



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0085731
(43) 공개일자 2013년07월30일

(51) 국제특허분류(Int. Cl.)

G10L 25/93 (2013.01) G10L 25/90 (2013.01)

(21) 출원번호 10-2012-0006750

(22) 출원일자 2012년01월20일

심사청구일자 2012년01월20일

(71) 출원인

세종대학교산학협력단

서울특별시 광진구 능동로 209 (군자동, 세종대학교)

(72) 발명자

김무영

서울특별시 강남구 도곡2동 도곡렉슬아파트 210동 402호

이명석

서울특별시 구로구 오류동 68-21 미래빌라 1동 209호

(74) 대리인

특허법인태백

전체 청구항 수 : 총 10 항

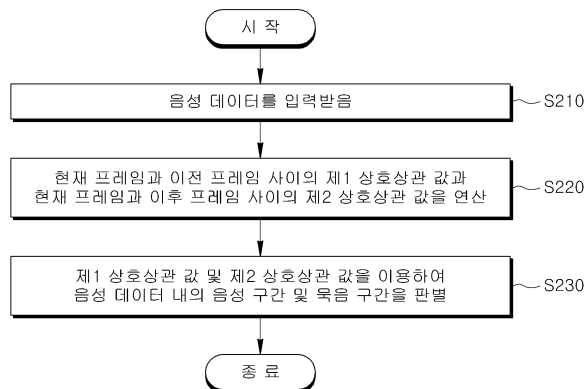
(54) 발명의 명칭 음성 영역 검출 방법 및 그 장치

(57) 요약

본 발명은 음성 영역 검출 방법 및 그 장치에 관한 것이다. 본 발명에 따르면, 음성 영역 검출 장치를 이용한 음성 영역 검출 방법에 있어서, 음성 데이터를 입력받는 단계와, 상기 음성 데이터를 구성하는 복수의 프레임들에 대해, 현재 프레임의 자기상관 값과 이전 프레임의 자기상관 값 사이의 제1 상호상관 값과, 상기 현재 프레임의 자기상관 값과 이후 프레임의 자기상관 값 사이의 제2 상호상관 값을 연산하는 단계, 및 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 이용하여 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별하는 단계를 포함하는 음성 영역 검출 방법을 제공한다.

상기 음성 영역 검출 방법 및 그 장치에 따르면, 음성 데이터에 대한 인접 프레임 간의 상관 관계를 이용함으로써 잡음 섞인 환경에서 음성 영역과 묵음 영역을 정확히 검출할 수 있는 이점이 있다.

대표도 - 도2



특허청구의 범위

청구항 1

음성 영역 검출 장치를 이용한 음성 영역 검출 방법에 있어서,

음성 데이터를 입력받는 단계;

상기 음성 데이터를 구성하는 복수의 프레임들에 대해, 현재 프레임의 자기상관 값과 이전 프레임의 자기상관 값 사이의 제1 상호상관 값과, 상기 현재 프레임의 자기상관 값과 이후 프레임의 자기상관 값 사이의 제2 상호상관 값을 연산하는 단계; 및

상기 제1 상호상관 값 및 상기 제2 상호상관 값을 이용하여 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별하는 단계를 포함하는 음성 영역 검출 방법.

청구항 2

청구항 1에 있어서,

복수의 음성 훈련 데이터들에 대하여 상기 제1 상호상관 값과 상기 제2 상호상관 값을 연산하는 단계; 및

상기 음성 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과, 상기 묵음 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들을 DB부에 저장하는 단계를 더 포함하는 음성 영역 검출 방법.

청구항 3

청구항 2에 있어서,

상기 음성 구간 및 묵음 구간을 판별하는 단계는,

상기 음성 데이터에 대해 연산된 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 상기 DB부 내에 저장된 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과 비교하여, 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별하는 음성 영역 검출 방법.

