(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) 。Int. Cl.
G10L 21/02 (2006.01)

(11) 공개번호 10-2007-0050699
(43) 공개일자 2007년05월16일

(21) 출원번호 10-2005-0108236
(22) 출원일자 2005년11월11일
심사청구일자 2005년11월11일

(71) 출원인 고려대학교 산학협력단
서울 성북구 안암동5가1 고려대학교 내

(72) 발명자 고한석
서울 용산구 서빙고동 신동아아파트 16-703
이윤재
부산 부산진구 개금2동 현대아파트 204동 1702호

(74) 대리인 현종철

전체 청구항 수 : 총 10 항

(54) 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법및 장치

(57) 요약

잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법 및 장치가 개시된다. 그 에너지 특징 보상 방법은 (a) 잡음이 없는 깨끗한 환경에서 수집한 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로 변환시키는 단계; (b) 인식 음성데이터의 에너지 최소값이 소정의 목표 최소값보다 큰지 체크하여, 상기 음성데이터의 에너지 최소값이 상기 목표 최소값보다 작으면, 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN을 수행하는 단계; 및 (c) 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면 에너지 특징을 소정의 방법으로 보상하는 단계를 포함함을 특징으로 한다.

본 발명에 의하면, 보다 다양한 환경에 더욱 강인한 에너지 특징 보상이 가능하게 되어 잡음 환경에서의 음성 인식률을 향상시킬 수 있다.

대표도

도 2

특허청구의 범위

청구항 1.

(a) 잡음이 없는 깨끗한 환경에서 수집한 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로 변환시키는 단계;

(b) 인식 음성데이터의 에너지 최소값이 소정의 목표 최소값보다 큰지 체크하여, 상기 음성데이터의 에너지 최소값이 상기 목표 최소값보다 작으면, 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN을 수행하는 단계; 및

(c) 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면 에너지 특징을 소정의 방법으로 보상하는 단계를 포함함을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법.

청구항 2.

제1항에 있어서, 상기 (c)단계는

(c1) 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면, 상기 최소 에너지가 소정의 문턱치보다 큰지 체크하는 단계; 및

(c2) 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크면 에너지를 차감하고, 그렇지 않으면 인식 데이터에서 문턱치 에너지 이하의 값에서 ERN 역함수를 이용하여 에너지를 변환하는 단계를 구비함을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법.

청구항 3.

제2항에 있어서, 상기 (a)단계의 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로의 변환은

에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half ERN 방법을 이용하여 수행함을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법.

청구항 4.

제3항에 있어서, 상기 문턱치는

[수학식 3]

$$Th = \alpha * Noisy_Min + (1-\alpha) * Noisy_Max$$

(여기서, 알파는 0에서 1사이의 임의의 값이며, Noisy_Min은 잡음의 최소값, Noisy_Max는 잡음의 최대값을 나타낸다.)

수학식 3에 의해 결정됨을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법.

청구항 5.

제3항에 있어서, 상기 에너지 차감은

음성이 입력될 때 첫 10개의 프레임은 은성이 없다고 가정할 때, 상기 인식 데이터의 첫 10 프레임을 이용하여 로그를 취하지 않은 잡음 에너지를 추정하여 에너지를 차감해준 후 다시 로그 함수를 취함을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법.

청구항 6.

잡음이 없는 깨끗한 환경에서 수집한 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로 변환시키는 에너지변환부;

인식 음성데이터의 에너지 최소값이 소정의 목표 최소값보다 큰지 체크하는 목표 최소값 비교부;

상기 목표 최소값 비교부에서 상기 음성데이터의 에너지 최소값이 목표 최소값보다 작으면, 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN을 수행하는 Half-ERN 변환부;

상기 목표 최소값 비교부에서 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면, 상기 최소 에너지가 소정의 문턱치보다 큰지 체크하는 문턱값 비교부;

상기 문턱값 비교부에서 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크면 에너지를 차감하는 에너지차감부; 및

상기 문턱값 비교부에서 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크지 않으면 인식 데이터에서 문턱치 에너지 이하의 값에서 ERN 역함수를 이용하여 에너지를 변환하는 역변환 ERN부를 더 포함함을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 장치.

청구항 7.

제6항에 있어서, 상기 에너지 변환부에서의 잡음환경 에너지와 유사한 환경으로의 변환은

에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half ERN 방법을 이용하여 수행함을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 장치.

청구항 8.

