발화에 걸쳐 계산된다. 선택적으로, 일부 음성 특징의 경우 무성음 부분(unvoiced parts)이 사용된다. 대안적으로 또는 추가로, 일부 음성 특징들에 대해서는 (예를 들어, 포먼트 및/또는 피치와 같은 특징들을 선택함으로써) 유성음 부분(voiced parts)만이 사용된다. 일부 실시예에서, 연속 세그먼트들 사이의 작은 갭(예를 들어, BPM 길이의 약 20% 미만, BPM 길이의 약 15% 미만, 및/또는 BPM 길이의 약 10% 미만, 및/또는 약 5% 미만)은 무시된다. 예를 들어, 특징 시퀀스(feature sequence)에서 생략된다. 잠재적으로 혈압에 대한 심장 박동 효과가 갭보다 길 것으로 추정되므로 작은 갭은 중요하지 않다. 대안적으로 또는 추가적으로, 세그먼트들 사이의 갭은 재구성(reconstructed) 및/또는 외삽(extrapolated)된다. 예를 들어, 세그먼트의 끝에서 기울기(slope)와 다음 세그먼트에서 컴플리멘터리 기울기(complementary slope)를 식별할 때, 음성 특징 및/또는 심장 특징의 발생은 외삽될 수 있다.

[0368] 선택적으로, 음성 특징은, 예를 들어 도 5b 내지 도 5c에 도시된 바와 같이, 단계(512)에서 정규화(normalized)된다. 정규화(normalization)의 잠재적 이점은 트랜지션 프레임(transition frames)과 같이 정보가 더 적은 프레임의 분석을 최소화하는 것이다. 일부 실시예에서, 프레임은 미리 결정된 규칙에 따라 단계(514)에서 제거되며, 규칙은, 선택적으로, 심장 정보를 나타내는 프레임의 확률에 관한 것이다.

[0369] 일부 실시예에서, 계산된 음성 특징은 단계(516)에서 각각의 프레임을 나타내는 음향 벡터로 결합된다. 선택적으로, 음향 벡터는 단일 벡터로 결합된 여러 음성 특징(예를 들어, MFCC, LPC, 웨이블릿, 피치 등, 일반적으로 음향 파라미터라 칭함) 및/또는 음성 특징 통계(voice feature statistics) 및/또는 미분(derivatives; 아래 식에서 "DIFF"로 명명 됨)을 포함한다. 예를 들어,

$$Acoustic_{VEC}(m) = \begin{Bmatrix} Acoustic_parameter\{1,m\} \\ Acoustic_parameter\{2,m\} \\ \dots \\ Acoustic_parameter\{12,m\} \\ Diff1(1,m) \\ Diff1(2,m) \\ \dots \\ Diff1(12,m) \\ Diff2(1,m) \\ Diff2(2,m) \\ \dots \\ Diff2(12,m) \\ \ddots \end{Bmatrix}$$

[0370]

[0371] 여기서 음향 벡터는 일련의 음향 파라미터(acoustic parameters)와 그들의 미분(derivatives)을 포함한다. - 1 차 미분에 대해 Diff1, 2차 미분에 대해 Diff2로 표시함.

[0372] 비 제한적인 예로서, 프레임 "m"의 음향 벡터는 다음과 같을 수 있다:

[0373] Acoustic_parameter(1..12,m)= LPC (1..12,m)

[0374] Acoustic_parameter(13..26,m)= MFCC (1..13,m)

[0375] Acoustic_parameter(27,m)= Pitch(m)

[0376] Acoustic_parameter(28,m)= RMS_ENERGY(m)

[0377] Acoustic_parameter(29..TBD,m)= SPECTRAL_COHERENCE_PITCH(m,t)

[0378] 비고 : 피치 주파수에서 인접한 프레임에 대한 일관성 값(coherence value to adjacent frames in pitch frequencies)

- [0379] Acoustic_parameter(29..TBD,m)= SPECTRAL_COHERENCE_FORMANTS(m,t)
- [0380] 비고 : 포먼트 주파수에서 인접한 프레임에 대한 일관성 값(coherence value to adjacent frames in formant frequencies)
- [0381] Acoustic_parameter(TBD..TBD,m)= WAVELET_TRANSFORM(m)
- [0382] Diff1(1..12,m)=first derivative of LPC(1..12,m)
- [0383] Diff2(1..13,m)=second derivative of MFCC(1..13,m)
- [0384] 선택적으로, 프레임당 음성 특징 벡터는 후속 분석, 예를 들어 심박수 정보 도출에 적용된다. 대안적으로 또는 추가적으로, 후속 분석은 분포 및/또는 주기성 계산을 포함한다. 대안적으로 또는 추가적으로, 선택적으로, 필터링 및/또는 분류 후에 음성 샘플은 음성 특징 또는 음향 벡터보다는 후속 분석을 위해 사용된다(the voice sample, optionally after filtering and/or classification, is used for subsequent analysis rather than the voice features or the acoustic vectors).
- [0385] 이제, 본 발명의 일부 실시 예에 따라, 음성 샘플의 정규화를 예시하는 도 5b를 참조한다. 음향 벡터는 심장 상태 및 스피치 음성학(speech phonetics)에 대한 정보를 포함하기 때문에, 심장 정보를 강조하기 위해 스피치 음성학 정보(speech phonetic information)를 제거(예를 들어, 정규화)하는 것이 잠재적으로 유리하다. 도 5b 및 5c는 음성의 (음성학 정보(phonetic information)로 인한) "자연적(natural)" 변화를 정규화하는 방법을 예시한다.
- [0386] 일부 실시예에서, 정규화 알고리즘은, 예를 들어 아래에 기술 된 바와 같이 다항식 회귀 곡선 피팅 방법(polynomial regression curve fitting method)을 사용하여 음성의 "자연적" 변화를 추정한다.
- [0387] 일부 실시예에서, 음성 특징 벡터가 획득되면, 음향 벡터 계수(acoustic vector coefficients)는 단계(542)에서 도출된다. 일부 실시예에서, 계수는 단계(544)에서 다항식 근사(polynomial approximation)에 적용된다. 다항식 함수(polynomial function)는 특징의 시간 변동(the time fluctuation of the features), 즉 시간의 함수로서의 파라미터(parameter as the function of the time)를 추정한다. 시간. 다항식에 의한 변동 함수 표현(the fluctuation function representation by a polynomial)은 "부드러운(smooth)" 함수로, 심장 펄스 효과에서 비롯된 짧은 변화가 아닌 "자연적(natural)" 긴 변화만 포함한다(accounts).
- [0388] 일부 실시예에서, 기울기 변화는 단계(546)에서 계산된다. 선택적으로, (프레임 당) 특징 벡터는 기울기 값 및 미리 결정된 임계치 또는 범위에 대한 위치에 기초하여 단계(548)에서 대역 통과 필터링(band pass filtered)된다. 일부 실시예에서, 다항식 추정은 음성의 "자연적" 변화, 즉 박성(phonation)으로 인한 음성의 변화를 반영하고, 선택적으로 음성 특징으로부터 차감된다. 예를 들어:
- [0389] $x(i)$ 를 스피치의 세그먼트 $i = 1 \dots N$ 에 대한 음성 특징 벡터 계수로하고,
- [0390] 이는 시간 $t(i)$ $i = 1 \dots N$ 에서 샘플링된다.
- [0391] 프레임 샘플링 간격: $T_s = 1 / F_r$
- [0392] 여기서 F_r 은 프레임 속도(frame rate)이고,
- [0393] M 차 다항식이 추정에 사용되어 추정 오차 $\varepsilon(i)$ 가 발생한다:
- [0394]
$$x(i) = \sum_{k=0}^m a_k * t(i) + \varepsilon(i)$$
- [0395] 이는 아래와 같이 적을 수 있다:
- [0396]
$$\begin{bmatrix} x(1) \\ \dots \\ x(N) \end{bmatrix} = \begin{bmatrix} t(1) & \dots & t(1) \\ \dots & \dots & \dots \\ t(N) & \dots & t(N) \end{bmatrix} * \begin{bmatrix} a_0 \\ \dots \\ a_m \end{bmatrix} + \begin{bmatrix} \varepsilon(1) \\ \dots \\ \varepsilon(N) \end{bmatrix}$$

- [0397] 또는
- [0398]
$$\vec{X} = T * \vec{a} + \vec{\varepsilon}$$
- [0399] 다항식 계수는 최소 제곱 추정을 사용하여 추정 할 수 있다:
- [0400]
$$\vec{a} = (T^T * T)^{-1} * T^T * \vec{X}$$
- [0401] 다항식 추정은 스피치의 "자연적" 변화를 반영하며 음성 기능에서 차감된다:
- [0402]
$$\tilde{x}(i) = x(i) - \sum_{k=0}^m a_k * t(i)$$
- [0403] 도 5c는 검정색 선이 시간에 따른 음성 특징 값을 나타내는 예시적인 정규화 프로세스를 도시한다(상단 그래프). 파란색 선(점선)은 3차 다항식을 사용하는 정규화-수정 함수(correction function)이다. 상단 그래프는 원래 신호와 보정 기능을 보여주고 하단 그래프는 정규화 후 신호를 보여준다. 일부 실시예에서, 일정한 사운드 기간(sound periods) 사이에 있는 스피치의 부분들은 분석으로부터 제거된다.
- [0404] 스피치 중에 음성 특징은 음성 정보 및 심장 정보를 모두 반영하는 것으로 추정된다. 스피치 신호에 대한 심장 펄스 영향은 섬세한 것으로 추정되기 때문에, 일부 실시예에서 정규화 알고리즘은, 선택적으로 일정한 사운드 기간(sound period) 동안 "자연적" 음성 특징에서 자연적인 변화를 감소시키기 위해 사용된다.
- [0405] 이러한 사운드 정규화 프로세스는, 선택적으로 본 발명의 일부 실시 예에서 발생한다. 일부 실시예에서, 정규화 알고리즘은 분석을 위해 더 많은 데이터를 고려하기 위해 전이 부분(transition parts)을 감소시킴 및/또는 단일 사운드 부분을 "부드럽게(smooths)" 한다. 도 8a 내지 도 8b는 피치 파라미터를 사용하는 다른 예를 도시한다.
- [0406] **분포 분석의 예시적인 상위 레벨 개요(Exemplary high level overview of distribution analysis)**
- [0407] 이제 도 6a를 참조하면, 본 발명의 일부 실시 예에 따라 분포 분석에 의해 음성 샘플로부터 심장 상태의 추정을 나타내는 일반적인 상위 레벨 개요가 도시되어 있다.
- [0408] 일부 실시예에서, 음성은 단계(602)에서 샘플링되고 단계(604)에서 적어도 하나의 음성 특징을 추출하기 위해 분석된다. 일부 실시예에서, 음성 특징의 시간에 따른 분포는 단계(606)에서 계산되고, 단계(608)에서 심장 상태를 결정하는데 사용된다. 분포 계산은 추출된 음성 특징의 변동성을 계산하는 것을 포함한다. 심장 박동의 영향을 받는 혈류가 음성에 영향을 미치고 부정맥이 발생하면 영향이 불규칙한 것으로 추정된다. 따라서 병적 심장 상태에서 음성 특징 값의 분포가 변화될 것으로 추정된다.
- [0409] **예시적인 분포 분석의 예시적인 상세 흐름도(Exemplary detailed flowchart of exemplary distribution analysis)**
- [0410] 본 발명의 일부 실시예에 따라, 음성 특징에 대한 예로서 피치를 도시하는 특징 분포 분석을 위한 상세한 프로세스를 나타내는 흐름도를 도시하는 도 6b를 참조한다.
- [0411] 일부 실시예에서, 선택된 음성 특징은 피치이다. 피치를 계산할 때 얻을 수 있는 장점은 심장 박동으로 인한 성대 질량의 변화로 인해 심장 박동이 나타나는 주요 음성 특징 중 하나라는 것이다. 일정한 음성 신호 동안 피치 파라미터는 비교적 작은 "자연적" 변화만을 갖는 것으로 추정된다. 본 명세서에 사용된 바와 같이, 자연적 변화는, 예를 들어 음성 전환(voice transition; 예를 들어, 호흡을 포함하여 소리 및/또는 음소 사이(e.g. between and in between sounds and/or phonemes, including for example breathing)) 동안 음성 특징에서 발생하는 변화를 지칭한다. 자연적 변화는 심장 박동과 관련이 없으며, 일부 실시예에서, 도 5b 내지 도 5c에 예시된 바와 같이 정규화 알고리즘에 의해 최소화된다. 비 심장 변경(non-cardiac alterations)에 더 취약한 다른 보컬 특징(vocal feature)보다 음성 분석에 피치를 사용하는 것은 잠재적 이점이다. 피치 특징(pitch feature)의 다른 잠재적 이점은 일반적으로 음성 신호 시간의 40% 이상에서 발견되며 자발적인 음성의 다른 사운드에 대해 안정적인 값을 유지할 가능성이 있다는 것이다.
- [0412] 일부 실시예에서, 음성은 단계(612)에서 샘플링되고 유성 세그먼트는 단계(614)에서 분류된다. 일부 실시예에서, 분류된 유성 세그먼트는 단계(616)에서 연결되어 연속적인 데이터를 수신한다. 일부 실시예에서, 피치는

단계(618)에서 연결된 유성 세그먼트로부터 추출된다.

- [0413] 선택적으로 피치의 자연적 변화(예를 들어, 혈류가 아닌 발성 기관(vocal organs)의 움직임 및/또는 동작으로 인한 결과)는 제거된다. 예를 들어, 도 5b 내지 도 5c에 기술되고 도시된 바와 같은 프로세스를 사용하여 제거된다. 일부 실시예에서, 세그먼트가 제거된 후, 나머지 유성 세그먼트는 단계(622)에서 연결되고 선택적으로 단계(624)에서 평탄화된다. 일부 실시예에서, 평탄화는 상이한 세그먼트 사이의 융합 점(fusion point)에서 전이(transition)를 평탄화하는 것을 포함한다.
- [0414] 일부 실시예에서, 피치 분포는 단계(626)에서 계산된다. 예를 들어, 분포 계산은 피치 값의 통계적 변동성 및/또는 예를 들어, 표준 편차 및/또는 고차 모멘트(high order moments), 및/또는 스큐 편차(skew deviation), 및/또는 평균 등과 같은 음성 특징에 대한 수학적 연산으로부터 얻어진 임의의 다른 통계 정보를 포함한다.
- [0415] 일부 실시예에서, 분포 값은 단계(630)에서 건강 상태 또는 단계(632)에서 병적 상태를 결정하기 위해 단계(628)에서 미리 결정된 임계 값과 비교된다. 예를 들어, 높은 분포는 불규칙한 영향을 암시하므로 임계 값을 초과하는 분포 값은 병적 상태로 간주 될 수 있다. 일부 실시예에서, 분포의 값은 그 형태(shape) 일 수 있다.
- [0416] **예시적인 단일-화자 방법(An exemplary single-speaker method)**
- [0417] 일부 실시예에서, 펄스 리듬의 변화는, 선택적으로 특정 시간의 단일 환자의 음성을 다른 시간의 동일한 환자의 음성과 비교하여 변화를 측정함으로써 검출된다.
- [0418] 일부 실시예에서, 환자는 상이한 시간에, 특히 환자의 심장이 부정맥일 때 및 환자의 심장 박동이 규칙적인 때, 동일한 소리(들) 및/또는 동일한 음소(들)을 발생하도록 요청 받는다.
- [0419] 심장 신호의 변화는 음성 신호의 분석에 의해 선택적으로 감지된다.
- [0420] 일부 실시예에서, 분석은 상이한 시간에 환자의 톤, 및/또는 음소, 및/또는 단어 및/또는 말의 레코딩에 대한 것이고, 선택적으로 환자의 동일한 톤 및/또는 음소, 및/또는 말의 라이브 보이싱(live voicing)과 비교이다 (the analysis is of recordings of tones, and/or phonemes, and/or words, and/or sayings of a patient at different times and optionally comparison to live voicing of the same tones, and/or phonemes, and/or words, and/or sayings of the patient).
- [0421] 일부 실시예에서, 만성 심장 질환과 같은 만성 질환 환자의 동일한 음소는 펄스가 규칙적 일 때, 선택적으로 분석되고, 분석 결과의 음향 특징은 라이브 보이싱의 음향 특징 또는 다른 시간에 동일한 환자의 다른 기록과 비교하기 위해 선택적으로 저장된다.
- [0422] 음향 특징은 선택적으로 다음 중 하나 이상을 포함한다:
- [0423] 에너지의 스펙트럼 분포, 피크, 피크 에너지, 피크 밴드 폭, 스펙트럼 기울기 등(spectral distribution of energy, peaks, peak energy, peak band width, spectral slope etc)과 같은 스펙트럼 기반 파라미터(Spectrum based parameters).
- [0424] Mel Frequency Cepstral Coefficients (MFCC);
- [0425] LPC 계수;
- [0426] 성문 펄스 파라미터(성문 펄스 상승 시간, 성문 펄스 하강 시간, 성문 펄스 개방 기간, 성문 펄스 종료 기간, 성문 펄스 주파수, 성문 펄스 피크 위치, 성문 펄스 스펙트럼, 성문 펄스주기의 지터, 성문 펄스 진폭의 지터 등)(Glottal pulse parameters (glottal pulse rise time, glottal pulse fall time, glottal pulse open period, glottal pulse close period, glottal pulse frequency, glottal pulse peak location, glottal pulse spectrum, jitter of glottal pulse periods, jitter of glottal pulse amplitudes and more));
- [0427] 피치 파라미터(들);
- [0428] 부분 상관 계수(partial correlation coefficients(PARCOR)), 이는 매개 변수 사이의 상관 관계의 척도이지만 반드시 모든 매개 변수 사이의 상관 관계는 아님(partial correlation coefficients, abbreviated PARCOR, which are a measure of correlation between the parameters, yet not necessarily between all the parameters); 및
- [0429] 웨이블릿 분석 파라미터(wavelet analysis parameters).