청구항 4

청구항 1에 있어서,

상기 음성 구간 및 묵음 구간을 판별하는 단계는,

상기 제1 상호상관 값 및 상기 제2 상호상관 값을 서로 합산하는 단계;

상기 합산된 값을 기 설정된 임계치와 비교하는 단계; 및

상기 합산된 값이 상기 임계치 미만이면 상기 묵음 구간으로 판별하고, 상기 임계치 이상이면 상기 음성 구간으로 판별하는 음성 영역 검출 방법.

청구항 5

청구항 1 내지 청구항 4 중 어느 한 항에 있어서,

상기 제1 상호상관 값 $\overline{\xi}_{i-}$ 과 상기 제2 상호상관 값 $\overline{\xi}_{i+}$ 은 아래의 수학식들로 연산되는 음성 영역 검출 방법:

$$\bar{\xi}_{i-} = \frac{\sum_{n=1}^N S_i(n) S_{i-1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i-1}^2(n)}}$$

$$\bar{\xi}_{i+} = \frac{\sum_{n=1}^N S_i(n) S_{i+1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i+1}^2(n)}}$$

여기서, $S_i(n)$ 은 상기 현재 프레임의 자기상관 값, $S_{i-1}(n)$ 은 상기 이전 프레임의 자기상관 값, $S_{i+1}(n)$ 은 상기 이후 프레임의 자기상관 값을 나타낸다.

청구항 6

음성 데이터를 입력받는 데이터 입력부;

상기 음성 데이터를 구성하는 복수의 프레임들에 대해, 현재 프레임의 자기상관 값과 이전 프레임의 자기상관 값 사이의 제1 상호상관 값과, 상기 현재 프레임의 자기상관 값과 이후 프레임의 자기상관 값 사이의 제2 상호상관 값을 연산하는 상관 값 연산부; 및

상기 제1 상호상관 값 및 상기 제2 상호상관 값을 이용하여 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별하는 구간 판별부를 포함하는 음성 영역 검출 장치.

청구항 7

청구항 6에 있어서,

복수의 음성 훈련 데이터들에 대하여 상기 제1 상호상관 값과 상기 제2 상호상관 값을 연산하는 훈련 데이터 연산부; 및

상기 음성 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과, 상기 묵음 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들을 DB부에 저장하는 DB부를 더 포함하는 음성 영역 검출 장치.

청구항 8

청구항 7에 있어서,

상기 구간 판별부는,

상기 음성 데이터에 대해 연산된 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 상기 DB부 내에 저장된 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과 비교하여, 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별하는 음성 영역 검출 장치.

청구항 9

청구항 6에 있어서,

상기 구간 판별부는,

상기 제1 상호상관 값 및 상기 제2 상호상관 값을 서로 합산하고, 상기 합산된 값을 기 설정된 임계치와 비교하여, 상기 합산된 값이 상기 임계치 미만이면 상기 묵음 구간으로 판별하고, 상기 임계치 이상이면 상기 음성 구

간으로 판별하는 음성 영역 검출 장치.

청구항 10

청구항 6 내지 청구항 9 중 어느 한 항에 있어서,

상기 제1 상호상관 값 $\overline{\xi}_{i-}$ 과 상기 제2 상호상관 값 $\overline{\xi}_{i+}$ 은 아래의 수학적식들로 연산되는 음성 영역 검출 장치:

$$\overline{\xi}_{i-} = \frac{\sum_{n=1}^N S_i(n) S_{i-1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i-1}^2(n)}}$$

$$\overline{\xi}_{i+} = \frac{\sum_{n=1}^N S_i(n) S_{i+1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i+1}^2(n)}}$$

여기서, $S_i(n)$ 은 상기 현재 프레임의 자기상관 값, $S_{i-1}(n)$ 은 상기 이전 프레임의 자기상관 값, $S_{i+1}(n)$ 은 상기 이후 프레임의 자기상관 값을 나타낸다.