제7항에 있어서, 상기 문턱치는

[수학식 3]

$$Th = \alpha * Noisy_Min + (1-\alpha) * Noisy_Max$$

(여기서, 알파는 0에서 1사이의 임의의 값이며, Noisy_Min은 잡음의 최소값, Noisy_Max는 잡음의 최대값을 나타낸다.)

수학식 3에 의해 결정됨을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 장치.

청구항 9.

제7항에 있어서, 상기 에너지 차감은

음성이 입력될 때 첫 10개의 프레임은 음성이 없다고 가정할 때, 상기 인식 데이터의 첫 10 프레임을 이용하여 로그를 취하지 않은 잡음 에너지를 추정하여 에너지를 차감해준 후 다시 로그 함수를 취함을 특징으로 하는 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 장치.

청구항 10.

제1항 내지 제5항 중 어느 한 항에 기재된 발명을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 음성인식에 관한 것으로서, 특히 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법 및 장치에 관한 것이다.

음성인식 분야에서 전통적인 에너지 정규화 방식은 음성신호의 각 프레임 중 최대의 값을 가지는 프레임의 로그 에너지를 모든 프레임에 빼주어 최대 에너지 값이 0이 되게 만들어주는 방법이다. 이 방법은 Hidden Markov Model Toolkit(HTK)에서 지원하는 방법이다. 하지만 이 방법은 잡음환경에 강인하지 못한 단점을 가진다. 최근에 나온 방법들 중에는 에너지 특징 또한 켑스트럼 특징과 같이 취급하여 평균과 분산 정규화 방법을 적용한 정규화 방법과 ERN(log-energy dynamic range normalization) 방법이 있다.

상기 ERN 방법은 깨끗한 환경의 에너지가 잡음에 의해 어떻게 변하는지 분석한 후, 깨끗한 환경의 에너지 특징을 잡음환경의 특징과 유사하게 변환시켜 주는 방법이며, 다양한 dynamic range(D.R)를 훈련데이터에 적용하여 음향 모델을 만든 후 인식데이터에도 D.R을 적용하여 최적의 인식 결과를 실험적으로 찾는 방식이다. 하지만 이 방법에서 사용하고 있는 음성 에너지의 최대값은 잡음에 영향을 받지 않는다는 가정이 높은 신호 대 잡음비 (signal to noise ratio, SNR)에만 적용이 되는 가정이므로 낮은 SNR 환경에서는 음성구간의 높은 에너지 차이를 줄여주지 못하게 되어 더 높은 음성 인식률을 기대할 수 없게 된다. 또한 실험적으로 찾은 D.R은 평균 인식률이 가장 높은 D.R 이지만 다양한 SNR에서 최고의 성능을 보이는 D.R 이 아니다. 그 이유는 높은 에너지 구간에서의 불일치뿐만 아니라 고정된 D.R 에 의해 낮은 구간의 에너지를 높은 에너지로 변환시켜도 실제 다양한 SNR 에 따라서 낮은 에너지 구간에서의 불일치가 생기기 때문이다. 실제로 높은 SNR과 낮은 SNR 에서 최적의 성능을 내는 D.R 이 다르게 되어 다양한 환경에 강인하지 못한 단점을 가진다.

발명이 이루고자 하는 기술적 과제

본 발명이 이루고자하는 기술적 과제는 인식데이터의 에너지 중 음성 인식에 중요한 상대적으로 높은 에너지 구간에서는 에너지 차감법을 이용하여 깨끗한 환경의 음성 에너지로 일관성 있게 변환 시켜 주기 위해, 깨끗한 환경의 훈련 데이터의 에너지 중 음성에 의한 높은 에너지들은 ERN 변환을 하지 않고 낮은 에너지 구간만 ERN 변환을 함으로써, 다양한 환경 및 다양한 신호 대 잡음비에 강인한 에너지 보상이 가능하게 하는, 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법 및 장치를 제공하는 것이다.