- [0430] 일부 실시예에서, 음향 특징을 계산하는 것은 다음 중 하나 이상을 포함한다:
- [0431] 특징 공간의 중심(centroids of feature space), 선택적으로 K-MEANS 클러스터 분석(K-MEANS clustering analysis)을 사용하는;
- [0432] 다양한 기능의 히스토그램 및/또는 확률 밀도 함수(Probability Density Function (PDF)). 일부 실시예에서, 하나 이상의 음향 특징(들)이 발화에 대해 계산되고, 음향 특징(들)의 분포(예를 들어 최소 및 최대 값, 표준 편차, 분포 형태 등)가 선택적으로 스피치의 심박수 병리를 지시하는데 이용됨(is optionally used as an indication of heart rate pathology in speech);
- [0433] 변조 파라미터(modulation parameters) - 발화에 대한 음향 파라미터(들)를 계산 한 후, 파라미터(들)의 스펙트럼은, 선택적으로 파라메트릭(parametric) 및/또는 비 파라메트릭(non-parametric) 방법을 사용하여 계산된다. 사용되는 스펙트럼의 매개 변수는 다음 중 하나 이상이다:
- [0434] 피크 위치, 피크 대역폭, 피크 에너지, 스펙트럼 기울기 등(peaks location(s), peak band-width(s), peak energy, spectral slope(s) etc);
- [0435] SVM(Support Vector Machine)을 사용한 특징 공간의 부분 공간 계산(sub-space calculation of feature space using a Support Vector Machine (SVM)); 및
- [0436] 기계 학습 접근법의 다른 클러스터링 방법(other clustering methods from a machine learning approach).
- [0437] 상기 음향 특징의 계산은, 선택적으로 건강한 사람의 스피치 및 병적 심장 상태를 가진 사람의 스피치의 다른 특성의 분포(distributions of different character)를 생성한다.
- [0438] 일부 실시예에서, 초기 학습/훈련 단계(an initial learning/training stage)에서, 선택적으로 건강한 규칙적인 심장 리듬(healthy regular cardiac rhythm) 및 병적 심장 리듬(pathological cardiac rhythm)에 대한 RR 간격의 길이의 분포를 선택적으로 포함하는 음향 신호의 통계적 모델이 계산되고, 음향 특징의 하나 이상의 통계적 측정치의 건강 값과 병적 값 사이의 구별을 위해 임계 값이 결정된다.
- [0439] 일부 실시예에서, 통계적 심장 상태 분류기는, 선택적으로 다음에 의해 훈련된다:
- [0440] 심장 건강 및 심장 건강에 해로운 환자의 음성 녹음(voice recordings)을 수집하는 단계;
- [0441] 여기에 언급된 바와 같이 녹음의 하나 이상의 음향 특징을 계산하는 단계;
- [0442] 전술한 기술 및/또는 추적된 기계 학습 기술을 이용하여 통계적 분류기를 훈련시키는 단계.
- [0443] 일부 실시예에서, 음성 녹음은 다른 사람들, 선택적으로 다수의 건강한 및/또는 다수의 건강하지 않은 화자에게 수집된다.
- [0444] 일부 실시예에서, 음성 녹음은 한 명의 사람에게 수집된다. 예를 들어, 사람이 건강하다고 알려진 시간에 수집된다.
- [0445] 일부 실시예에서, 건강한 스피치의 변화가 검출되는 경우, 사람은, 선택적으로 잠재적으로 건강하지 않은 심장 상태를 갖는 것으로 분류된다.
- [0446] 일부 실시예에서, 음성 녹음은, 선택적으로 또한 한 명의 사람에게 수집된다. 예를 들어, 사람이 건강하지 않은 것으로 알려진 시간에, 선택적으로 병적 심장 상태로 알려진 시간에 수집된다.
- [0447] 일부 실시예에서, 스피치가 건강하지 않은 녹음(들)과 유사한 것으로 검출될 때, 사람은, 선택적으로 건강하지 않은 심장 상태를 갖는 것으로 분류된다.
- [0448] 일부 실시예에서, 초기 학습/훈련 단계에서, 선택적으로 건강한 규칙적인 심장 리듬(healthy regular cardiac rhythm) 및 병적 심장 리듬(pathological cardiac rhythm)에 대한 RR 간격의 길이의 분포를 선택적으로 포함하는 음향 신호의 통계적 모델이 계산되고, 음향 특징의 하나 이상의 통계적 측정치의 건강 값과 병적 값 사이의 구별을 위해 임계 값이 결정된다.
- [0449] 본 발명의 일부 실시예에 따라 심장 상태를 진단하는 것을 학습하는 방법의 단순화된 흐름도인 도 6c를 참조한다.
- [0450] 도 6C는 다음을 포함하는 방법을 도시한다:

- [0451] 음성 샘플링(Sampling voice; 640);
- [0452] 음성 특징 추출(Extracting voice features; 642);
- [0453] 선택적으로, 여러 발화의 음성 특징을 결합(Optionally combining voice features from several utterances; 644).
- [0454] 클러스터 분석 수행(Performing cluster analysis; 646);
- [0455] 선택적으로 코드북 생성(Optionally producing a codebook; 648).
- [0456] 일부 실시예에서, 샘플링은 6Hz 내지 100KHz를 포함하는 주파수 범위에서 음성을 디지털화함으로써 선택적으로 수행된다.
- [0457] 일부 실시예에서, 음성의 샘플링은, 선택적으로 심장이 건강한 화자에 의한 스피치 및/또는 적어도 화자(들)이 비병적 심장 리듬(a non-pathological heart rhythm)을 가진 때의 스피치에 대해 수행된다.
- [0458] 일부 실시예에서, 음성의 샘플링은, 선택적으로 병적 심장 리듬을 갖는 화자(들)에 의해 발언된 것으로 알려진 스피치를 포함하고, 선택적으로 병적 심장 리듬을 갖는 화자의 스피치의 예로서 인식된다.
- [0459] 일부 실시예에서, 여러 발화로부터의 음성 특징의 선택적인 결합(the optional combining of voice features from several utterances)은 건강한 화자(들)의 발화 및/또는 발화자(들)가 비병적 심장 리듬을 갖는 때의 발화에 대해 선택적으로 수행된다.
- [0460] 일부 실시예에서, 여러 발화로부터의 음성 특징들의 선택적인 결합은 건강한 화자(들)의 발화에 및/또는 화자(들)가 비 병적 심장 리듬을 가질 때 및 병적 심장 리듬을 가진 화자의 발언에 대해 분리되서 선택적으로 수행될 수 있다(the optional combining of voice features from several utterances is optionally performed separately for utterances of healthy speaker(s) and/or when the speaker(s) have a non-pathological heart rhythm and separately for speech of a speaker or speakers having a pathological heart rhythm).
- [0461] 일부 실시예에서, 클러스터 분석(cluster analysis)은 건강한 화자(들)의 발화 및/또는 화자(들)가 비 병적 심장 리듬을 가질 때 선택적으로 수행된다.
- [0462] 일부 실시예에서, 클러스터 분석은 건강한 화자(들)의 발화 및/또는 비 병적 심장 리듬을 가질 때의 발화에 대해 및 병적 심장 리듬을 갖는 화자(들)자의 스피치에 대해 분리되어 선택적으로 수행될 수 있다.
- [0463] 일부 실시예에서, 코드북은, 선택적으로 건강한 화자(들) 및/또는 화자(들)가 비 병적 심장 박동을 가질 때를 나타낸다.
- [0464] 일부 실시예에서, 코드북은, 선택적으로 건강한 화자(들)의 스피치 및 병적 심장 리듬을 갖는 화자(들)의 스피치에 대한 분류를 포함한다.
- [0465] 일부 실시예에서, 클러스터 분석은 K-평균 분석(K-means analysis)에 의해 선택적으로 수행된다.
- [0466] 본 발명의 일부 실시예에 따라 심장 상태를 분류하는 방법의 단순화된 흐름도인 도 6d를 참조한다.
- [0467] 도 6d는 다음을 포함하는 방법을 도시한다:
- [0468] 음성을 샘플링하는 단계(650);
- [0469] 음성 특징 추출하는 단계(652);
- [0470] 선택적으로, 음성 특징을 선택하는 단계(654);
- [0471] 하나 이상의 코드북 값(들)으로부터 음성 특징의 거리를 계산하는 단계(656);
- [0472] 거리를 임계 값과 비교하는 단계(658); 및
- [0473] 심장 상태 분류하는 단계(659).
- [0474] 일부 실시예에서, 샘플링은 6Hz 내지 100KHz를 포함하는 주파수 범위에서 음성을 디지털화함으로써 선택적으로 수행된다.
- [0475] 일부 실시예에서, 심장 상태를 분류하는 단계는 건강한 심장 상태의 하나 이상의 코드북 값으로부터의 임계 값보다 큰 거리에 기초하여 결정된다(the classifying cardiac condition is optionally determined based on

the distance being smaller than a threshold value from one or more codebook values of a healthy cardiac condition).

- [0476] 일부 실시예에서, 심장 상태를 분류하는 단계는 건강한 심장 상태의 하나 이상의 코드북 값으로부터의 임계 값보다 작은 거리에 기초하여 결정된다.
- [0477] 일부 실시예에서, 심장 상태를 분류하는 단계는 질환이 있는 심장 상태의 하나 이상의 코드북 값으로부터의 임계 값보다 큰 거리에 기초하여 결정된다.
- [0478] 일부 실시예에서, 심장 상태를 분류하는 단계는 질환이 있는 심장 상태의 하나 이상의 코드북 값으로부터의 임계 값보다 작은 거리에 기초하여 결정된다.
- [0479] **연결된 유성 세그먼트의 예(Example of concatenated voiced segments)**
- [0480] 도 7a 내지 도 7b를 참조한다. 본 발명의 일부 실시예에 따라, 도 7a는 음성 샘플에 대한 예를 도시하며, 이는 음성 세그먼트를 제거하도록 분류된 후 나머지 세그먼트가 도 7b에 연결되어 있다.
- [0481] 예시된 분류에서, 도 7a의 음성 샘플은 유성 세그먼트(voiced segments; 702) 및 무성 세그먼트(unvoiced segments; 704)에 대해 분류되었다. 도 7b에 도시된 그래프는 나머지 유성 세그먼트(702)가 연결된 것을 나타낸다. 세그먼트의 연결은 본 발명의 일부 실시예에서만 수행되며, 음성 특징의 시간 순서가 덜 중요한 경우에, 반면 각 지점에서의 행동 및/또는 전체로서의 포인트가 분석될 때 일반적으로 제공됨에 유의해야 한다.
- [0482] **데이터 평탄화의 예(Example of data smoothing)**
- [0483] 이제 평활화 된 음성 샘플로부터 추출 된 피치 데이터에 대한 예를 도시하는 도 8a 내지도 8b를 참조하면, 도 8a 및 도 8b는 음성의 "자연적" 변화의 평탄화를 도시한다. 빨간색 선은 시간이 지남에 따른 음성 특성을 나타내며, 이 경우 시간에 따른 피치를 나타내며, 파란색 선은 수정 함수(correction function)를 나타낸다.
- [0484] 도 8b는 정규화 후의 피치 신호(pitch signal)를 도시한다. 일반적으로, 음성 특징으로부터 "자연적" 변화를 제거한 후에, 심장 정보를 추출에 높은 정확도가 획득된다.
- [0485] **피치 추출의 예(Example of pitch extraction)**
- [0486] 본 발명의 일부 실시예에 따라, 필터링되고 평탄화된 후, 음성 샘플로부터 추출된 피치 데이터에 대한 예를 도시하는 도 9a 내지도 9c를 참조하며, 도 9a는 연결된 유성 세그먼트에 대한 예를 도시한다. 도 9b는 본 발명의 일부 실시예에 따라, 도 9a의 경우에 있어 연결된 유성 세그먼트의 피치 추출에 대한 예를 도시한다. 점선 부분은, "자연적" 변화가 중요하고 심장이 기여한 변화를 가릴 수 있기 때문에, 분석에서 생략된 개별 세그먼트 사이의 "융합(fusion)" 지점이다.
- [0487] **음성 기능 통계 분석 방법을 사용한 심박수 병리의 확률에 대한 예시 적 결정 (Exemplary determination of probability for heart rate pathology using voice features statistical analysis method)**
- [0488] 일부 실시예에서, "건강한" 음성 특징 (및/또는 그 통계 및/또는 수학적 오퍼레이션(mathematical operation)) 분포는 건강한 화자(즉, 화자 독립 모델(a speaker independent model))의 트레이닝 세트로부터 계산된다. 대안적으로 또는 추가로, "건강한" 음성 특징 분포 (및/또는 그 통계 및/또는 수학적 오퍼레이션)는 건강한(예를 들어, 충분히 주기적인) 상태로 알려진 동안 테스트된 화자로부터 계산된다(화자 의존 모델(a speaker dependent model)).
- [0489] AF 동안(도 10b 및 도 10d) 및 건강한 사람들의 스피치 동안(도 10a 및 도 10c) 기록된 스피치의 피치의 정규화된 표준 편차 값의 분포의 히스토그램의 예를 도시하는 도 10a 내지 도 10d를 참조한다. 본 발명의 일부 실시예에 따라 음성 특징 분포 분석을 이용하여 심박수 병리에 대한 확률의 결정을 예시한다(exemplifying the determination of probability for heart rate pathology using voice features distribution analysis, in accordance with some embodiments of the invention).
- [0490] 도 10a (및 도 10c - 도 10a와 동일한 분포를 더 높은 해상도로 나타냄)은 건강한 대상으로부터 도출된 스피치의 트레이닝 세트의 피치에 대해 계산된 정규화된 표준 편차이고, 도 10b (및 도 10d - 도 10b 와 동일한 분포를 더 높은 해상도로 나타냄)은 AF 진단된 대상에 대한 예시이다. 이 예에서, 피치 파라미터는 다른 사운드에 비해 상대적으로 높은 안정성으로 인해 선택되었으며 자연 스피치 기간의 40% 이상이 존재한다.
- [0491] 일부 실시예에서, 피치의 정규화된 표준 편차의 임계 값이 AF를 검출하기 위해 사용된다. 대안적으로 또는 추

가로, 건강한 대상에 대해 트레이닝된 통계 분포가 형성되고 건강하지 않은 대상에 대해 트레이닝된 통계 분포가 형성되고, 선택적으로 현재 분포는 각각의 트레이닝된 분포에 통계적으로 매칭된다. 선택적으로, 진단은 매치하기 위한 더 높은 통계적 확률에 기초하여 추정된다.

[0492] 예를 들어, i 번째 스피치의 프레임, $i = 1 \cdots N$ 에 대해 계산된 피치 값의 경우:

[0493] 표준 편차는 음성 신호의 여러 부분에 걸쳐 계산된다. 부분 지속 시간(the parts duration)은 사용되는 음향 파라미터의 가변성에 따라 (짧은 병적 에피소드를 검출하기 위해) 최소로 선택된다.

[0494]

$$STDEV = \sqrt{\frac{1}{N} \sum_{i=1}^N (Pitch(i) - \mu)^2}$$

[0495]

$$\mu = \frac{1}{N} \sum_{i=1}^N Pitch(i).$$

[0496] 이 예에서는, 일반 펄스에 대해 85%가 $STD < 0.02$ 이고, AF-레코딩(AF-recording)에 대해 60%가 $STD > 0.02$ 인 것을 볼 수 있다(For this example it can be seen that 85% of the $STD < 0.02$ for regular pulse, and for 60% of the AF-recordings $STD > 0.02$).

[0497] **예시적인 심박수 추출(Exemplary heart rate extraction)**

[0498] 일부 계산에서, 일부 실시예에서, 분석된 대상의 심박수가 사용된다. 다음은 본 발명의 일부 실시예에 따른, 대상의 음성 샘플로부터 예시적인 심장 박동 주파수 추출이다.

[0499] 일부 실시예에서, 심박수는 스펙트럼 분석에 기초하여 계산된다(예를 들어, 도 14b의 블록(1416)에 도시된 바와 같이). 예를 들어, 음성 특징에 대한 이산 푸리에 변환에 기초하여 계산된다. 다음 예에서, $\vec{V}$ 는 T개의 연속적인 스피치 프래그먼트(예를 들어, 잡음 및/또는 침묵이 없는 스피치 프레임의 프래그먼트)의 p개의 음성 특징의 매트릭스를 나타낸다:

[0500]

$$\vec{V} = \begin{bmatrix} v(t, 1) & v(t+1, 1) & \dots & v(T, 1) \\ \vdots & \vdots & \ddots & \vdots \\ v(t, p) & v(t+1, p) & \dots & v(T, p) \end{bmatrix}$$

[0501] 매트릭스는 특정 시간의 다양한 음성 특징을 포함할 수 있다. 예를 들어:

[0502] $V(t, 1)$ - 시간 t에서 피치 값일 수 있음.

[0503] $V(t, 2)$ - 시간 t에서 피치 미분(pitch derivative)일 수 있음.

[0504] $V(t, 3)$ - 시간 t에서 MFCC 계수일 수 있음.

[0505] 이산 푸리에 변환(a discrete Fourier transform)은 다음을 사용하여 각 계수 (매트릭스의 행(row of the matrix))에 대해 계산될 수 있다.

[0506]

$$\vec{V}(w) = \begin{bmatrix} V(w, 1) \\ \vdots \\ V(w, p) \end{bmatrix} = \begin{bmatrix} \left\| \sum_{n=0}^{N-1} v(t, 1) * e^{-i\omega t/N} \right\| & \dots \\ \vdots & \ddots \\ \left\| \sum_{n=0}^{N-1} v(t, p) * e^{-i\omega t/N} \right\| & \dots \end{bmatrix}.$$

[0507] 각 행은 하나의 음성 특징 파라미터의 푸리에 변환의 절대 값(absolute value)과 같다. 심장 박동수는 심장 박동수의 주파수에서 피크로 주파수 도메인(frequency domain)에 표시된다. 선택적으로, BPM(분당 비트 수(Beats Per Minute) - 심박수의)은 여러 음성 특징에 대해 계산되는데, 평균값 또는 평균 계산된 값을 사용하여 계산되거나, 계산된 모든 특징 계수에서 가장 빈번한 값의 과반수로 계산된다. 일부 실시예에서, 비합리적인

값(non-reasonable values)은 필터링된다(예를 들어, 약 30bpm 미만 및 약 300bpm 초과). 대안적으로 또는 추가로, 매우 불규칙적인 심장 박동을 위해 스펙트럼 에너지는 대역 제한되지 않는다(not band limited).

[0508] 예시적인 펄스 간격 추출(Exemplary pulse interval extraction)

[0509] 본 발명의 일부 실시예에 따라, 펄스 간격(pulse interval)을 계산하기 위한 일반적인 흐름도를 도시한 도 11a를 참조한다.

[0510] 일부 실시예에서, 음성 샘플(1102)은 단계(1104)에서 적어도 두 개의 심장 펄스를 식별하기 위해 사용된다. 일부 실시예에서, 펄스 발생 사이의 시간 간격(time interval)은 단계(1106)에서 계산된다. 일부 실시예에서, 심장 상태는 추출된 사건 간격에 기초하여 단계(1108)에서 추정된다.