명세서

기술분야

[0001] 본 발명은 음성 영역 검출 방법 및 그 장치에 관한 것으로서, 보다 상세하게는 인접 프레임 간의 상관 관계를 이용하여 음성 영역과 묵음 영역을 판별할 수 있는 음성 영역 검출 방법 및 그 장치에 관한 것이다.

배경기술

[0002] 기존에 음성 데이터로부터 음성 영역을 검출하는 방법에 관하여 다양하게 개시되어 있다. 그 중에서 특허공개 제1999-0039422호는 G.729 음성 부호화기를 위한 음성 활성화 측정 방법에 관한 것이다. 이는 음성 프레임의 에너지를 문턱값과 비교하여 묵음 구간 및 실 음성 구간으로 구분하는 구성을 포함한다.

[0003] 그런데, 이러한 종래의 방법에 따르면, 단순히 문턱값을 이용하여 음성 영역을 검출하므로 에러가 발생할 수 있으며, 잡음이 섞이면서 특징 벡터들이 크게 변하여 정확한 음성 영역의 검출이 어려운 단점이 있다.

발명의 내용

해결하려는 과제

[0004] 본 발명은 음성 데이터에 대한 인접 프레임 간의 상관 관계를 이용함으로써 잡음 섞인 환경에서 음성 영역과 묵음 영역을 정확히 검출할 수 있는 음성 영역 검출 방법 및 그 장치를 제공하는데 목적이 있다.

과제의 해결 수단

[0005] 본 발명은, 음성 영역 검출 장치를 이용한 음성 영역 검출 방법에 있어서, 음성 데이터를 입력받는 단계와, 상기 음성 데이터를 구성하는 복수의 프레임들에 대해, 현재 프레임의 자기상관 값과 이전 프레임의 자기상관 값 사이의 제1 상호상관 값과, 상기 현재 프레임의 자기상관 값과 이후 프레임의 자기상관 값 사이의 제2 상호상관

값을 연산하는 단계, 및 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 이용하여 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별하는 단계를 포함하는 음성 영역 검출 방법을 제공한다.

[0006] 여기서, 상기 음성 영역 검출 방법의 일 실시예에서는, 복수의 음성 훈련 데이터들에 대하여 상기 제1 상호상관 값과 상기 제2 상호상관 값을 연산하는 단계, 및 상기 음성 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과, 상기 묵음 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들을 DB부에 저장하는 단계를 더 포함할 수 있다. 이러한 일 실시예의 경우, 상기 음성 구간 및 묵음 구간을 판별하는 단계는, 상기 음성 데이터에 대해 연산된 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 상기 DB부 내에 저장된 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과 비교하여, 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별할 수 있다.

[0007] 또한, 상기 음성 영역 검출 방법의 다른 실시예의 경우, 상기 음성 구간 및 묵음 구간을 판별하는 단계는, 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 서로 합산하는 단계와, 상기 합산된 값을 기 설정된 임계치와 비교하는 단계, 및 상기 합산된 값이 상기 임계치 미만이면 상기 묵음 구간으로 판별하고, 상기 임계치 이상이면 상기 음성 구간으로 판별할 수 있다.

[0008] 그리고, 상기 제1 상호상관 값 $\bar{\xi}_{i-}$ 과 상기 제2 상호상관 값 $\bar{\xi}_{i+}$ 은 아래의 수학식들로 연산될 수 있다.

$$\bar{\xi}_{i-} = \frac{\sum_{n=1}^N S_i(n) S_{i-1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i-1}^2(n)}}$$

[0009]

$$\bar{\xi}_{i+} = \frac{\sum_{n=1}^N S_i(n) S_{i+1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i+1}^2(n)}}$$

[0010]

[0011] 여기서, $S_i(n)$ 은 상기 현재 프레임의 자기상관 값, $S_{i-1}(n)$ 은 상기 이전 프레임의 자기상관 값, $S_{i+1}(n)$ 은 상기 이후 프레임의 자기상관 값을 나타낸다.