발명의 구성

상기 기술적 과제를 해결하기 위한 본 발명에 의한 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법은, (a) 잡음이 없는 깨끗한 환경에서 수집한 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로 변환시키는 단계; (b) 인식 음성데이터의 에너지 최소값이 소정의 목표 최소값보다 큰지 체크하여, 상기 음성데이터의 에너지 최소값이 상기 목표 최소값보다 작으면, 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN을 수행하는 단계; 및 (c) 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면 에너지 특징을 소정의 방법으로 보상하는 단계를 포함함을 특징으로 한다. 상기 (c)단계는 (c1) 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면, 상기 최소 에너지가 소정의 문턱치보다 큰지 체크하는 단계; 및 (c2) 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크면 에너지를 차감하고, 그렇지 않으면 인식 데이터에서 문턱치 에너지 이하의 값에서 ERN 역함수를 이용하여 에너지를 변환하는 단계를 구비함이 바람직하다. 상기 (a)단계의 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로의 변환은 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half ERN 방

법을 이용하여 수행함이 바람직하다. 상기 에너지 차감은 음성이 입력될 때 첫 10개의 프레임은 음성이 없다고 가정할 때, 상기 인식 데이터의 첫 10 프레임을 이용하여 로그를 취하지 않은 잡음 에너지를 추정하여 에너지를 차감해준 후 다시 로그 함수를 취함이 바람직하다.

상기 기술적 과제를 해결하기 위한 본 발명에 의한 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 장치는, 잡음이 없는 깨끗한 환경에서 수집한 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로 변환시키는 에너지 변환부; 인식 음성데이터의 에너지 최소값이 소정의 목표 최소값보다 큰지 체크하는 목표 최소값 비교부; 상기 목표 최소값 비교부에서 상기 음성데이터의 에너지 최소값이 목표 최소값보다 작으면, 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN을 수행하는 Half-ERN 변환부; 상기 목표 최소값 비교부에서 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면, 상기 최소 에너지가 소정의 문턱치보다 큰지 체크하는 문턱값 비교부; 상기 문턱값 비교부에서 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크면 에너지를 차감하는 에너지차감부; 및 상기 문턱값 비교부에서 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크지 않으면 인식 데이터에서 문턱치 에너지 이하의 값에서 ERN 역함수를 이용하여 에너지를 변환하는 역변환 ERN부를 더 포함함을 특징으로 한다.

상기 에너지 변환부에서의 잡음환경 에너지와 유사한 환경으로의 변환은 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN 방법을 이용하여 수행함이 바람직하다. 상기 에너지 차감은 음성이 입력될 때 첫 10개의 프레임은 음성이 없다고 가정할 때, 상기 인식 데이터의 첫 10 프레임을 이용하여 로그를 취하지 않은 잡음 에너지를 추정하여 에너지를 차감해준 후 다시 로그 함수를 취함이 바람직하다.

그리고 상기 기재된 발명을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.

이하, 첨부된 도면들을 참조하여 본 발명에 따른 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법 및 장치에 대해 상세히 설명한다.

도 1은 본 발명에 따른 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 장치의 구성을 블록도로 도시한 것으로서, 에너지 변환부(100), 목표 최소값 비교부(120), Half-ERN변환부(140), 문턱값 비교부(160), 에너지 차감부(180) 및 역변환 ERN부(190)를 포함하여 이루어진다.

상기 에너지 변환부(100)는 잡음이 없는 깨끗한 환경에서 수집한 음성 훈련데이터의 에너지 특징을 잡음환경의 에너지와 유사한 환경으로 변환시킨다. 상기 에너지 변환부(100)에서의 잡음환경 에너지와 유사한 환경으로의 변환은 바람직하게는 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN 방법을 이용하여 수행한다.

상기 목표 최소값 비교부(120)는 인식 음성데이터의 에너지 최소값이 소정의 목표 최소값보다 큰지 체크한다. Half-ERN 변환부(140)는 상기 목표 최소값 비교부(120)에서 상기 음성데이터의 에너지 최소값이 목표 최소값보다 작으면, 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키는 Half-ERN을 수행한다.

상기 문턱값 비교부(160)는 상기 목표 최소값 비교부(120)에서 인식 음성데이터의 에너지 최소값이 목표 최소값보다 작지 않으면, 상기 최소 에너지가 소정의 문턱치보다 큰지 체크한다.

상기 에너지 차감부(180)는 상기 문턱값 비교부(160)에서 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크면 에너지를 차감한다. 상기 에너지 차감은 음성이 입력될 때 첫 10개의 프레임은 음성이 없다고 가정할 때, 상기 인식 데이터의 첫 10 프레임을 이용하여 로그를 취하지 않은 잡음 에너지를 추정하여 에너지를 차감해준 후 다시 로그 함수를 취함이 바람직하다. 상기 역변환 ERN부(190)는 상기 문턱값 비교부(160)에서 인식 데이터의 에너지 최소값이 소정의 문턱치 에너지보다 크지 않으면 인식 데이터에서 문턱치 에너지 이하의 값에서 ERN 역함수를 이용하여 에너지를 변환한다.