[0511] 선택적으로, 펄스는 서로 연속적이다. 대안적으로, 시간 간격이 충분히 짧게 계산되면, 비 연속 펄스는 심장 상태를 결정하기에 충분하다.

[0512] 일부 실시예에서, 분석은 각 프레임에서 개별적으로 실시된다. 대안적으로 또는 추가로, 분석은 복수의 프레임의 구성된 시퀀스(composed sequence of a plurality of frames)에서 실행된다. 일부 실시예에서, 식별된 펄스의 시퀀스가 컴파일되며(compiled), 선택적으로 펄스가 다른 프레임에서 발생하더라도 펄스의 시간 시퀀스(time sequence)를 유지한다. 일부 실시예에서, 식별된 펄스들 사이의 시간 간격은 심장 상태를 결정하기 위해 기준 및/또는 임계 값과 비교된다.

[0513] 예시적인 상세한 펄스 간격 추출 (Exemplary detailed pulse interval extraction)

[0514] 본 발명의 일부 실시예에 따라, 상세한 펄스 간격 추출을 예시하는 흐름도인 도 11b를 참조한다.

[0515] 일부 실시예에서, 음성 샘플(1112)은 단계(1114)에서 유성 세그먼트(voiced segments)를 추출하도록 분류된다. 일부 실시예에서, 유성 세그먼트는 약 0.5초 내지 약 3초의 길이를 포함한다. 일부 실시예에서, 유성 세그먼트는 단계(1116)에서 심장 펄스를 식별하기 위해 이용된다. 예를 들어, 도 11c 및 11d에 더 상세하게 도시된 바와 같이 R 펄스(R pulses)를 식별하는데 사용된다. 일부 실시예에서, 식별된 펄스 사이의 시간 간격은 단계(1118)에서 계산된다.

[0516] 일부 실시예에서, 적어도 두 개의 펄스가 단일 음성 세그먼트(single voice segment)에서 식별된다. 대안적으로, 음성 세그먼트 당 적어도 세 개의 펄스, 또는 적어도 넷, 또는 적어도 다섯, 또는 적어도 여섯, 또는 적어도 일곱, 또는 적어도 여덟 개의 펄스가 식별된다. 일부 실시예에서, 이 세그먼트는 비 연속적인 부분을 포함한다. 일부 실시예에서, 서로 연속적인 펄스만이 고려된다. 대안 적으로, 대상의 심박수와 같거나, 작은 시간 간격만이 고려된다.

[0517] 선택적으로, 식별된 펄스들 사이의 짧은 시간 간격 시퀀스(short sequences of time interval)는 단계(1120)에서 기준(reference)과 매칭되고, 기준과 매칭할 확률이 계산된다. 일부 실시예에서, 기준은 동일한 대상에서, 그러나 건강한 상태 동안 식별된 펄스의 시간 간격이다. 대안적으로 또는 추가로, 기준은 건강 및/또는 병적 상태의 특성(characteristic)인 것으로 알려진 펄스의 시간 간격을 포함한다. 대안적으로 또는 추가적으로, 기준은 다른 건강 및/또는 병적 음성 샘플로부터 도출된 복수의 시간 간격 시퀀스의 데이터베이스를 포함한다. 일부 실시예에서, 기준 패턴(reference pattern)은 건강 및 병적 음성 샘플의 데이터베이스를 사용하여 예비 단계(preliminary stage)에서 획득된다.

[0518] 일부 실시예에서, 값을 매치시킬 확률(probability to match value)은 단계(1122)에서의 임계 값과 비교되고, 예를 들어 병적 시퀀스와 매칭할 때, 임계 값 미만의 값은 단계(1130)에서 건강한 상태 결정으로 이어지고, 임계 값 초과 인 값은 단계(1132)에서의 병적 상태 결정으로 이어진다. 일부 실시예에서, 임계 값은 미리 결정된 고정된 값이다. 대안적으로, 임계 값은, 예를 들어 기계 학습을 사용하는 것과 같은 다수의 데이터 소스(source of data)에 기초하여 동적으로 결정되는 변수이다. 예를 들어, 다른 테스트에서 부정맥을 결정할 높은 확률이 검출된 경우, 단계(1122)에서의 시간 간격에 기초하여 부정맥을 결정하기위한 임계 값이 더 낮을 수 있다.

[0519] 예시적인 R 펄스 위치 식별(Exemplary R pulse location identification)

[0520] 이제 시간 축 상의 R 펄스 위치의 식별을 예시하는 흐름도를 도시하는 도 11c를 참조한다.

[0521] 따라서, 일부 실시예에서, 음성 프레임의 분류 (도 5a 내지 도 5c에 도시 됨) 후 및 선택적으로 무음 및/또는 정보를 가지지 않은 프레임 및/또는 전이 프레임(transition frames)을 생략한다.

[0522] 일부 실시예에서, 스피치의 연속 세그먼트에 대해, 스피치의 프레임을 나타내는 음향 벡터가 획득된다. 일부 실시예에서, 각 프레임 음향 벡터로부터 세그먼트의 다른 모든 프레임까지의 거리가 계산된다. 잠재적으로, 거리는 조직 유연성(tissue flexibility)을 반영하는 음성 장기 조직(speech organ tissues)의 비 연속적이거나 알려지지 않은 작은 변화로 인한 음성 신호의 변화, 및/또는 근육 긴장 및/또는 혈관의 질량 및 부피의 변화로 인한 자연적 변화에 관한 정보를 보유한다. 연속된 프레임 사이의 거리는 작을 것으로 예상되고, 더 분리된 프레임일수록 커질 것으로 예상된다.

[0523] 일부 실시예에서, 단계(1150)에서 각각의 프레임 (t)에 대한 거리는, 선택적으로 모든 스피치 프레임들에 대해 이전(후방) 프레임들 및 다음(전방) 프레임에 대해 계산된다.

[0524] 예를 들어, 두 음향 벡터 사이의 거리 함수(distance function)는 다음과 같이 계산할 수 있다:

$$Acoustic_{vec(m)} = \begin{Bmatrix} Acoustic_parameter\{1,m\} \\ Acoustic_parameter\{2,m\} \\ \dots \\ Acoustic_parameter\{12,m\} \\ Diff1(1,m) \\ Diff1(2,m) \\ \dots \\ Diff1(12,m) \\ Diff2(1,m) \\ Diff2(2,m) \\ \dots \\ Diff2(12,m) \end{Bmatrix} = \begin{Bmatrix} ap(1,m) \\ ap(2,m) \\ ap(3,m) \\ \vdots \\ \ddots \\ ap(36,m) \end{Bmatrix} =$$

[0525]

[0526] {여기서 음향 벡터에는 일련의 음향 파라미터와 그 미분(1차 미분은 Diff1, 2차 미분은 Diff2로 표시됨)이 포함됨}

[0527] $D(t1,t2)=D(Acoustic_vec(t1),Acoustic_vec(t2))=$

$$=\frac{1}{12}\sum_{p=1}^{12}[(ap(p,t1)-ap(p,t2))*W(p)]^2$$

[0529] 여기서 W(p)는 정규화 및 스케일링에 사용되는 웨이트 계수(weighing coefficients)이다.

[0530] 일부 실시예에서, 단계(1152)에서 전방 거리 시리즈(the forward distance series) 및/또는 후방 거리 시리즈(the backward distance series)의 국소 최소 값(local minima)이 식별된다.

[0531] 전방($t > t1$)에서의 포인트($t=t1$)로부터 측정된 거리의 국소 최소값을 찾는다:

$$t2 = ARGMIN_j\{D(t,j)\}, j > t1$$

[0532]

[0533] 그리고 후방에서는($t < t1$),

$$t3 = ARGMIN_j\{D(t,k)\}, k < t1$$

[0535] 일부 실시예에서, 국소 최소 값까지의 거리의 합은 프레임 t에 기초하여 단계(1154)에서 RR 간격을 추정하는데 사용된다:

$$\widetilde{RR}_{interval}^t = a(t3 + t2)$$

[0536]

[0537] 여기서 a는 상수이다.

[0538] 본 발명의 일부 실시예에 따른 RR 간격 추정의 예시적인 그래프를 도시하는 도 11d를 참조한다.

[0539] 도 11d는 X축(1107)을 따른 시간 (t)의 함수로서 Y축(1106)에서 D(t)를 나타내는 그래프(1100)를 포함한다.

[0540] 그래프(1100)는 시간 t1에서 D(t1)(1101), t1 이전 시간 t2에서 D(t)의 제 1 최소값(1102), 및 t1 이후 시간 t3에서 D(t)의 제 2 최소값(1103)을 도시한다.

[0541] 도 11d는 $RR_{interval} = (t3-t2)$ 로서 최소값(1102, 1103)을 검출하는 것에 기초하여 추정되는 RR 간격을 도시

한다.

- [0542] 일부 실시예에서, 여러 RR 간격이 추정된다.
- [0543] 일부 실시예에서, 다음과 같은 하나 이상의 기준에 따라, 일부 추정된 RR 간격만이 사용된다:
- [0544] (1) 국소 최소값의 모양 ("깊음(deep)" 또는 "평평한(flat)")
- [0545] (2) RR 간격 유효성. 일부 실시예에서, RR 간격은 합리적인 범위내에서(within reason) BPM 값을 추정하기에 적합한 경우에만 선택적으로, 사용된다. (비제한적인 예로서, 45 BPM보다 크고 250 BPM보다 작은)
- [0546] (3) 가장 가능성이 높은 RR을 선택하라. 일부 실시예에서, RR 간격 길이는 선택적으로 계산되고, RR 간격 길이의 분포는 선택적으로 계산된다. 선택적으로, RR 간격은 평균 RR 간격 길이(an average RR interval length)를 둘러싼 특정 임계 값에 속하도록 선택된다. 예를 들어, 평균의 대략 $\pm 5\%$, 10% , 20% , 33% , 50% 와 같이 임계 값에 속하도록 선택된다. 선택적으로, 평균 심박수를 둘러싼 심박수에 해당하는 RR 간격 길이의 특정 범위에 속하는 RR 간격이 선택된다. 예를 들어, 대략 $\pm 5\%$, 10% , 20% , 33% , 50% 범위에서 선택된다. 일부 실시예에서, 이러한 분석은, 선택적으로 스피치의 세그먼트의 분석 이후에 이루어진다. 선택적으로, 전체 대화를 분석하고 가능한 RR 간격의 목록을 획득 =한 후에 이루어진다.
- [0547] 일부 실시예에서, 가장 가능한 패턴(most probable pattern)을 찾아서 시퀀스 분석이 수행된다. 예를 들어, 규칙적인 심박수 - 유사한 RR 간격이 발견될 수 있다. 불규칙한 심박수의 경우 유사한 RR 간격은 거의 없거나 발견될 수 없다.
- [0548] (4) 다른 기준을 사용하여 "심장 박동"과 관련 없는 자연적 스피치 변화의 확률을 추정하라.
- [0549] 일부 실시예에서, 국소 최소값이 예리하지 않은 경우(예를 들어, 긴 기간(long period)) 추정된 거리는 후속 분석으로부터 제거되고 다음 최소값이 고려된다. 일부 실시예에서, 일반적으로 짧은 프레임 세그먼트에 대해, 프레임 세그먼트 당 하나의 RR 간격이 계산된다. 대안적으로, 복수의 RR 간격은 일반적으로 더 긴 세그먼트에 대해 계산된다.
- [0550] 일부 실시예에서, RR 간격은 모든 프레임 $\{t = 0, 1, \dots, T\}$ 에 대해 추정된다. 선택적으로, 전체 스피치 샘플로부터 도출된 RR 간격은 단계(1156)에서 계산된 모든 특징에 대해 평균화되고, 선택적으로 전체 스피치 샘플의 RR 간격의 전체 추정치는 단계(1158)에서 도출된다.
- [0551] **음성 샘플에서 RR 간격 식별의 예 (Example of RR interval identification in a voice sample)**
- [0552] 일정한 모음(vowel)에서 MFCC 타입 음성 특징들에 대한 프레임 대 프레임 거리의 예를 나타내는 도 12a가 이제 참조되며, 본 발명의 일부 실시예에 따라, 상단 그래프는 (각 프레임에서 계산된 MFCC 음향 특징의) 프레임 대 프레임 거리 패턴을 나타내고, 하단 그래프는 기준으로써, ECG 그래프를 나타낸다. 도 12a에 도시된 화살표는 두 개의 프레임, 즉 프레임 (t) 및 프레임 (i)을 나타낸다. 주어진 프레임 {t}의 경우, 프레임 대 프레임 {i}의 거리는 로컬 최소 점이 RR 간격마다 나타날 수 있는 주기적인 특성을 갖는다. 즉, $D(t, i)$ 는 사이의 거리가 RR 간격과 동일한 국소 최소값을 가진다. 상단 그래프에서 하단 그래프까지의 화살표는 프레임간 거리의 국소 최소값(상단 그래프)이 주기적으로 RR 간격(하단 그래프)과 유사한 위치에 나타남을 보여준다.
- [0553] **일관성을 사용하는 음성 샘플에서 RR 간격 식별의 예(Example of RR interval identification in a voice sample using coherence)**
- [0554] 일부 실시예에서, 스피치의 변동의 척도로서 일관성 값(coherence value)을 사용하여 상기 방법이, 선택적으로 사용된다.
- [0555] 상기 방법 및 정의를 이용하여:
- [0556]
$$D(t1, t2) = D(\text{Acoustic_vec}(t1), \text{Acoustic_vec}(t2)) =$$
- [0557]
$$= \frac{1}{12} \sum_{p=1}^{12} [(\text{ap}(p, t1) - \text{ap}(p, t2)) * W(p)]^2$$
- [0558] 일관성에 대해:
- [0559]
$$D(t1, t2) = \text{Coherence}(\text{Speech}(t1:t1+tw), \text{Speech}(t2:t2+tw))$$
- [0560] 일부 실시예에서, 스펙트럼 교차 일관성(spectral cross coherence)은 하나의 시간 t1에서 시작하고, 다른 하나

는 시간 t2에서 시작하는, 두 개의 스피치 프레임들 사이에서 계산된다.

[0561] $X(n)=\text{speech}(t_1, t_1+1, t_1+2, \dots)$

[0562] $Y(n)=\text{speech}(t_2, t_2+1, t_2+2, \dots)$

[0563] 여기서, T_w = 분석하는 프레임 사이즈(frame size of analysis)

$$C_{xy}(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f) * G_{yy}(f)}$$

$$G_{xy}(f) = \frac{1}{2\pi} \sum R_{xy}(n) * e^{-i2\pi fn}$$

$$D(t_1, t_2) = \sum_{f=\text{pitch or } f=\text{formants}} C_{xy}(f)$$

[0567] 이제 본 발명의 일부 실시예에 따른 일관성 계산의 예시적인 그래프를 도시하는 도 12b를 참조한다.

[0568] 도 12b는 X축(1237)을 따라 밀리 초(milliseconds) 단위의 시간 (t)의 함수로서 Y축(1236)의 일관성 값(coherence value)을 나타내는 그래프(1230)를 포함한다.

[0569] 그래프(1230)는 시간 t1의 포인트(1231)에서 최소 값에 도달하고, ~820msec 거리의 시간 t2(1232)에서 주파수 최대 시프트(maximal shift in frequency)에 해당하는 피치 주파수에 대한 일관성을 보여준다. 시간 t2는 환자의 펄스 주기 855msec에 가깝다.

[0570] RR 간격 데이터의 예(Example of RR interval data)

[0571] 본 발명의 일부 실시예에 따라, 연속 펄스의 발생이 증가함에 따라 감도 변화 대 심장 상태 결정의 특이성(the change in sensitivity versus the specificity of heart condition determination)에 대한 예를 나타내는 도 13을 이제 참조한다.

[0572] 그래프는 Y축에 병적으로 정확하게 진단한 사람의 확률을 P(D)로, X축에 병적으로 잘못 진단한 사람의 확률을 P(FA)(1에서 확률을 뺀 값)을 보여준다. 도시된 그래프는 연속 펄스 수가 증가할 때 감도(sensitivity)가 어떻게 증가하는지를 예시한다. 또한, 연속 펄스 수가 증가하면 특이성(specificity)과 감도(sensitivity)도 증가한다는 것을 알 수 있다.

[0573] 계산은, 소수의 RR 펄스로부터의 AF 평가 방법을 포함하는, 전체 부정맥 확률 통합 섹션(the section of overall Arrhythmia probability integration)에서 나타난, 아래에 표시된 알고리즘에 기초한다.

[0574] 예시적인 주기성 결정 개요(Exemplary periodicity determination overview)

[0575] 본 발명의 일부 실시예에 따라, 심장 상태를 결정하기 위해 사용된 심박수의 주기성을 결정하는 프로세스를 나타내는 흐름도인 도 14a를 참조한다. 심박수가 주기적이라고 가정하면, 혈관 질량 및/또는 부피의 변화는 반주기적(semi-periodic_ (또는 정상적인 심박수의 주기(periodic for normal sinus heart rate))일 것으로 예상된다. 따라서, 심장 활동의 타이밍은 프레임 대 프레임 거리 패턴에서 주기적 변화를 발견함으로써 잠재적으로 검출된다.

[0576] 일부 실시예에서, 음성은 단계(1402)에서 샘플링되고 적어도 하나의 음성 특징은 단계(1404)에서 추출된다. 일부 실시예에서, 적어도 하나의 음성 특징의 주기성은, 예를 들어 스펙트럼 분석 및/또는 자기 상관을 이용하여 단계(1406)에서 계산된다. 일부 실시예에서, 심장 상태는 음성 특징의 주기성에 기초하여 단계(1408)에서 추정된다. 이는 대부분의 에너지가 심박수 주파수에 집중되어 있는 것으로 추정되기 때문에, 규칙적인 맥박수는 스펙트럼/자기 상관 도메인(시간에 따른 음성 특징의(of a voice feature over time))에서 "좁은(narrow)" 에너지를 생성하는 것으로 추정됨을 의미한다. 불규칙한 맥박(주로 AF)은 아마도 한 주파수에서 집중된 에너지를 갖지 않을 것이며, 아마도 평균 심박수 근처의 스펙트럼/자기 상관 값은 덜 집중될 것이다.