[0012] 그리고, 본 발명은, 음성 데이터를 획득하는 데이터 입력부와, 상기 음성 데이터를 구성하는 복수의 프레임들에 대해, 현재 프레임의 자기상관 값과 이전 프레임의 자기상관 값 사이의 제1 상호상관 값과, 상기 현재 프레임의 자기상관 값과 이후 프레임의 자기상관 값 사이의 제2 상호상관 값을 연산하는 상관 값 연산부, 및 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 이용하여 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별하는 구간 판별부를 포함하는 음성 영역 검출 장치를 제공한다.

[0013] 여기서, 상기 음성 영역 검출 장치의 일 실시예에서는, 복수의 음성 훈련 데이터들에 대하여 상기 제1 상호상관 값과 상기 제2 상호상관 값을 연산하는 훈련 데이터 연산부, 및 상기 음성 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과, 상기 묵음 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들을 DB부에 저장하는 DB부를 더 포함할 수 있다. 이러한 일 실시예의 경우, 상기 구간 판별부는, 상기 음성 데이터에 대해 연산된 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 상기 DB부 내에 저장된 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과 비교하여, 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별할 수 있다.

[0014] 또한, 상기 음성 영역 검출 장치의 다른 실시예의 경우, 상기 구간 판별부는, 상기 제1 상호상관 값 및 상기 제2 상호상관 값을 서로 합산하고, 상기 합산된 값을 기 설정된 임계치와 비교하여, 상기 합산된 값이 상기 임계치 미만이면 상기 묵음 구간으로 판별하고, 상기 임계치 이상이면 상기 음성 구간으로 판별할 수 있다.

발명의 효과

- [0015] 본 발명에 따른 음성 영역 검출 방법 및 그 장치에 따르면, 음성 데이터에 대한 인접 프레임 간의 상관 관계를 이용함으로써 잡음 섞인 환경에서 음성 영역과 묵음 영역을 정확히 검출할 수 있는 이점이 있다.

도면의 간단한 설명

- [0016] 도 1은 본 발명의 실시예에 따른 음성 영역 검출 장치의 구성도이다.
 도 2는 도 1의 장치를 이용한 음성 영역 검출 방법을 나타내는 흐름도이다.
 도 3은 도 2의 일 실시예를 위한 흐름도이다.
 도 4는 도 2의 다른 실시예를 위한 흐름도이다.
 도 5는 도 4에 대응되는 개념도이다.

발명을 실시하기 위한 구체적인 내용

- [0017] 그러면 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다.
- [0018] 본 발명은 음성 영역 검출 장치 및 방법에 관한 것으로서 배경 잡음에 강인한 음성 영역 검출 알고리즘을 제공한다.
- [0019] 도 1은 본 발명의 실시예에 따른 음성 영역 검출 장치의 구성도이다. 상기 음성 영역 검출 장치(100)는 데이터 입력부(110), 상관 값 연산부(120), 구간 판별부(130), 훈련 데이터 연산부(140), DB부(150)를 포함한다.
- [0020] 도 2는 도 1의 장치를 이용한 음성 영역 검출 방법을 나타내는 흐름도이다. 이하에서는 본 발명에 따른 음성 영역 검출 방법에 관하여 상세히 설명한다.
- [0021] 먼저, 상기 데이터 입력부(110)에서는 음성 영역 검출 대상에 해당되는 테스트용 음성 데이터를 입력받는다(S210). 이러한 데이터 입력부(110)는 음성 수집모듈(미도시)로부터 미리 획득된 음성 데이터를 전송받아 입력받거나, 직접 음성 데이터를 수집하여 입력받을 수 있다. 상기 음성 데이터는 시간 흐름에 따라 복수의 프레임들로 구분될 수 있다.
- [0022] 이후, 상기 상관 값 연산부(120)에서는 상기 획득된 음성 데이터를 구성하는 복수의 프레임들에 대해, 현재 프레임의 자기상관 값과 이전 프레임의 자기상관 값 사이의 제1 상호상관 값, 그리고 상기 현재 프레임의 자기상관 값과 이후 프레임의 자기상관 값 사이의 제2 상호상관 값을 각각 연산한다(S220).
- [0023] 여기서, 자기 상관이란 Auto-Correlation을 의미하고, 상호 상관이란 Cross-Correlation을 의미한다. 자기 상관 함수는 음성 신호의 하모닉(Harmonic; 고조파)을 확인할 수 있는 방법에 해당된다.