도 2는 본 발명에 따른 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법을 흐름도로 도시한 것으로서, 도 2를 참조하여 본 발명에 의한 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법을 설명하기로 한다.

도 3은 잡음이 없는 깨끗한 환경과 10dB 자동차 잡음 환경에 의한 시간에 따른 음성 에너지의 변화를 도시한 것이다. 도 3을 참조하면, 깨끗한 환경의 에너지는 에너지가 낮은 구간에서 그 변화가 매우 크다는 것을 볼 수 있다. 이와 같은 사실을 기반으로 기존의 ERN 방법에서는 훈련데이터에 대해서는 먼저 깨끗한 환경의 훈련데이터의 에너지 특징을 ERN 변환식을 사용하여 잡음환경의 에너지로 변환하여 추정한다.

$$new_log_energy_i = log_energy_i + \frac{T_Min - Min}{Max - Min} * (Max - log_energy_i)$$

여기서, Min 및 Max는 입력 음성 데이터의 log 에너지의 최소값 및 최대값을 나타내고, log_energy_i는 변환전 음성데이터의 log 에너지, $new_log_energy_i$는 변환후의 음성데이터의 log 에너지를 나타낸다.

도 4는 수학식 1을 도식적으로 그린 것이다. 도 3에 도시된 바와 같이 깨끗한 환경의 에너지 최대값은 잡음에 영향을 받지 않는다고 가정하면 에너지 최대값을 고정시킨 채로 낮은 에너지 구간에 보다 많은 비중을 두어 변환시킨다.

그러나 본 발명에서는 종래의 ERN 과는 달리 에너지의 최소값과 최대값의 중간값 이하만 ERN 변환식에 의해 변환시키며 이를 Half-ERN 이라 칭한다.

음성에너지의 최대값이 잡음에 영향을 덜 받는 것은 로그 함수의 특성 때문인데, 낮은 신호 대 잡음비 환경에서는 음성 에너지의 최대값도 잡음에 의해 변화하게 된다. 따라서 기존의 ERN 알고리즘으로는 이러한 문제점을 해결 할 수가 없다. 그래서 본 발명에서는 훈련데이터의 상대적으로 높은 에너지들은 변환시키지 않고 그대로 유지시켜 깨끗한 환경의 에너지 상태로 두게 된다. 그리고 인식 데이터로부터 잡음 에너지를 추정하여 높은 에너지 구간에서 차감해주기 때문에 인식 환경의 SNR에 상관없이 상대적으로 잡음 환경의 높은 에너지들을 신뢰성 있게 깨끗한 환경의 에너지로 변환시킬 수 있게 된다.

목표 최소값 T_Min 은 수학식 2와 같이 고정된 Dynamic Range(D.R)에 의해 결정되는 값으로 에너지의 최소값이 T_Min 값까지 올라가게 되는 선형변환이다.

$$D.R(dB) = 10 * \frac{Max(\log - energy_i)}{T_Min}$$

인식데이터에 대해서는 다음과 같은 과정으로 에너지 특징을 보상해 준다. 먼저, 음성이 입력되면(200단계), 에너지 최소값과 목표 최소값을 비교하여(210단계), 에너지 최소값이 목표 최소값 보다 낮으면 훈련데이터와 같은 Half-ERN 을 적용한다(220단계). 기존의 ERN 에서는 인식 데이터의 최소 로그 에너지가 목표 최소값보다 높으면 어떤 과정도 수행하지 않아 낮은 구간에서 여전히 발생하는 불일치를 해결하지 못했지만 본 발명에서는 낮은 에너지 구간에서의 불일치를 더욱 줄이기 위해 다음과 같은 과정을 수행한다.

만약 인식 데이터의 에너지 최소값이 목표 최소값보다 높으면(210단계), 수학식 3에 의해 결정되는 문턱치에 따라 다른 알고리즘을 적용한다.

$$Th = \alpha * Noisy_Min + (1 - \alpha) * Noisy_Max$$

여기서, 알파는 0에서 1사이의 임의의 값이며, Noisy_Min 및 Noisy_Max은 잡음환경에서 테스트 데이터의 log 에너지의 최소값 및 최대값을 나타낸다. Half-ERN 에서는 알파가 0.5였으나 잡음환경에서는 0.75로 둔다. 그 이유는 잡음환경에 의해 낮은 에너지가 변화가 크기 때문에 문턱치의 비율을 최소 에너지 값에 더 비중을 둘 필요가 있기 때문이다.