[0577] 예시적인 상세한 음성 특징의 주기성 결정(Exemplary detailed periodicity determination of the voice)

features)

- [0578] 본 발명의 일부 실시예에 따라, 음성 샘플에 대한 그 영향의 타이밍에 기초한 심장 활동의 타이밍을 결정하기 위한 흐름도인 도 14b를 참조한다.
- [0579] 일부 실시예에서, 음성은 단계(1412)에서 샘플링되고 적어도 하나의 음성 특징이 단계(1414)에서 추출된다. 일부 실시예에서, 음성 샘플은 샘플링된 대상의 심박수를 추출하는데 이용된다. 대안적으로 또는 추가적으로, 심박수는 심박수 모니터(a heart rate monitor)에 의해 측정된다. 일부 실시예에서, 음성 특징의 자기 상관은 심박수 주파수를 포함하는 복수의 주파수에서 단계(1418)에서 계산된다. 자기 상관은 선택적으로 심박수의 주파수 주위의 미리 결정된 범위에서 특성화된다. 일부 실시예에서, 단계(1420)에서 피크 형태 파라미터(peak shape parameter)를 식별하고 그 값을 결정함으로써, 특성화가 제공된다. 예를 들어, 피크 형태 파라미터는 피크의 대역폭 및/또는 진폭 일 수 있다. 일부 실시예에서, 파라미터 값은 단계(1430)에서의 건강 상태 또는 단계(1432)에서의 병적 상태를 결정하기 위해 단계(1422)에서의 임계 값과 비교된다.
- [0580] **음성 특징의 성문 펄스 분석의 예(Example for Glottal pulse analysis of voice features)**
- [0581] 성문 펄스 분석(Glottal pulse analysis)은 입, 혀, 입술 및/또는 구강/비강과 같은 다른 언어 기관의 영향없이 잠재적으로 성대(glottal cords)의 움직임으로 인한 맥박을 계산하는 기술이다.
- [0582] 본 발명의 일부 실시예에 따라, 성문 파(glottal wave)의 단순화된 예시인 도 14c를 참조한다.
- [0583] 도 14c는 목구멍 및 성문(glottis)의 단면(1441, 1442, 1443, 1444, 및 1445)을 포함하는 그림(1440)을 도시한다.
- [0584] 도 14c는 또한 X축(1452)을 따른 시간 (t)의 함수로서, Y축(1451)상의 공기 흐름 값(airflow values)을 나타내는 선(1453)을 나타내는 그래프(1450)를 포함한다.
- [0585] 목구멍 및 성문(glottis)의 단면(1441, 1442, 1443, 1444, 및 1445)의 그림(1440)은 각각 다른 성문(glottis) 상태를 보여주며 그래프(1450)는 위의 상태에 대응되는 공기 흐름 값(airflow values)을 보여준다.
- [0586] 일부 예에 따라 도 14c는 다음을 도시한다:
- [0587] 폐쇄된 성문(a closed glottis)에 대응하는 선(1453) 상의 제1 위치(1454);
- [0588] 개방되는 성문(an opening glottis)에 대응하는 선(1453) 상의 제2 위치(1455);
- [0589] 개방 성문(an open glottis)에 대응하는 선(1453) 상의 제3 위치(1456); 및
- [0590] 폐쇄되는 성문(a closing glottis)에 대응하는 선(1453) 상의 제4 위치(1457), 이 위치의 성문은 위치(1454)에 선행하여 폐쇄되는 성문을 나타내는 선(1453)보다 빠른 속도로 폐쇄됨.
- [0591] 폐쇄된 성문(a closed glottis)에 대응하는 선(1453) 상의 제5 위치(1458);
- [0592] 성문 파는 스피치 신호로부터 선택적으로 추출된다.
- [0593] 일부 실시예에서, 스피치 신호로부터의 성문 펄스(a glottal pulse) 계산은 선택적으로 심박수 상태(a heart rate condition)를 결정하기 위해 사용된다. 첫 번째 단계는 선택적으로 성문 펄스를 얻는 것이다.
- [0594] 일부 실시예에서, 성문 펄스는 상기 언급된 "Estimation of the glottal pulse from speech or singing voice"이라는 제목의 문헌에 기재된 바와 같이 선택적으로 계산된다.
- [0595] 일부 실시예에서, 다음 파라미터 중 하나 이상이 성문 펄스 신호(glottal pulse signal)로부터 계산된다:
- [0596] 성문 상승 시간(a glottal rise time) (10%에서 90 %로 시간 증가);
- [0597] 성문 하강 시간(a glottal fall time) (90 %에서 10 %로 시간 단축);
- [0598] 성문의 개방/폐쇄 기간(an open/close duration for the glottis);
- [0599] 주기 (성문 펄스 신호의 피크 사이의 거리);
- [0600] 성문 펄스 신호의 피크 높이 및/또는 폭; 및
- [0601] 성문 펄스 신호의 여러 주기의 스펙트럼.

- [0602] 일부 실시예에서, 파라미터의 통계는 선택적으로 비교되고, 결정은 선택적으로 심박수 변동성으로 이루어진다.
- [0603] 일부 실시예에서, 전술한 파라미터들 중 하나 이상이 본 명세서에 설명된 프로세싱 기술에서 "스피치 특징(speech features)"으로 이용된다.
- [0604] 성문 신호(glottal signal)를 사용하는 것의 잠재적 이점은 심장 펄스와 관련이 없는 (혀 및/또는 입 움직임과 같은) 음성 기관의 "자연적" 변화가 잠재적으로 고려되지 않고, 잠재적으로 보다 정확한 추정이 획득될 수 있는 점이다.
- [0605] **음성 특징의 스펙트럼 분석의 예(Example for spectral analysis of voice features)**
- [0606] 본 발명의 일부 실시예에 따라, 음성 특징의 스펙트럼 특성을 이용하여 심박수 병리에 대한 확률의 결정을 나타내는 도 15a 내지 도 15d를 참조하면, 도 15a 및 도 15c는 높은 진폭(high amplitude), 낮은 폭의 피크(low width peak)를 특징으로 하는 건강한 패턴을 나타내고, 도 15b 및 15d는 낮은 진폭(low amplitude), 높은 폭의 피크(high width peak)를 특징으로 하는 AF 패턴을 도시한다. 빨간색 원은 심박수의 스펙트럼 피크를 나타낸다 (앞서 설명한 방법으로 계산됨).
- [0607] 심장 펄스가 음성을 변조하는 것으로 추정되기 때문에, 상이한 심장 박동 리듬이 음성 특징 스펙트럼 패턴에서 상이한 패턴을 생성할 것으로 예상된다. 스펙트럼 피크의 대역폭은 심장 파라미터의 주기적인 범위(the periodic extent)와 일치할 것으로 예상된다. 예를 들어, AF 펄스는 매우 불규칙하며 스펙트럼 영역에서 다양한 피크를 생성하므로 더 넓은 대역폭을 갖는다. 피크 대역폭은 비교적 크며, AF 및 기타 심장 박동 장애에 공통적인 비 주기적 변화를 반영한다. 이에 비해, 규칙적인 펄스, 즉 주기적 펄스는 음성의 일정한 변조를 유발할 것으로 예상되며, 일반적으로 스펙트럼 영역에서, 아마도 정확히 심박수에서, 하나의 메인 피크를 생성한다. 피크는 주기적 변화(예를 들어, 동리듬 타입 심장 펄스(sinus type heart pulse))를 반영하기 위해 좁을 수 있다.
- [0608] 도 15a 내지 도 15d에 도시된 예는 MFCC 파라미터의 1차 미분을 사용하여 계산된다. 미분 연산은, "자연적"(음성학적(phonetic)) 스피치 정보와 내용을 대부분 제거하기 때문에, 잠재적으로 유용하다. 이 방법은 스펙트럼 분석에 대안적으로 또는 추가적으로 자기 상관 계수를 사용할 수 있음을 유의해야 한다.
- [0609] **예시적인 전체 부정맥 확률 통합(Exemplary overall Arrhythmia probability integration)**
- [0610] 일부 실시예에서, 예를 들어, 도 3의 블록 362 내지 364 및 도 4의 블록 460 내지 466에 도시된 바와 같이, 심장 펄스에서 계산된 불규칙성의 확률(probability of irregularity)은 (i) 음성 특징 분포(voice feature distribution) (ii) 음성 특징 주기성(voice feature periodicity) (iii) 펄스 간격 데이터(pulse interval data) 중 적어도 하나에 기초한다.
- [0611] 일부 심장 병리는 일반적으로 높은 심박수로 나타난다. 일부 실시예에서, 높은 심박수는 70BPM 초과, 또는 90BPM 초과, 또는 100BPM 초과, 또는 120 BPM초과로 정의된다. 일부 실시예에서, 이들 병리에 대해, 낮고 안정적인 BPM이 검출될 때(예를 들어, 단일 주파수 주위에 집중된 스펙트럼 도메인에서의 높은 에너지), 이들 특정 병리에 대한 더 낮은 확률이 추정 될 수 있다.
- [0612] 선택적으로, 펄스 데이터 간격의 시리즈의 총 변동은 다음과 같이 계산된다. 예를 들어,
- $$diff(i) = (T_{i+1} - T_i) / average(T_i)$$
- $$TV = \sum_{i=1} diff(i)^2$$
- [0613]
- [0614] 심장 펄스의 시간(time of heart pulse): T_i
- [0615] 두 펄스 사이 거리(RR): $|T_{i+1} - T_i|$
- [0616] 실험적으로 AF 심박수와 정상 심박수의 TV 값(TV values)이 다른 값을 갖는 것으로 밝혀졌다. 이것은 짧은 일련의 심장 펄스(a short series of Heart pulses)에 기초한 AF 검출의 임계 값을 결과로 낳았다(예를 들어, 도 11b의 블록 1122 내지 1130, 및 1132에 예시된 바와 같이).
- [0617] 일부 실시예에서, 총 변동이 미리 결정된 임계 값을 초과하면, AF가 검출된다. 선택적으로, 특정 양성 추정

(positive estimations; 예를 들어, 심장 병리에 대한 높은 확률)에 대해, 예를 들어, 사용자에게 통지를 보내고 및/또는 간병인에게 통지함으로써 추가적인 건강 진단이 권장된다.

[0618] 일부 실시예에서, 선택적으로 적은 확신의 추정(less certain estimations; 예를 들어, 심장 병리에 대한 중간 확률)을 위해 추가 음성 테스트가 개시된다. 일부 실시예에서, 대상은 추가적인 검사를 위해, 사전 선택된 일부 부부 스피치 소리를 발음하도록 요청된다. 선택적으로, 미리 선택된 스피치 소리(예를 들어, 3 내지 6 초의 음성 - WahW 의 긴 반복)이 추가 정보를 획득하기 위해 이용된다.

[0619] 일부 실시예에서, 심장 병리의 가능성이 낮은 대상의 경우, 대상의 데이터가 선택적으로 후속 시험에서의 추가 분석을 위해 또는 건강한 트레이닝 모델 데이터베이스(training model database)에 추가하기 위해 저장된다.

[0620] 일부 실시예에서, 병적 상태를 가질 확률이 확립되면, 대상은 잠재적으로 검출 감도를 증가시키는 추가 진단 방법을 받을 수 있다. 예를 들어, 대상은 병원 기반 검진 및/또는 PPG(photoplethysmography)와 같은 보충적인 가정용 어플리케이션(complementary home-use applications) 및/또는 휴대용 ECG 모니터(예를 들어, 스마트폰에서 이용될 수 있는)를 이용하도록 언급될 수 있다. 대상체에게 보완적인 가정용 어플리케이션 중 하나를 적용하도록 통지함으로써, 정확한 진단에 대한 가능성을 증가시킨다.

[0621] 일부 실시예에서, 보완적 어플리케이션은 수동으로 작동된다. 선택적으로, 보완적 어플리케이션은, 예를 들어, 하루에 한 번, 하루에 두 번, 하루에 세 번 등의 미리 정해진 요법에 따라, 선택적으로 병적 상태를 진단받을 확률에 따라 작동된다. 예를 들어, 병적 상태를 가질 확률이 더 높으면, 더 많은 수의 보완적 어플리케이션 동작이 바람직하다. 대안적으로 또는 추가로, 보완적 어플리케이션은 병적 상태 추정시 즉시 작동된다.

[0622] **예시적인 치료 모니터링(Exemplary treatment monitoring)**

[0623] 본 발명의 일부 실시예에 따라 본 명세서에 개시된 시스템 및 방법에 의해 제공되는 가능한 치료 및 모니터링(treatment and monitoring)의 타임 라인을 도시하는 도 16을 참조한다.

[0624] 일부 실시예에서, 검사(screening; 1602)은 다수의 사람들 또는 단일 사람에게 제공된다. 일부 실시예에서, 충분한 미리 결정된 확률로 심장 상태의 결정 후, 치료가 제공된다. 대안적으로 또는 추가적으로, 충분한 미리 결정된 확률로 심장 상태를 결정한 후, 제2 진단 테스트 또는 예를 들어 ECG 모니터링 및/또는 PPG(photoplethysmography) 및/또는 미리 결정된 발성(vocalization)을 포함하는 제2 음성 샘플과 같은 검증 테스트가 수행된다. 선택적으로, 치료는 전형적으로 일정 기간에 걸친 약학 치료(1630)를 포함한다. 대안적으로 또는 추가적으로, 치료는 일반적으로 단일 발생(single occurrence)으로 제공되는 절차적 치료(1640)를 포함한다.

[0625] 일부 실시예에서, 대상이 치료를 시작하려고 할 때, 단계(1604)에서 그의 심장 상태의 모니터링이 치료가 시작되기 전에 제공된다. 선택적으로, 상태의 중증도를 모니터링하며 이러한 모니터링 후에 치료가 변경될 수 있습니다. 대안적으로 또는 추가로, 단계(1606)에서, 선택적으로 약물의 효과를 모니터링하기 위해 치료 동안 모니터링이 제공된다. 대안적으로 또는 추가적으로, 선택적으로 상태의 재발을 식별하기 위해, 치료 후 단계(1608)에서 모니터링이 제공된다.

[0626] 일부 실시예에서, 모니터링 스케줄은 대상의 진단된 상태에 기초한다. 예를 들어, 매일 몇 시간 동안 심방 세동이 있는 심한 AF의 경우, 하루에 한 번 대상을 샘플링하는 것으로 충분할 수 있습니다. 반면에, 가벼운 AF 사례로 대상을 진단한 경우 더 검출을 달성하기 위해 더 빈번한 샘플링이 요구된다.

[0627] **일반**

[0628] 본 명세서에 사용된 용어 "약"은 $\pm 25\%$ 를 지칭한다.

[0629] 용어 "포함하다", "포함하는", "갖는" 및 이들의 활용은 "포함하지만 이에 한정되지 않는"을 의미한다.

[0630] "구성되는"이라는 용어는 "포함하고 이에 한정되는"을 의미한다.

[0631] "본질적으로 구성되는"이라는 용어는 구성, 방법 또는 구조가 추가 성분, 단계 및/또는 부품을 포함할 수 있지만, 추가 성분, 단계 및/또는 부품이 청구된 구성, 방법 또는 구조물의 기본적인 새로운 특성을 실질적으로 변경하지 않는 경우에만 해당된다.

[0632] 본 명세서에서 사용된 바와 같이, 단수 형태 문맥상 명백하게 다르게 지시되지 않는 한 복수의 언급을 포함한다.

- [0633] 본 명세서 전체에서, 본 발명의 다양한 실시예는 범위 형식으로 제시될 수 있다. 범위 형식의 설명은 단지 편의 및 간결성을 위한 것이며 본 발명의 범위에 대한 제한으로 해석되어서는 안된다는 것을 이해해야 한다. 따라서, 범위의 설명은 모든 가능한 하위 범위 및 그 범위 내의 개별 수치를 구체적으로 개시한 것으로 간주되어야 한다. 예를 들어, 1 내지 6과 같은 범위의 설명은 1 내지 3, 1 내지 4, 1 내지 5, 2 내지 4, 2 내지 6, 3 내지 6 등 뿐만 아니라 예를 들어, 1, 2, 3, 4, 5 및 6과 같은 범위 내의 개별 숫자와 같은 하위 범위를 구체적으로 개시한 것으로 간주되어야 한다. 이는 범위의 폭에 관계없이 적용된다.
- [0634] 수치 범위가 본 명세서에 지시될 때마다, 지시된 범위 내에 임의의 인용된 숫자 (분수 또는 정수)를 포함하는 것으로 의도된다. 제1 지시 번호와 제2 지시 번호 "사이의 범위" 및 제1 지시 번호 "부터" 제2 지시 번호의 "까지의 범위"는 본 명세서에서 상호 교환적으로 사용되며 제1 및 제2 지시 번호 및 그 사이의 모든 분수 및 정수를 포함하는 것으로 이해되어야 한다.
- [0635] 명확성을 위해, 별도의 실시예들과 관련하여 설명된 본 발명의 특정 특징들은 또한 단일 실시예에서 조합하여 제공될 수 있다는 것으로 이해되어야 한다. 반대로, 간결하게, 단일 실시예와 관련하여 설명된 본 발명의 다양한 특징들은 개별적으로 또는 임의의 적절한 하위 조합으로 또는 본 발명의 임의의 다른 실시 예에서 적합한 것으로 제공될 수 있다. 다양한 실시예들과 관련하여 설명된 특정 특징들은 실시 예들이 그러한 요소들 없이 동작하지 않는 한, 이들 실시예들의 필수 특징으로 간주되지 않아야 한다.
- [0636] 본 명세서에 기술되고 하기 청구 범위의 청구된 바와 같이 본 발명의 다양한 실시예는 뒤따르는 예시로부터 뒷받침된다.
- [0637] 예시
- [0638] **음성 샘플을 사용한 심장 상태 결정의 예(Example for heart condition determination using a voice sample)**
- [0639] 상기 상세한 설명과 함께 본 발명의 일부 실시예를 비 제한적인 방식으로 도시하는 하기 예시를 참조한다.
- [0640] 일부 실시예에서, 인간의 음성이 샘플링된다. 일부 실시예에서, 음성 샘플은 원하는 음성 세그먼트를 추출하기 위해 분석된다. 예를 들어, 침묵 기간(silence periods) 제거 및/또는 유성 세그먼트(voiced segments)만 유지될 수 있다. 선택적으로 특정 종류의 모음(vowel)만 유지된다. 예를 들어, /ah/ 모음만, 또는 /eh/ 모음만, 또는 /oh/ 모음만, 또는 /uh/ 모음만 또는 /eeh/ 모음만 유지된다. 일부 실시예에서, 유지된 음성 세그먼트들은 연결되어 잠재적으로 시간 순서를 따르는 것이 아니라 연속적인 데이터로 결과가 나타난다. 대안적으로, 유지된 음성 세그먼트는 시간에 따른 시퀀스를 유지하면서 세그먼트의 타이밍 정보를 유지한다.
- [0641] 일부 실시예에서, 피치와 같은 음성 특징은, 유성 세그먼트(voiced segments)를 획득하기 위해 필터링되고 연결된 음성 세그먼트(voice segments)로부터 추출된다. 일부 실시예에서, 피치의, 음성 캐비티(voice cavities) 및 성대(vocal cords)의 자연적 변동으로 인한 변화 및 스피치의 발성(phonation)의 변화인 자연적 변동이 제거된다. 일부 실시예에서, 자연적 변동은, 선택적으로 심장 박동 관련 변화를 검출하는데 더 높은 민감도를 달성하도록 필터링된다. 일부 실시예에서, 조정된 피치를 갖는 음성 세그먼트는 자연적 변동이 제거된 후에 연결된다. 선택적으로 연속적인 음성 세그먼트의 융합 지점(fusion points)이 평탄화된다.
- [0642] 일부 실시예에서, 평탄화되고 연결된 피치의 데이터는 정규화된 표준 편차 (the normalized standard deviation(N-STD))를 도출하기 위해 분석된다. 일부 실시예에서, N-STD의 값은 임계 값과 비교되며, 이보다 높으면 샘플링된 대상이 병적 심장 상태를 갖는 것으로 결정되고 그 미만으로 샘플링된 대상은 건강한 심장 상태를 갖는 것으로 결정된다.
- [0643] 일부 실시예에서, 동일한 음성 샘플은 제2 분석을 거친다. 선택적으로 음성 세그먼트가 추출된다. 일부 실시예에서, R 펄스는 각각의 유지된 음성 세그먼트에서 식별된다. 선택적으로, R 펄스는 제거된 음성 세그먼트에 연속적인 R 펄스의 적어도 일부를 식별함으로써 외삽된다. 일부 실시예에서, RR 간격은 연속적인 R 펄스를 위치시킴으로써 계산된다. 예를 들어, 동일한 음성 세그먼트에 하나 이상의 R 펄스를 위치시킴으로써 및/또는 무시될 수 있는 짧은 갭을 갖는 상이한 음성 세그먼트 상에 R 펄스를 위치시킴으로써 및/또는 외삽된 펄스(extrapolated pulse)에 연속적인 R 펄스를 식별함으로써 계산된다.
- [0644] 일부 실시예에서, RR 간격은 기준(reference), 예를 들어 AF와 같은 병적 심장 상태의 특징 인 것으로 알려진 RR 간격을 갖는 기준과 비교된다. 대안적으로 또는 추가로, 기준은 동일한 샘플 대상의 건강한 RR 간격 시퀀스를 포함한다. 대안적으로 또는 추가로, 기준은 상이한 건강 및/또는 병적 대상의 샘플을 포함한다.
- [0645] 일부 실시예에서, 기준과 일치하는 RR 간격 시퀀스의 확률은 임계 값과 비교된다. 예를 들어, RR 간격의 확률