- [0024] 이러한 S220 단계에 따른 제1 상호상관 값($\overline{\xi}_{i-}$)과, 제2 상호상관 값($\overline{\xi}_{i+}$)의 연산은 아래의 수학식 1 및 수학식 2를 사용한다.

수학식 1

$$\overline{\xi}_{i-} = \frac{\sum_{n=1}^N S_i(n) S_{i-1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i-1}^2(n)}}$$

[0025]

수학식 2

$$\overline{\xi}_{i+} = \frac{\sum_{n=1}^N S_i(n) S_{i+1}(n)}{\sqrt{\sum S_i^2(n)} \sqrt{\sum S_{i+1}^2(n)}}$$

[0026]

[0027] 이러한 수학식 1 및 수학식 2에서, $S_i(n)$ 는 상기 현재 프레임의 자기상관 값이고, $S_{i-1}(n)$ 은 상기 이전 프레임의 자기상관 값이며, $S_{i+1}(n)$ 은 상기 이후 프레임의 자기상관 값을 나타낸다. 그리고, n 은 시간축의 인덱스 값으로서 1에서 N 사이의 정수 값을 갖는다.

[0028] 상기 수학식 1,2에 의해 구해지는 제1 상호상관 값과, 제2 상호상관 값은 모두 0과 1 사이의 값($0 \leq \overline{\xi}_{i-} \leq 1$, $0 \leq \overline{\xi}_{i+} \leq 1$)을 갖는다. 여기서, 0에 가까울수록 묵음 영역에 해당될 확률이 높고, 1에 가까울수록 음성 영역에 해당될 확률이 높다.

[0029] 이후, 상기 구간 판별부(130)에서는 상기 수학식 1,2를 통해 연산된 제1 상호상관 값($\overline{\xi}_{i-}$)과, 제2 상호상관 값($\overline{\xi}_{i+}$)을 이용하여 상기 음성 데이터 내의 음성 구간 및 묵음 구간을 판별한다(S230).

[0030] 여기서, 음성 구간 및 묵음 구간의 판단의 목적은 음성 데이터로부터 음성 부분만을 추출하여 다양한 용도의 시스템에 사용하기 위한 것이다. 음성의 추출의 성능이 높을수록 시스템의 오류를 줄이고 시스템의 운용 효율 및 신뢰도를 높일 수 있다.

[0031] 이하에서는 상기 음성 구간 및 묵음 구간을 판별하기 위한 서로 다른 두 방법의 실시예를 설명한다.

[0032] 그 일 실시예의 경우는, 별도의 음성 훈련 데이터들로부터 각 구간 별 상호상관 값의 데이터 범위를 학습한 다음, 실제 테스트를 위한 음성 데이터의 프레임별 상호상관 값을 기 학습된 데이터와 대조하여, 음성 데이터 내의 전체 프레임에 대해 해당 프레임이 음성 구간인지 묵음 구간인지 판별하는 방식이다.

[0033] 다른 실시예의 경우는, 상기와 같은 훈련 데이터가 필요 없이, 단순히 수학식 1,2에 의한 상호상관의 합산된 값을 기 설정된 임계치와 비교하여, 음성 데이터 내의 해당 프레임이 음성 구간인지 묵음 구간인지 판별하는 방법이다.

[0034] 도 3은 도 2의 일 실시예를 위한 흐름도이다. 먼저, 도 3을 참조로 하여 상기 일 실시예에 대하여 상세히 설명하면 다음과 같다. 이러한 일 실시예의 경우 상기 음성 영역 검출 장치(100)는 상기 훈련 데이터 연산부(140), DB부(150)의 구성을 필요로 한다.

[0035] 우선, 상기 훈련 데이터 연산부(140)에서는 복수의 음성 훈련 데이터들에 대하여, 상기 제1 상호상관 값과 상기 제2 상호상관 값을 수학식 1,2를 사용하여 연산한다(S310).

[0036] 여기서, 음성 훈련 데이터란 전체 프레임에 대하여 음성 구간 및 묵음 구간에 대한 정보가 미리 확보된 데이터를 의미한다. 훈련 데이터를 1개만 사용해도 되지만, 오류율을 낮추기 위해서는 여러 개의 훈련 데이터를 통해 상호상관 값을 구하여 결과에 반영하는 것이 유리하다.

[0037] 여기서, 임의의 훈련 데이터의 프레임에 대해 상호상관 값이 계산된 결과 예는 표 1을 참조한다. 표 1은 설명의 편의상 4개의 프레임에 대한 결과만 나타낸 것이다.

표 1

제1 상호상관 값	제2 상호상관 값	구간
0.9	0.8	음성
0.1	0.2	묵음
0.7	0.4	묵음
0.7	0.5	음성

[0038]

[0039]

이러한 표 1과 같은 데이터들을 모두 취합하면 음성 구간에 해당되는 개별 상호상관 값들, 묵음 구간에 해당되는 개별 상호상관 값들을 획득할 수 있다. 이러한 모델의 학습에는 통계적 방식이 사용될 수 있는데, 예를 들면 GMM(Gaussian mixture model; 가우시안 혼합 모델) 방식이 적용될 수 있다. 물론, 본 발명이 반드시 이에 한정되는 것은 아니다.

[0040]

이후, 상기 DB부(150)에서는 상기 음성 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들, 그리고 상기 묵음 구간에 해당하는 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들을 저장한다(S320).

[0041]

즉, 사전에 학습을 위해 사용되는 상기 음성 훈련 데이터들은 전체 프레임에 대해 이미 음성 구간과 묵음 구간을 알고 있는 데이터에 해당된다. 이러한 점을 이용하여, 음성 구간에 해당되는 제1 상호상관 값의 범위와 제2 상호상관 값의 범위, 그리고 묵음 구간에 해당되는 제1 상호상관 값의 범위와 제2 상호상관 값의 범위를 미리 획득하여 저장할 수 있다. 이렇게 저장된 정보는 앞서 S230 단계 시에 참고 데이터로 활용된다.

[0042]

이에 따라, 본 발명의 일 실시예에서 상기 음성 구간 및 묵음 구간을 판별할 때(S230)에는, S220 단계 시에 테스트용 음성 데이터에 대해 획득된 값을, S320 단계 시에 훈련 데이터에 대해 저장된 값들과 비교함으로써, 음성 구간 및 묵음 구간을 검출한다.

[0043]

다시 말해서, 일 실시예에서 상기 구간 판별부(130)는 테스트용 음성 데이터 내의 프레임에 대해 연산된 상기 제1 상호상관 값 및 상기 제2 상호상관 값을, 상기 DB부(150) 내에 저장된 훈련 데이터들의 제1 상호상관 값들 및 제2 상호상관 값들과 비교함으로써, 테스트용 음성 데이터 내의 프레임들에 대하여 음성 구간 및 묵음 구간을 판별 및 구분할 수 있게 된다.

[0044]

이와는 달리, 상기 다른 실시의 경우, 상기 음성 영역 검출 장치(100)는 상기 훈련 데이터 연산부(140), DB부(150)의 구성이 전혀 필요 없이 음성 구간 및 묵음 구간을 판별할 수 있다.