이 병적 시퀀스와 비교 될 때 임계 값보다 높고, 및/또는 건강한 시퀀스와 비교할 때 임계 값보다 낮으면 병적 상태가 결정된다. 대안적으로 또는 추가적으로, RR 간격의 확률이 병적 시퀀스와 비교 될 때 임계 값 미만이고, 및/또는 건강한 서열과 비교할 때 임계 값보다 높으면, 건강 상태가 결정된다.

[0646] 일부 실시예에서, 음성 샘플은 제3 분석을 거친다. 분석 수행 순서는 다양 할 수 있으며, 각각의 분석은 다른 두 분석과 독립적으로 또는 조합하여 단독으로 수행될 수 있음(an independent manner from the other two analyses, in combination or alone)에 유의해야 한다.

[0647] 일부 실시예에서, 음성 샘플은 음성 특징을 추출하기 위해 분석된다. 일부 실시예에서, 선택적으로 음성 샘플 자체로부터 심박수를 추출함으로써 심박수가 결정된다. 일부 실시예에서, 심박수 주파수에서의 음성 특징의 자기 상관 및/또는 스펙트럼이 계산된다. 보컬 영역(vocal region)으로 들어오고 나가는 혈류는 아마도 심장 박동과 관련이 있다. 음성 특징의 변동성의 적어도 일부는 이러한 혈액의 흐름 및/또는 심실 박동 및/또는 심방 박동에 의해 영향을 받을 것으로 추정된다. 따라서, 건강한 주기적인 심박수가 있는 경우, 심박수에서의 자기 상관 분석 및/또는 스펙트럼 분석은 심박수에서 보컬 혈액 영향(vocal blood influence)의 높은 주기성을 나타내는 피크 형태를 제공 할 것으로 추정된다. 반면에 부정맥이 있으면, 심박수에서 자기 상관은 희미한 피크(smeared peak)를 제공할 것이며, 이는 혈액의 영향이 그리 주기적이지 않음을 나타낸다.

[0648] 일부 실시예에서, 피크 형태의 파라미터는 임계 값과 비교된다. 예를 들어, 피크 형태의 파라미터는 대역폭, 및/또는 높이 및/또는 피크의 총 면적을 포함할 수 있다. 일부 실시예에서, 건강 또는 병리적 심장 상태를 결정하기 위해 피크 형태 파라미터를 임계 값과 비교한다.

[0649] 일부 실시예에서, 상기 분석에 기초한 건강 및/또는 병리적 심장 상태에 대한 가중 확률(weighted probability)이 계산된다.

[0650] 일부 실시예에서, 만성 질환은 음성 파라미터의 변화에 의해 잠재적으로 검출된다. 어떤 경우에는 심방 세동(AF)이 좌심방의 확장을 유발한다(대부분의 병적 AF 사례에서 발생). 확장은 후두 신경(recurrent laryngeal nerve)에 영향을 미쳐 지속적으로 목소리를 변화시킨다. 변화는 피치 신호의 파라미터에서 잠재적으로 나타나며 분석 및 검출 될 수 있다.

[0651] 이를 통해 심장이 정상 속도로 뛰는 경우에도 만성 상태의 AF 환자를 감지 할 수 있다.

[0652] 심박수 추출의 예(Example for heart rate extraction)

[0653] 본 발명의 일부 실시예에 따라, 음성 샘플로부터의 심박수 추출을 위한 예를 도시한 도 17을 참조한다. 심장 박동과 관련된 음성 특징은 심장 박동 주파수에서 피크 값을 갖는 것으로 추정된다. 예를 들어, 도 17은 74 BPM을 갖는 대상의 Wahw 소리의 음성 샘플의 분석을 예시한다. 분석을 위해 선택된 음성 특징은 dMFCC(MFCC 벡터의 1차 미분의 절대 값) 고 에너지(peak)는 0.84Hz에서 74bpm이다(dMFCC: absolute value of first derivative of MFCC vector High energy (peak) at 0.84Hz which is 74bpm).

[0654] 각 프레임으로부터 음성 특징을 도출한 후, 일부 실시 예에서, 적어도 하나의 음성 특징의 주기적인 패턴은 음성 특징의 시간에 따른 시퀀스에서 식별된다(periodic patterns of at least one voice feature are identified in a sequence over time of the voice feature). 예를 들어, 주기적 패턴의 식별은 스펙트럼 분석 및/또는 자기 상관 분석을 이용하여 수행 될 수 있다. 일부 실시예에서, 선택적으로 최대인, 분당 비트 수(BPM)의 심박수는 식별된 주기 패턴으로부터 추정된다.

[0655] 본 명세서에 기술되고 하기 청구 범위의 청구된 본 발명의 다양한 실시예는 뒤따르는 예시로부터 실험적 및 계산된 서포팅을 받는다.

[0656] 예시

[0657] 상기 상세한 설명과 함께 본 발명의 일부 실시예를 비 제한적인 방식으로 설명하는 하기 예를 참조한다.

[0658] IRB 승인 임상 연구는 이스라엘의 주요 의료 센터에서 AF 상태이면서 정상 심박수의 AF 환자를 대상으로 실시되었다.

[0659] 58명의 AF 환자 중 26명이 카디오버전(cardioversion)을 받았고, (a) AF 에피소드 동안 및 (b) 카디오버전 2시간 후 정상적인 심박수 상태 동안, 2회 기록되었다. 카디오버전은 비정상적으로 빠른 심박수 또는 기타 심장 부정맥이 전기 또는 약물을 사용하여 정상 리듬으로 전환되는 의학적 절차이다. 표 1은 현재까지 수집된 ECG/음성 기록의 수를 요약한 것이다.

표 1

그룹	환자 수
AS 및 정상 심박수 상태가 동일한 환자 (카디오버전 전후)	32
AF 환자(전체)	58
정상 심박수 상태인 대상	23

[0660]

본 발명의 일부 실시예에 따른 화자 독립적 검출 결과(speaker-independent detection results)를 나타내는 그래프인 도 18a를 참조한다.

[0662]

도 18a는 AF 상태를 갖는 58명의 환자 및 정상 심장 리듬(동리듬(sinus rhythm))을 갖는 23명의 환자의 기록이, AF 또는 정상 심장 리듬의 판단을 위한 상이한 임계 값에서, 본 발명의 예시적인 실시예에 의해 분석된 실험의 그래프 (1801)를 도시한다.

[0663]

그래프(1801)는 임계 값의 정성적 값에 대응하는 X축(1802) 및 환자의 상태가 정확하게 분석된 환자 기록의 비율(0 내지 1 범위)에 대응하는 Y축(1803)을 갖는다.

[0664]

도 18a는 진 양성(true positive, AF 조건을 갖는) 심장 상태 분류의 제1 라인(1804) 및 진 음성(true negative, 동리듬(sinus rhythm)) 심장 상태 분류의 제2 라인(1805)을 도시한다.

[0665]

도 18a는 또한, 동일한 특정 임계 값이 여기서 구체적으로 기술된 결과를 제공하는 2개의 예시적인 포인트를 도시한다: 음성 분석에 기초하여 AF의 92% 진 양성(true positive) 식별을 나타내는 제 1 포인트(1806) 및 70%의 진 음성(true negative) 식별(즉, 음성 분석에 기초한 정상적인 동리듬(sinus rhythm) 식별)을 나타내는 제2 포인트(1807).

[0666]

도 18a는 화자 독립적 결과, 즉 다수의 화자에 대해 훈련되고 알려지지 않은 환자의 심장 상태를 검출하는데 사용되는 시스템의 결과를 도시한다. 제시된 결과는 사람 음성에 기초한 심장 상태 추정에 성공한 것으로 간주된다.

[0667]

본 발명의 일부 실시예에 따른 화자 의존적 검출 결과(speaker-dependent detection results)를 나타내는 그래프인 도 18b를 참조한다.

[0668]

도 18b는 AF 또는 정상 심장 리듬에 대한 결정의 상이한 임계 값에서 본 발명의 일부 실시예에 의해 분석된, AF 조건 및 정상 동리듬(sinus rhythm) 상태에서 기록된 AF 상태를 갖는 32명의 환자의 기록이 있는 실험의 그래프 (1811)를 도시한다.

[0669]

그래프 (1811)는 임계 값의 정성적 값에 대응하는 X축(1812) 및 환자의 상태가 정확하게 분석된 환자 기록의 비율에 대응하는 Y축(1813)을 갖는다.

[0670]

도 18b는 진 양성(true positive, AF 조건을 갖는) 심장 상태 분류의 제1 라인(1814) 및 진 음성(true negative, 동리듬) 심장 상태 분류의 제2 라인(1815)을 도시한다.

[0671]

도 18a는 또한, 동일한 특정 임계 값이 여기서 구체적으로 기술된 결과를 제공하는 2개의 예시적인 포인트를 도시한다: 음성 분석에 기초하여 AF의 92% 진 음성(true negative, 동리듬(sinus rhythm) 식별을 나타내는 제1 포인트(1814) 및 70%의 진 양성(true positive, AF 조건을 갖는) 식별을 나타내는 제2 포인트(1815).

[0672]

도 18b는 화자 의존적 결과, 즉 동일한 화자의 심장 상태를 검출하기 위해, 이용되는 알려진 특정 심장 상태(건강 또는 AF)에서 취한 단일 화자의 하나 이상의 음성 기록(들)에 의해 훈련된 시스템의 결과를 도시한다. 제시된 결과는 사람의 음성에 기초한 심장 상태를 추정이 성공한 것으로 간주된다.

[0673]

비교 벤치마크(Comparative Benchmarks)

[0674]

본 발명의 일부 실시예와 비교하여, AF 에피소드를 포착하기 위해 기회 기반(opportunistic basis, 예를 들어, 일주일에 한 번)으로 사용되는 단일 타임 포인트(single time-point) ECG 장치(또는 이벤트 레코더(event recorder))에 걸리는 시간이 추정된다. 아래 표 2에는 다양한 AF 부담(AF burden) 및 테스트 주파수 시나리오(testing frequency scenarios)에서 시뮬레이션을 실행할 때 계산 결과가 요약되어 있다.

[0675]

AF 음성 바이오 마커(AF voice biomarker)가 비 자발적인 스피치의 적어도 두 개 이상의 샘플에서 검출될 때 AF 검출을 등록하도록 시뮬레이터가 구성되었다. 이러한 구성에서, 본 발명의 실시예는 94%의 검출 민감도(진 양

성(true positive)) 및 92%의 검출 특이성(동리듬 검출의 진 음성(true negative detection of sinus rhythm)의 결과를 나타냈다.

[0676] 표 2에 나타난 바와 같이, 하루 6회의 10분 에피소드의 AF 부담에 기초하여, 시뮬레이션에 의하면 ECG 기반의 주간 모니터링이 163일 이내에 AF를 검출하는 반면, 본 발명의 실시예는 8일 이내에 AF를 검출하는 것으로 나타났다.

표 2

[0677]

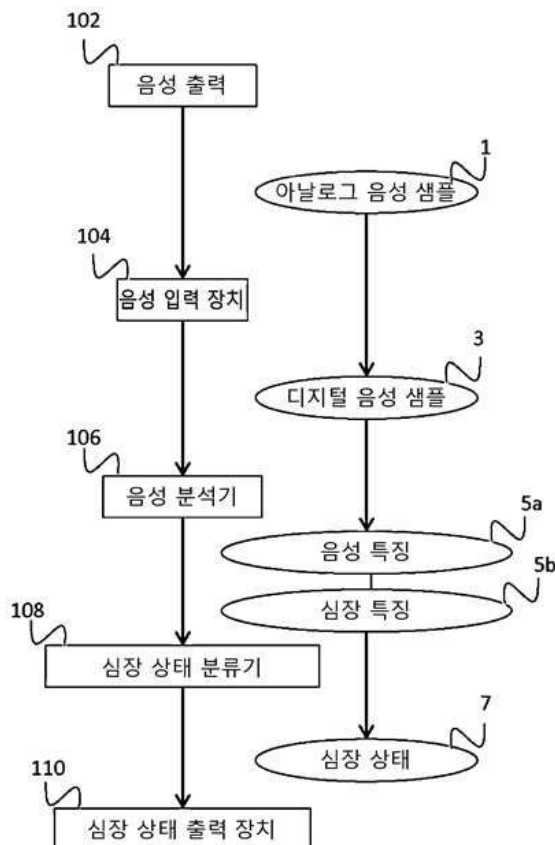
#	AF 시나리오		ECG		실시예	
	부담(분/일)	에피소드 기간(분)	주기	AF 검출을 위한 일 수	주기	AF 검출을 위한 일 수
1	60	10	1/주	163	하루 5번의 전화 x 각 4분	8
2	15	5	1/주	546	하루 5번의 전화 x 각 4분	29

[0678] 본 발명은 특정 실시예와 관련하여 설명되었지만, 많은 대안, 수정 및 변형이 당업자에게 명백할 것이다. 따라서, 첨부된 청구 범위의 사상 및 넓은 범위에 속하는 모든 그러한 대안, 수정 및 변형을 포함하도록 의도된다.

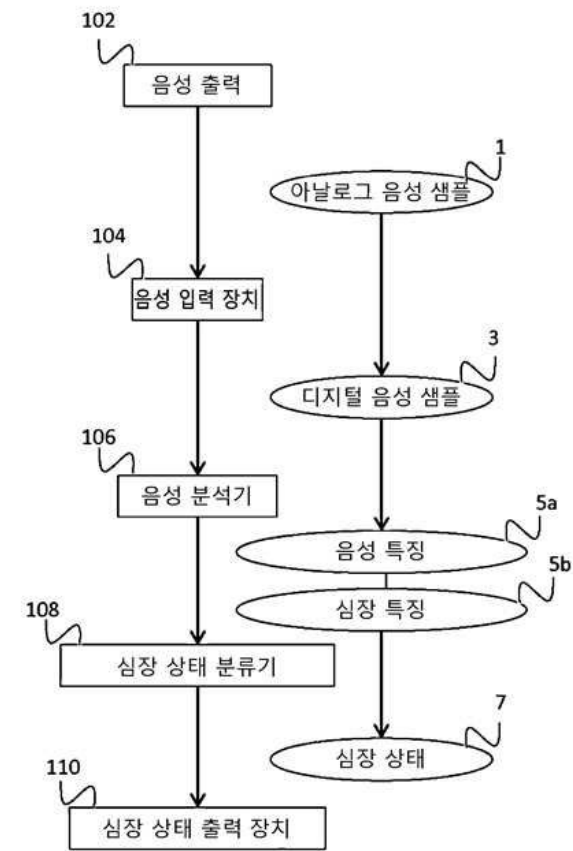
[0679] 본 명세서에 언급된 모든 간행물, 특허 및 특허 출원은 각각의 개별 간행물, 특허 또는 특허 출원이 본원에 참조로 포함된 것으로 구체적이고 개별적으로 지시된 것과 동일한 정도로 본 명세서에 참조로 전체적으로 포함된다. 또한, 본 출원에서 임의의 참조의 인용 또는 식별은 이러한 참조가 본 발명의 선행 기술로서 이용 가능하다는 인정으로 해석되어서는 안된다. 섹션 제목이 이용된 내용에 대해, 제목은 내용을 제한하는 것으로 해석해서는 안된다.