[0045]

도 4는 도 2의 다른 실시예를 위한 흐름도이다. 도 5는 도 4에 대응되는 개념도이다. 도 4 및 도 5를 참조로 하여 상기 일 실시예에 대하여 상세히 설명하면 다음과 같다.

[0046]

먼저, 상기 구간 판별부(130)는 앞서 S220 단계에서 구하여진 제1 상호상관 값 ($\overline{\xi}_{i-}$)과 상기 제2 상호상관 값 ($\overline{\xi}_{i+}$)을 수학적 식 3과 같이 서로 합산한다(S410).

수학적 식 3

$$\overline{\xi}_i = \overline{\xi}_{i-} + \overline{\xi}_{i+}$$

[0047]

[0048]

이렇게 합산된 값은 0과 2 사이의 값($0 \leq \overline{\xi}_i \leq 2$)을 갖는다. 이후, 상기 합산된 값($\overline{\xi}_i$)을 0과 1 사이의 값으로 정규화한다(S420). 이러한 정규화 과정은 이후 임계치와의 비교 시에 계산의 용이성 및 정확도를 위한 과정이다.

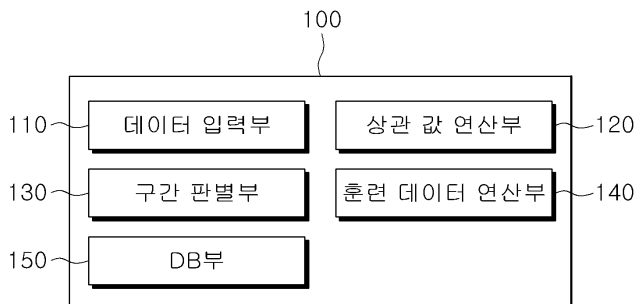
- [0049] 다음, 상기 정규화된 값을 기 설정된 임계치(Threshold)와 비교한다(S430). 물론, 상기 임계치는 0과 1 사이의 임의의 값을 갖는다.
- [0050] 만약, 상기 정규화된 값이 임계치 이상이면 해당 프레임을 상기 음성 구간으로 판별한다(S440). 또한, 상기 정규화된 값이 임계치 미만이면 상기 묵음 구간으로 판별한다(S450).
- [0051] 이러한 다른 실시예의 경우, 훈련 데이터를 학습시키는 과정이 전혀 불필요하므로 연산 과정이 간소화되고 연산 속도가 빨라지는 이점이 있다.
- [0052] 이상과 같은 본 발명의 실시예에 따른 음성 영역 검출 방법 및 장치에 따르면, 음성 데이터에 대한 인접 프레임 간의 상관 관계를 이용함으로써 잡음 섞인 환경에서 음성 영역과 묵음 영역을 정확히 검출할 수 있는 이점이 있다.
- [0053] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 다른 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의하여 정해져야 할 것이다.

부호의 설명

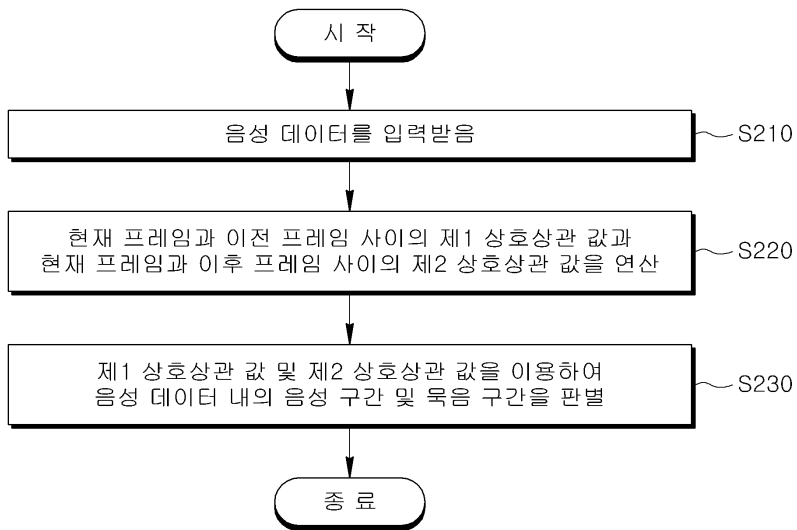
- [0054] 100: 음성 영역 검출 장치 110: 데이터 입력부
120: 상관 값 연산부 130: 구간 판별부
140: 훈련 데이터 연산부 150: DB부

도면

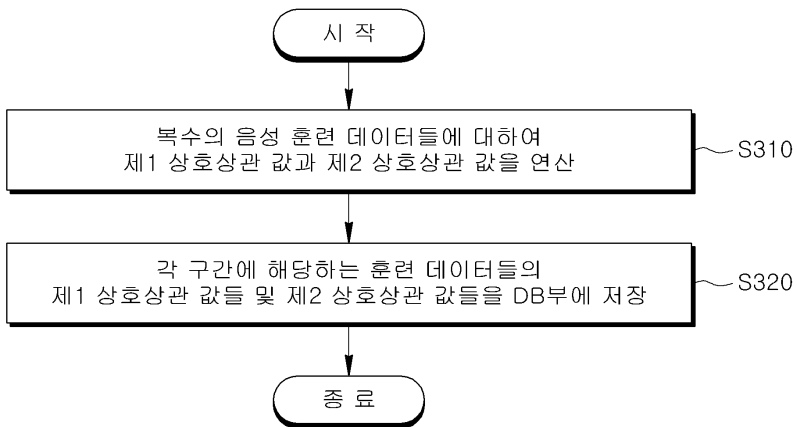
도면1



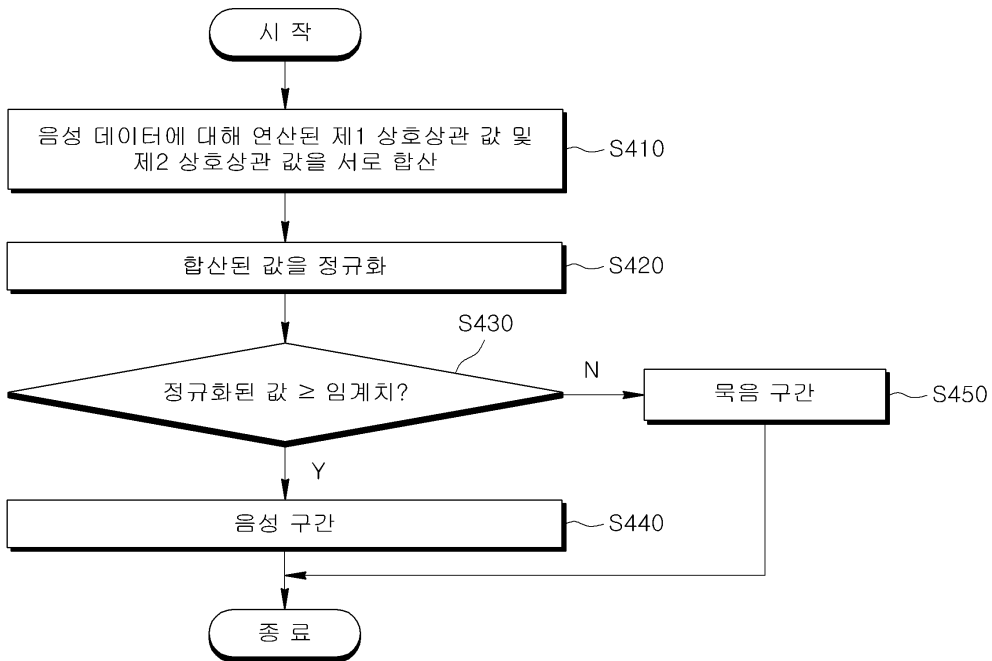
도면2



도면3



도면4



도면5