도면

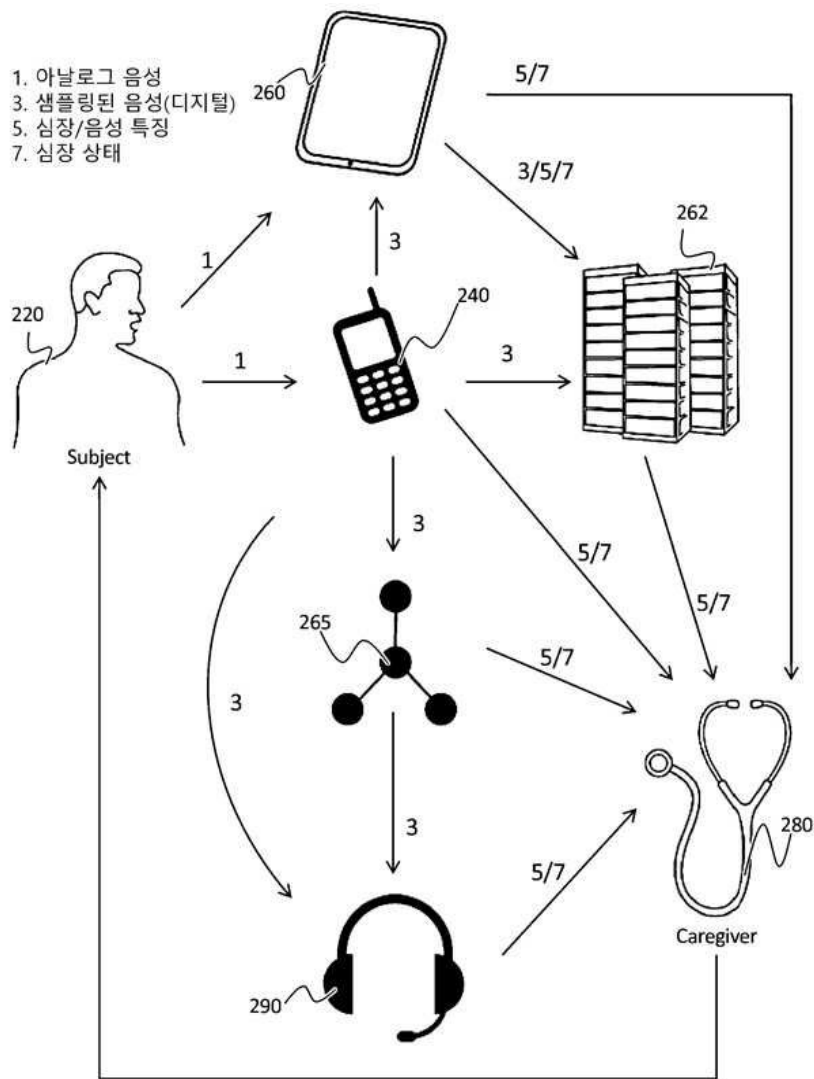
도면1a



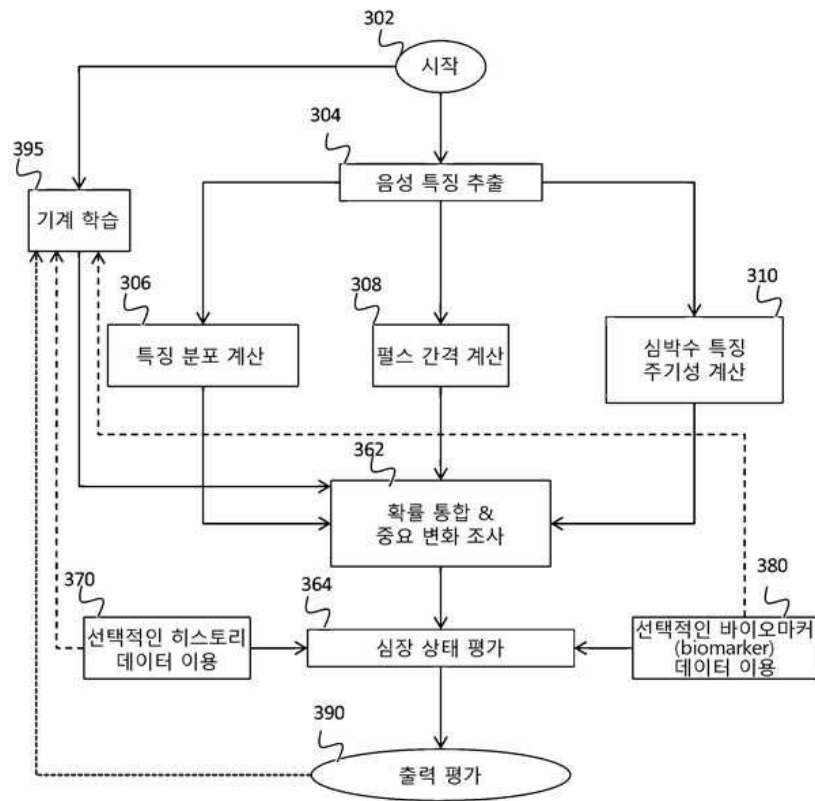
도면1b



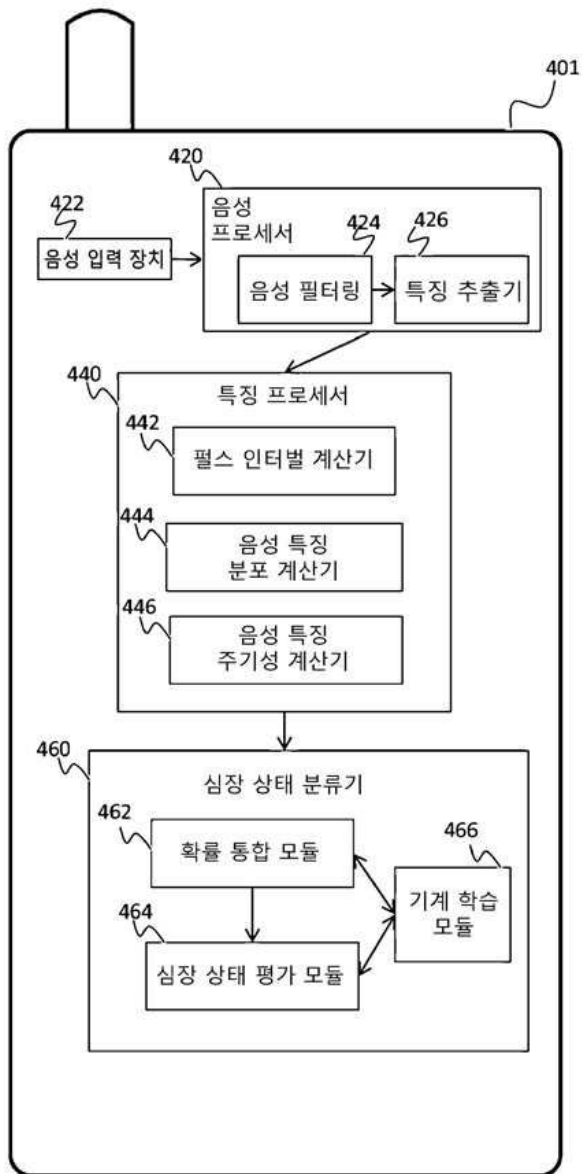
도면2



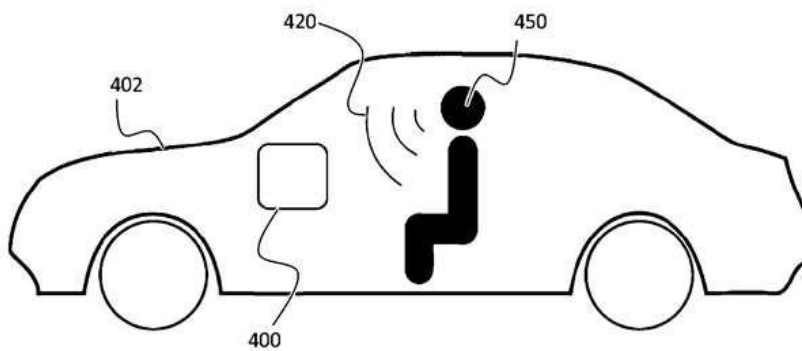
도면3



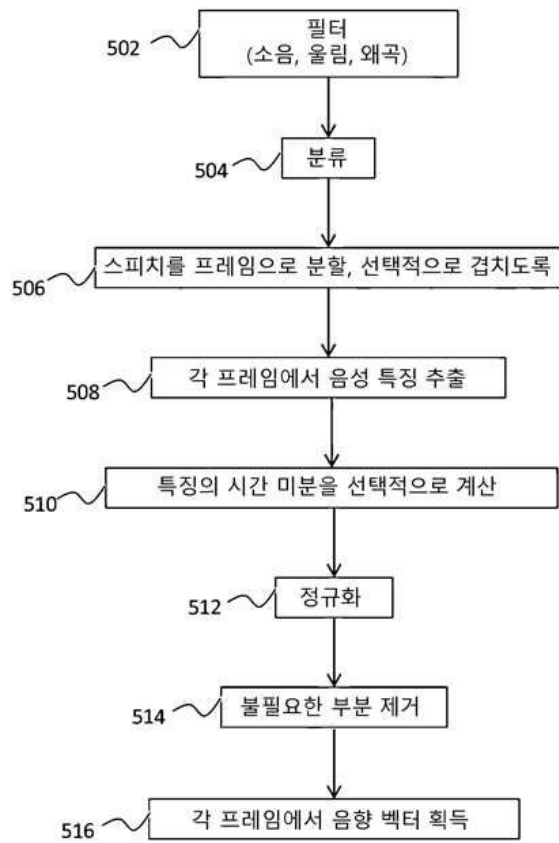
도면4a



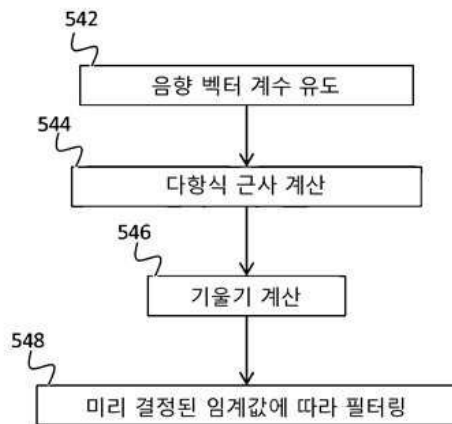
도면4b



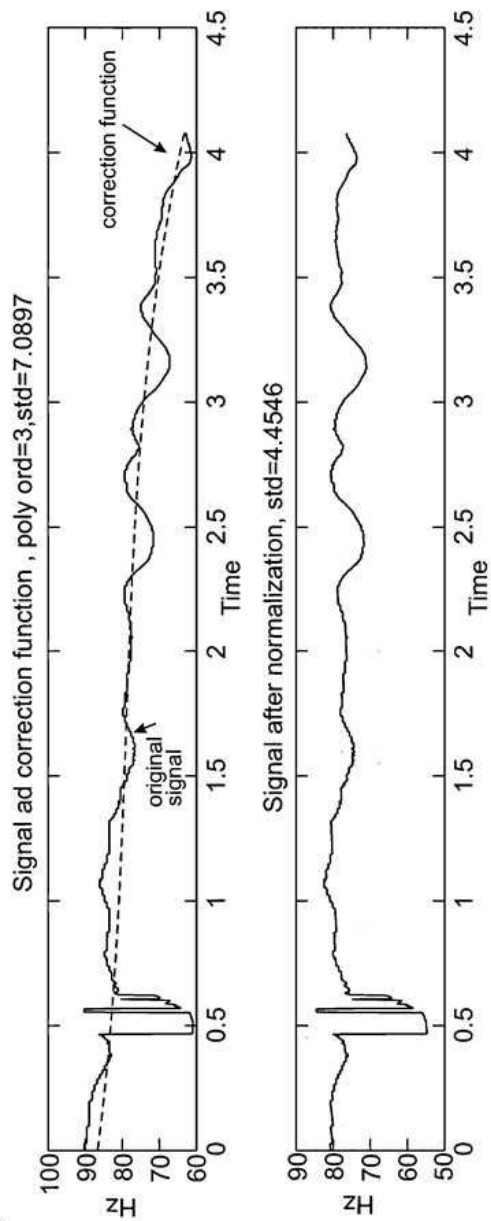
도면5a



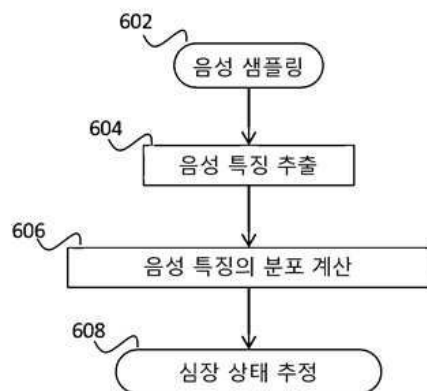
도면5b



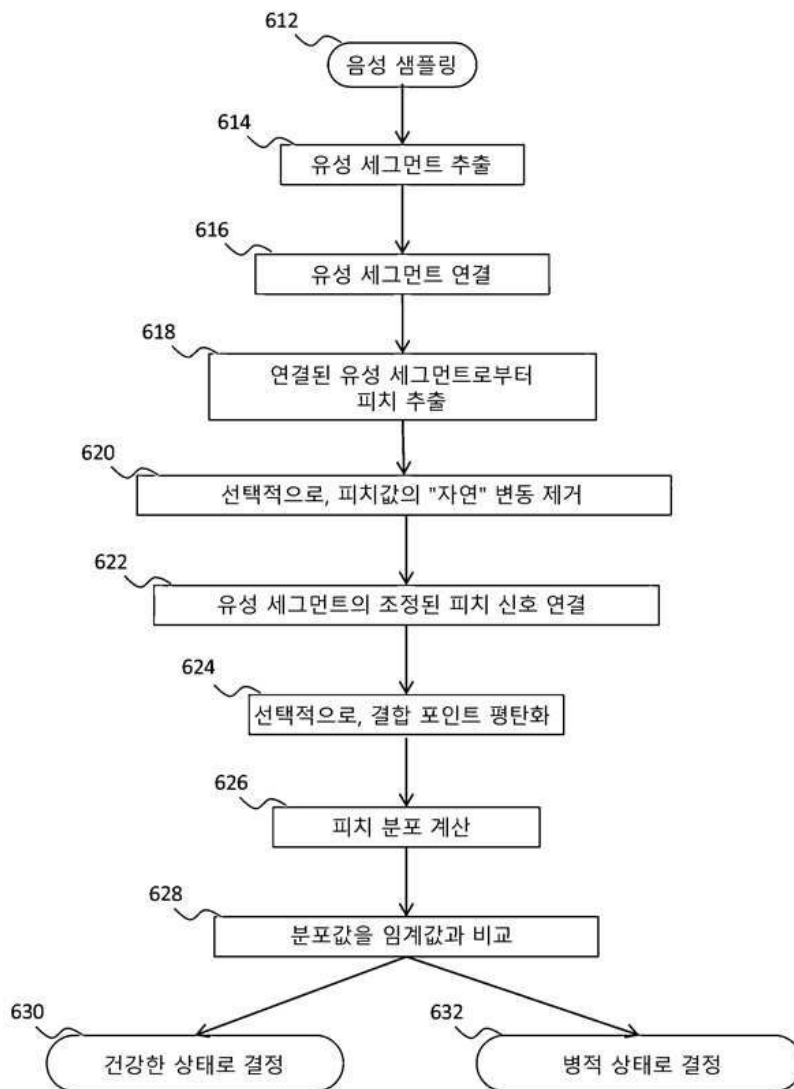
도면5c



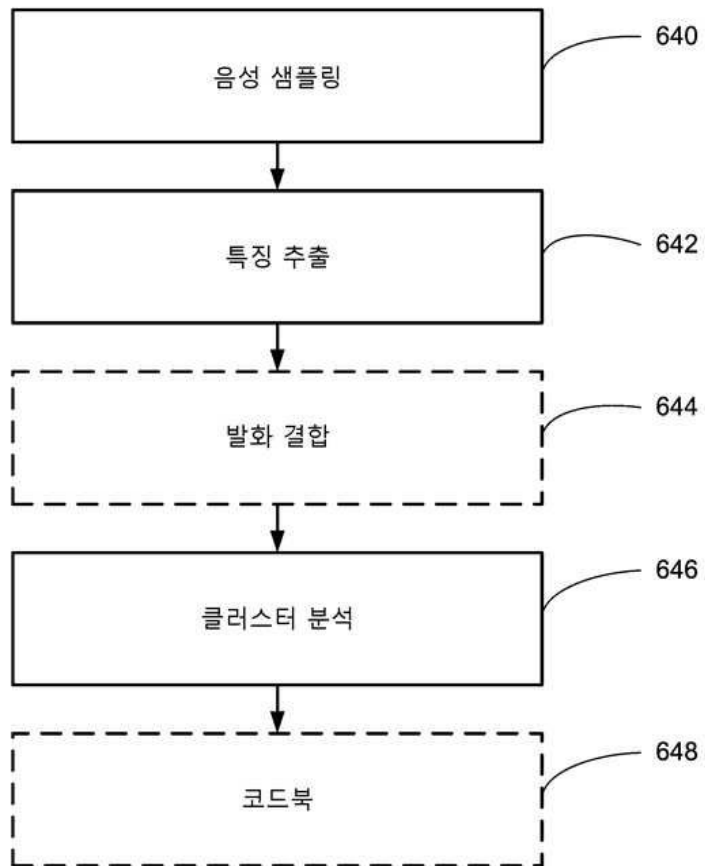
도면6a



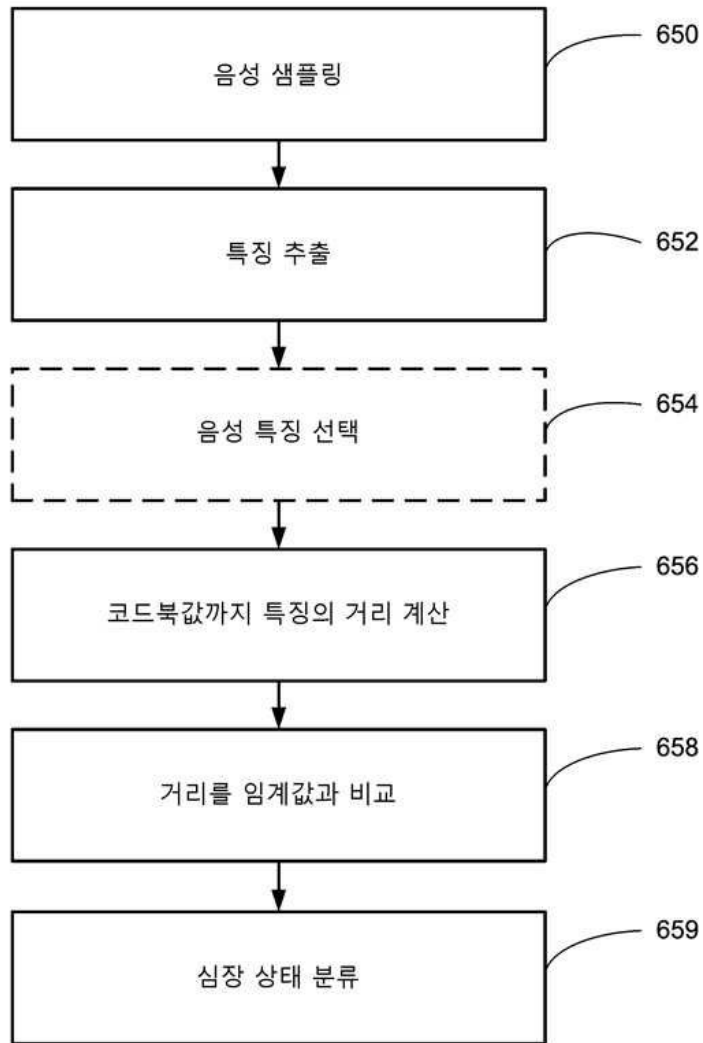
도면6b



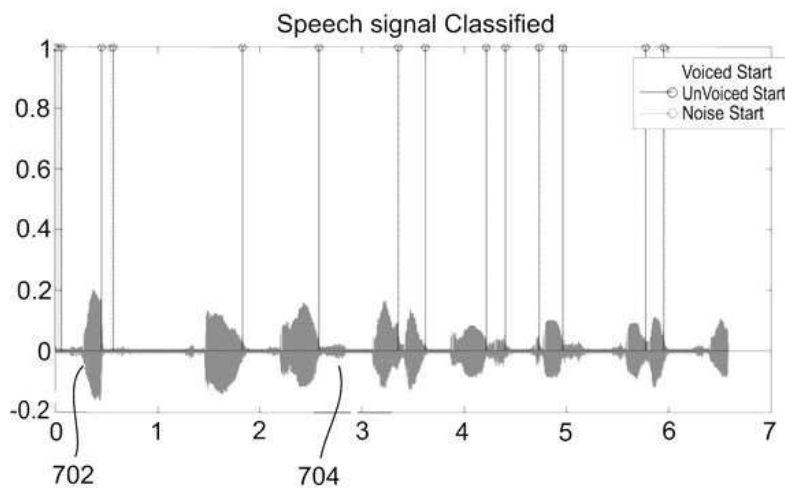
도면6c



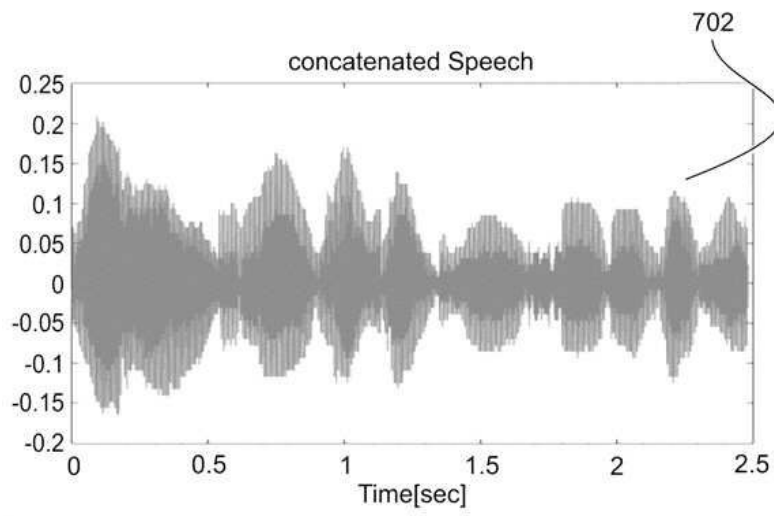
도면6d



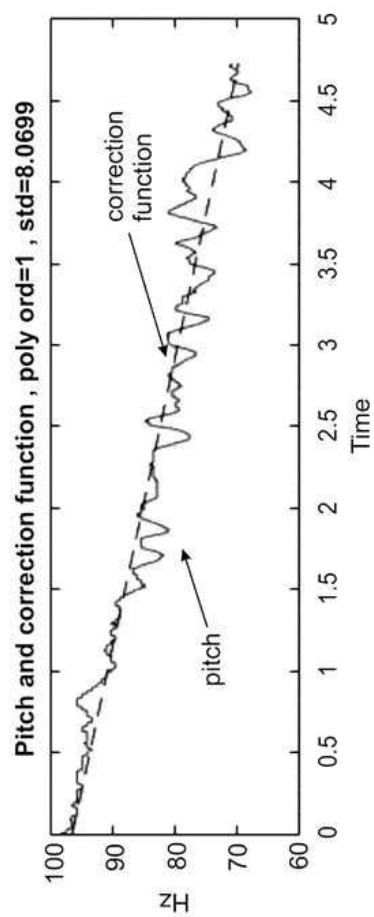
도면7a



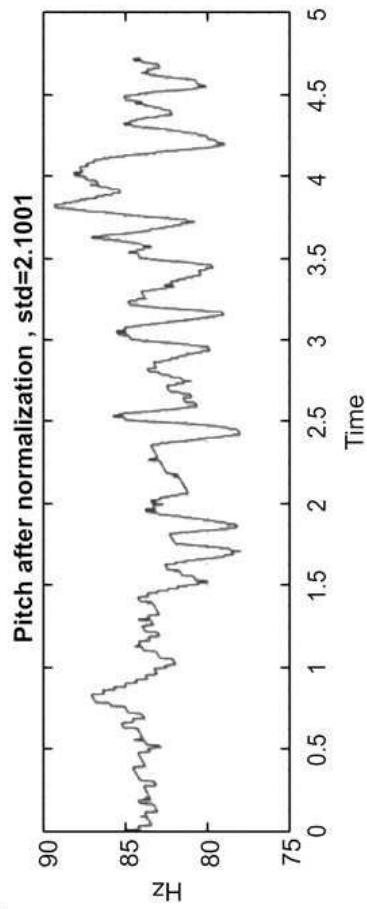
도면7b



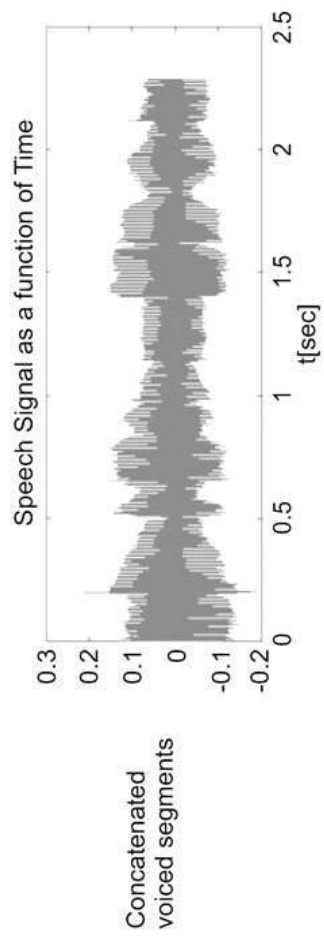
도면8a



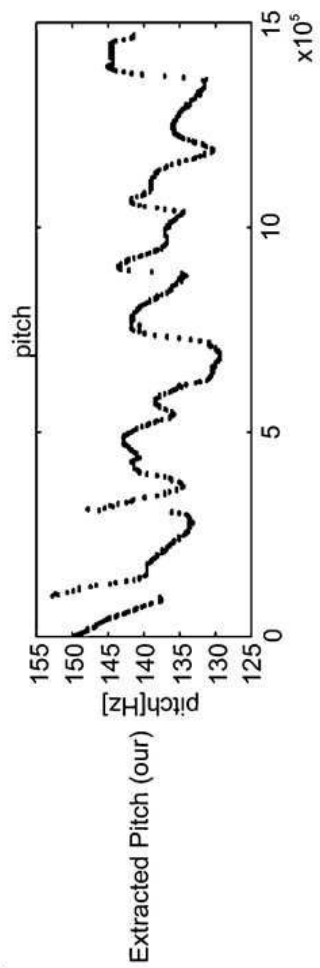
도면8b



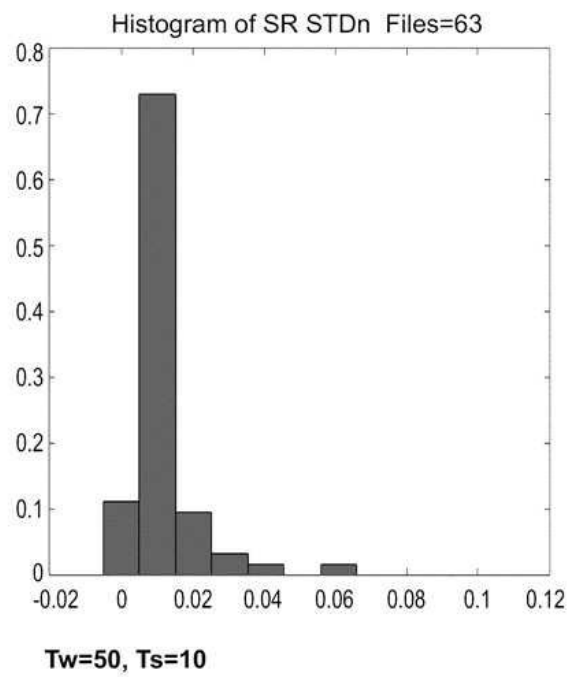
도면9a



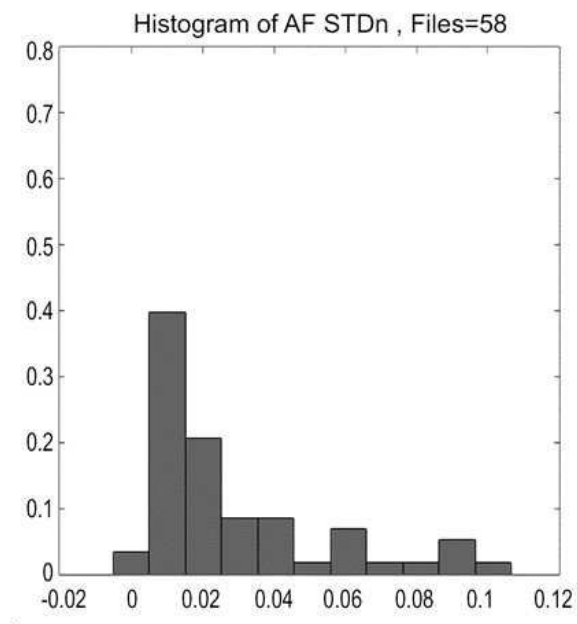
도면9b



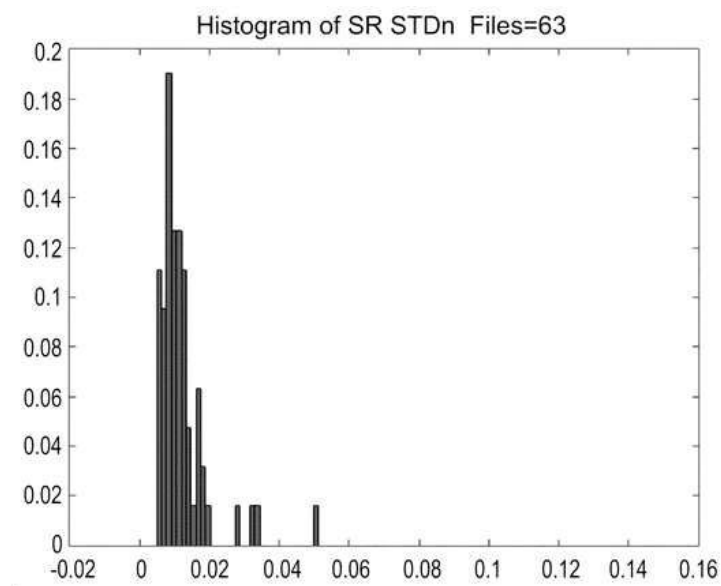
도면10a



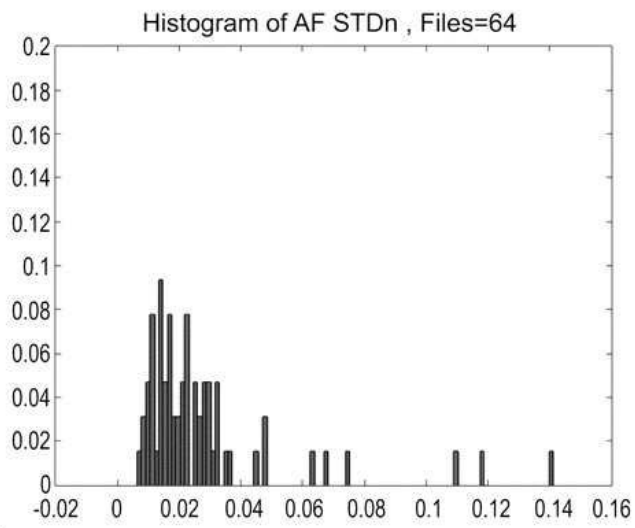
도면10b



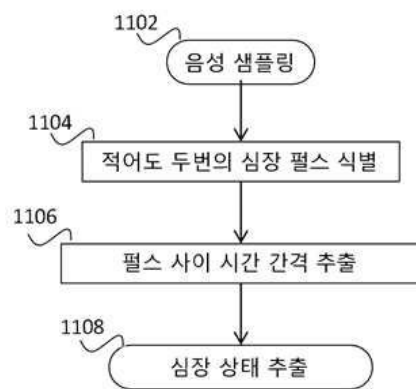
도면10c



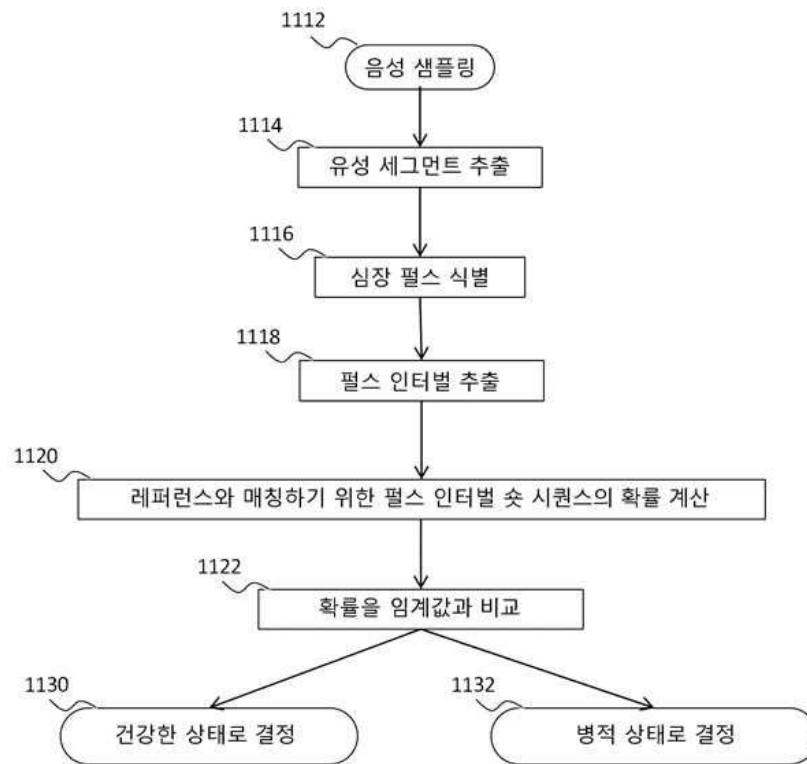
도면10d



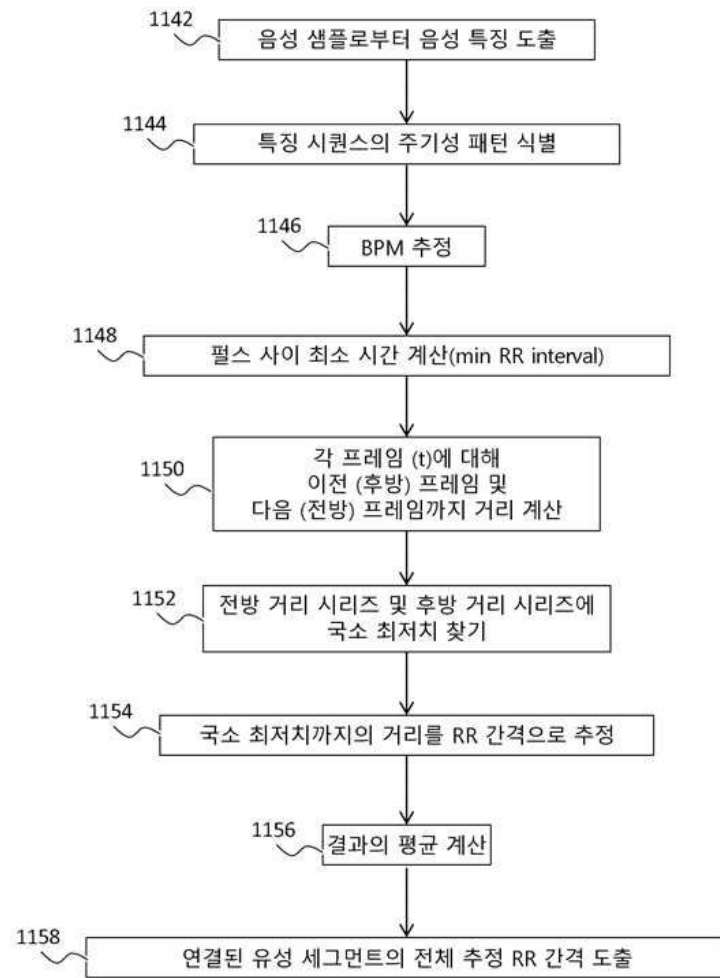
도면11a



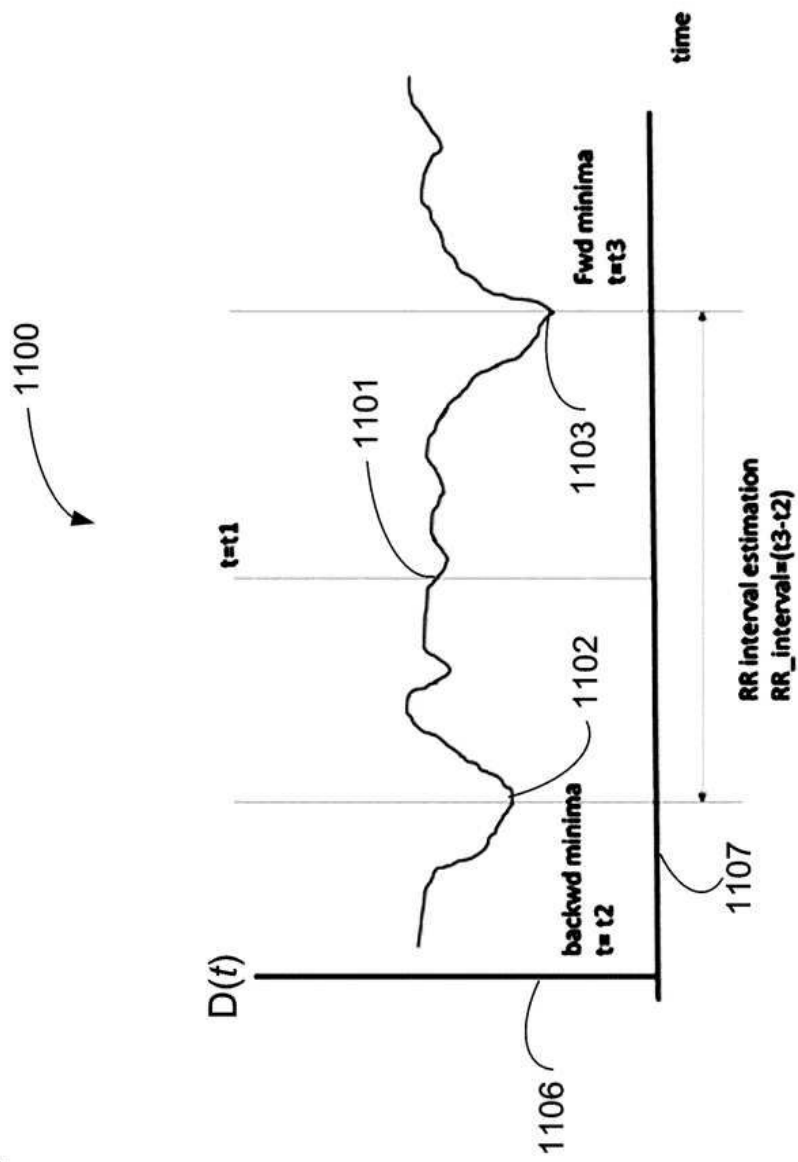
도면11b



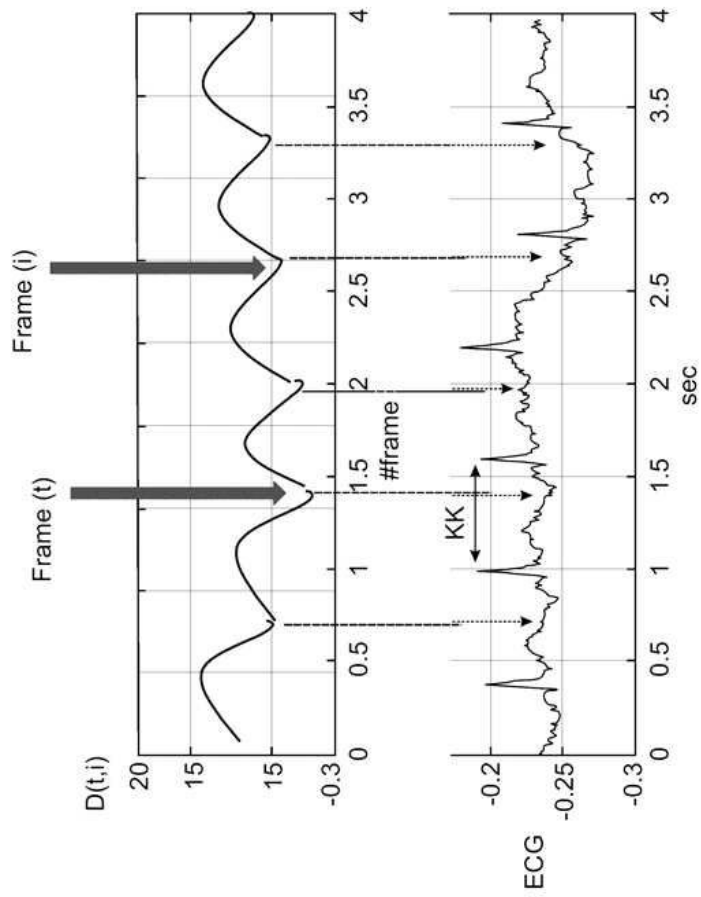
도면11c



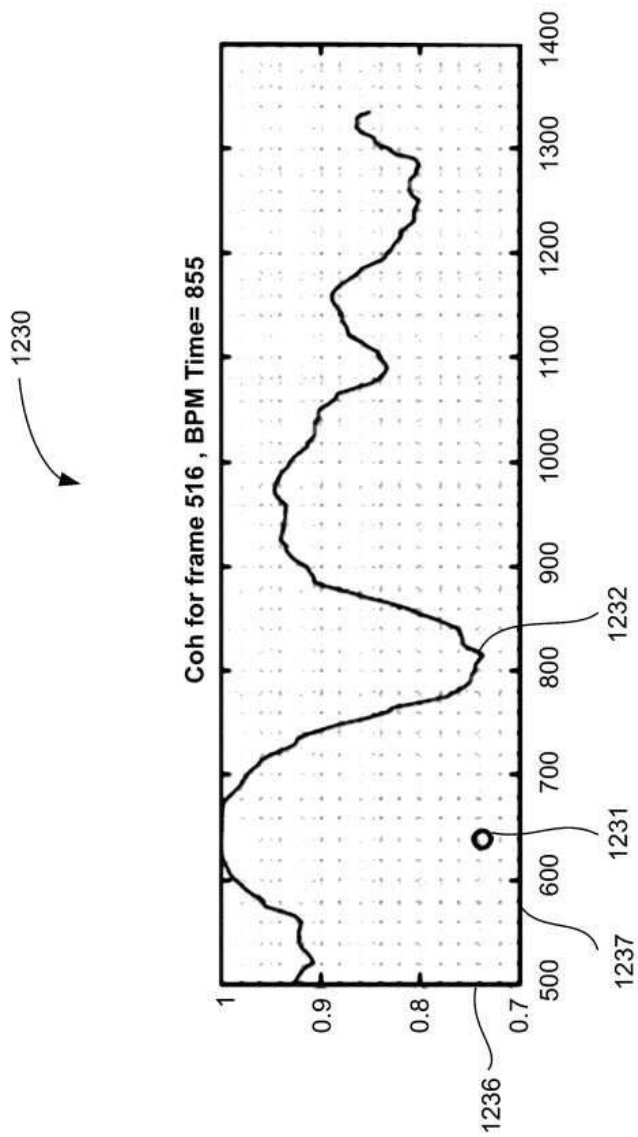
도면11d



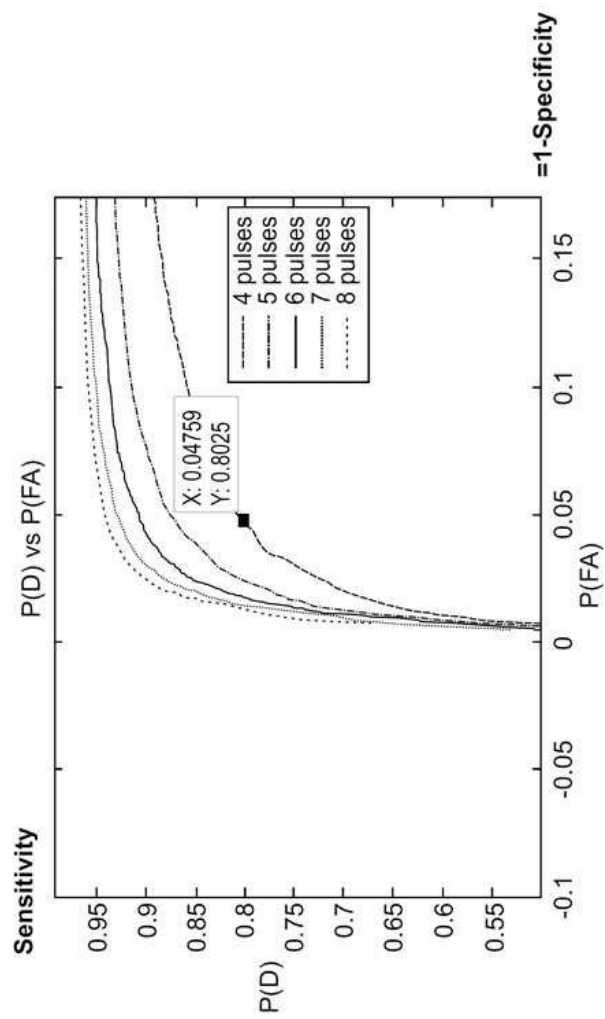
도면12a



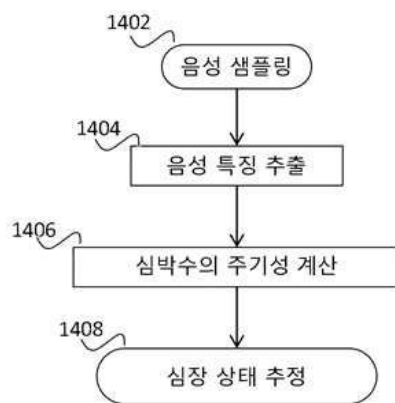
도면12b



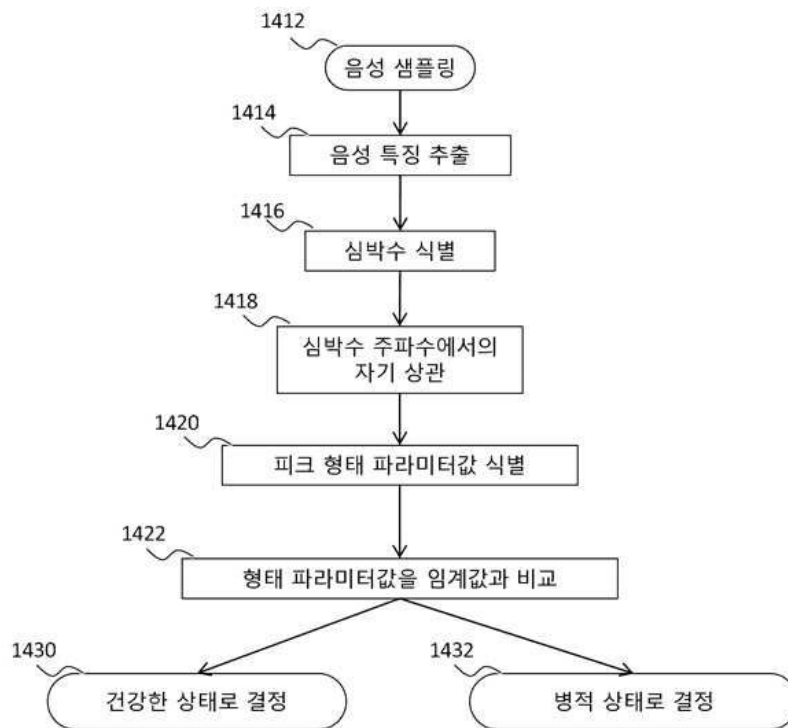
도면13



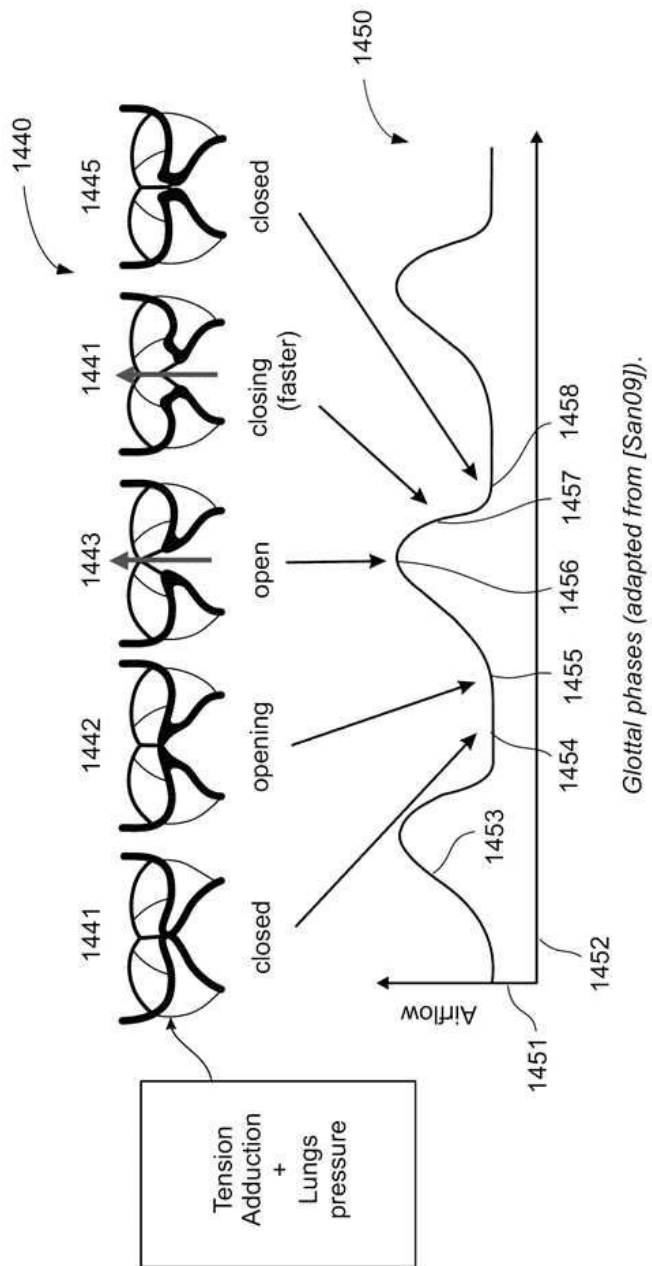
도면14a



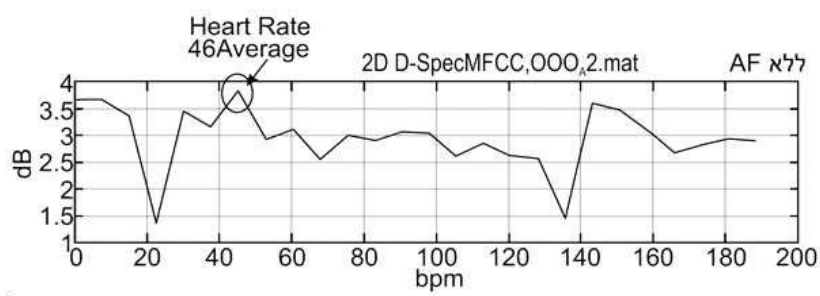
도면14b



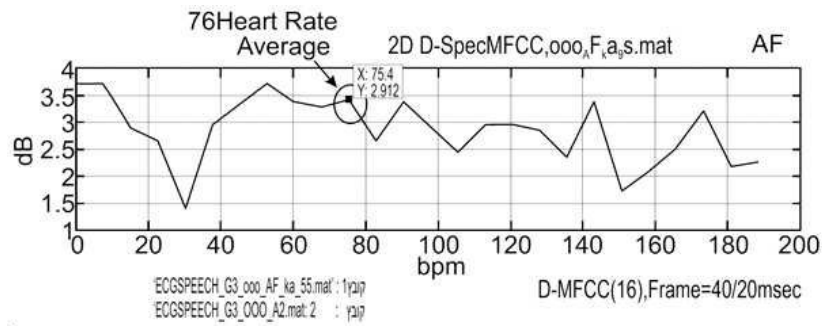
도면14c



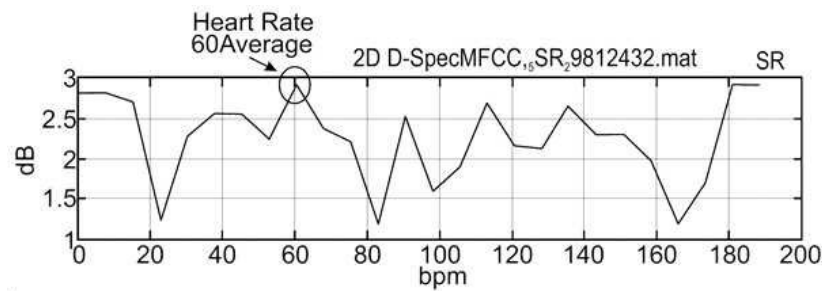
도면15a



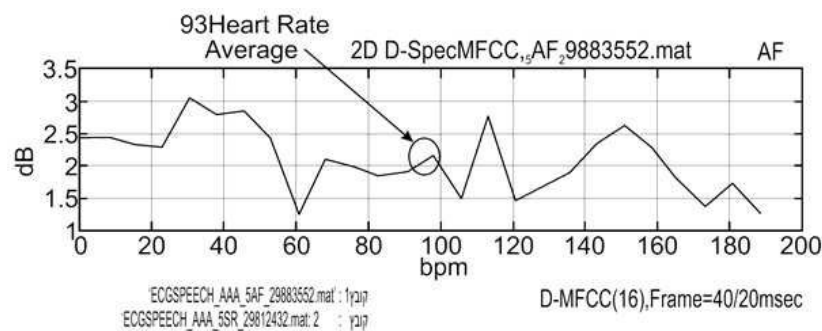
도면15b



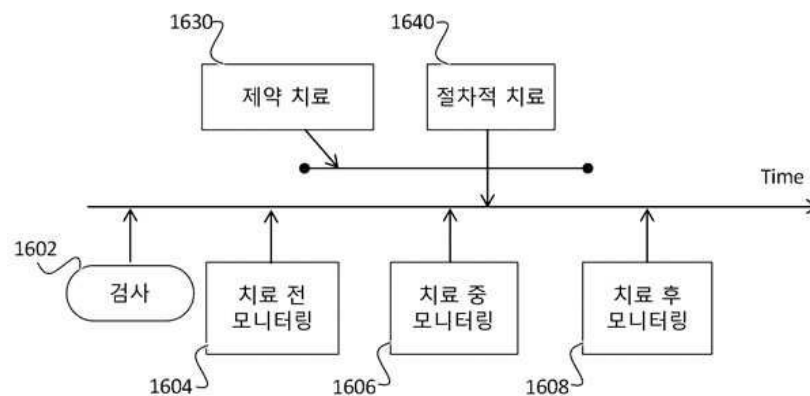
도면15c



도면15d



도면16



专利名称(译)	语言常规筛查心脏病		
公开(公告)号	KR1020190113968A	公开(公告)日	2019-10-08
申请号	KR1020197026858	申请日	2018-02-12
发明人	하우프먼, 일미야후 고렌, 알론 아타르, 엘리 사박, 핀하스		
IPC分类号	A61B5/00 A61B5/02 G10L17/26 G10L25/66		
CPC分类号	A61B5/7275 A61B5/0022 A61B5/02 A61B5/4803 A61B5/7264 G10L17/26 G10L25/66 G10L15/22		
优先权	62/457914 2017-02-12 US		
外部链接	Espacenet		

摘要(译)

公开了一种分析对象的语音样本以确定心脏状况（例如心律不齐）的方法，该方法包括从语音样本中提取至少一个语音特征，针对至少一个语音特征的心脏检测状况的影响，并根据该影响确定心脏状况。另外，公开了相关的装置和方法。