먼저 앞에서 언급하였듯이, 최소에너지와 문턱치를 비교하여(230단계), 문턱치 이상의 값에 대해서는 에너지 차감법을 적용한다(240단계). 로그 함수가 적용되기 전의 에너지는 깨끗한 환경의 에너지와 잡음에너지가 부가적인 관계이므로 음성이 없다고 가정되는 처음 10 프레임의 평균 에너지를 잡음에너지로 추정하여 차감해준다. 이 에너지 차감법을 통해 상대적으로 높은 에너지들은 잡음 환경과 SNR에 상관없이 일관성있게 깨끗한 환경의 에너지로 변환되는 것이다. 에너지 차감법 적용시 차감해준 값이 너무 작거나 음수가 되는 것을 방지하기 위해 로그 에너지 값이 5가 되도록 제한을 둔다. 이 값은 실험적으로 깨끗한 환경의 묵음구간의 최소 에너지를 구한 값이다.

상기 230단계에서 비교결과, 최소에너지가 문턱치보다 크지 않으면 즉 문턱치 이하의 값에 대해서는 최소 에너지가 목표 최소값 보다 높으므로 수학식 4와 같은 ERN 변환 함수의 역변환식을 이용하여 낮은 구간의 에너지들을 더 낮은 에너지로 변환시켜 최소 에너지가 목표 최소값까지 내려오도록 하게 한다.(250단계)

$$estimated_e_i = \frac{noisy_e_i - K * Max}{(1-K)} \quad , \quad K = \frac{Min - T_Min}{Max - T_Min}$$

estimated_e_i는 새로 추정된 에너지를 나타내고, noisy_e_i는 잡음환경에서 변환되기 전의 현재 log 에너지를 나타낸다.

도 5는 기존의 ERN 방법과 본 발명에 의한 방법의 결과를 비교한 것을 도시한 것이다. 도 5에서 굵은 선이 깨끗한 환경의 에너지 특징이며, 점선이 에너지 특징이 보상된 것이다.

기존의 방법인 도 5의 위 그래프에서는 비음성 구간인 낮은 에너지 구간에서 여전히 불일치가 일어나며 에너지가 큰 음성 구간에서도 불일치가 보이는 것을 확인 할 수 있다.

반면, 본 발명에 의한 방법을 나타내는 도 5의 아래 그래프에서, 음성 구간에 속하는 상대적으로 높은 에너지들은 훈련데이터 및 인식데이터 모두 깨끗한 환경의 에너지가 되어 유사하게 되고, 상대적으로 낮은 에너지 구간에서는 도 5의 위 그래프의 점선에 해당하는 에너지를 아래쪽으로 끌어내림으로써 깨끗한 환경의 훈련 데이터의 에너지와 잡음환경의 인식 데이터의 에너지간의 불일치가 더욱 줄어드는 것을 확인할 수 있다.

[표 1]

	clean	20dB	15dB	10dB	5dB	0dB	-5dB	평균
Baseline	98.96	97.41	90.04	67.01	34.09	14.46	9.39	60.60
ERN	98.90	96.72	94.27	85.54	59.23	27.02	10.71	72.56
개발된방법	98.96	97.67	95.44	88.1	68.51	34.48	11.01	76.84

또한 표 1의 실험결과에서 보듯이 실제로 음성 인식에서도 본 발명에 의한 방법이 기존의 ERN 방법보다 다양한 SNR에 보다 강인하고 전체 평균 인식률도 보다 향상된 것을 확인할 수 있다.

본 발명은 컴퓨터로 읽을 수 있는 기록 매체에 컴퓨터(정보 처리 기능을 갖는 장치를 모두 포함한다)가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록 장치의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등이 있다.

본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의해 정해져야 할 것이다.

발명의 효과

본 발명에 따른 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법 및 장치에 의하면, 보다 다양한 환경에 더욱 강인한 에너지 특징 보상이 가능하게 되어 잡음 환경에서의 음성 인식률을 향상시킬 수 있다.

본 발명에 의하면, 훈련데이터의 음성 부분에 해당하는 높은 에너지를 그대로 두어 인식 데이터가 입력으로 들어오면 인식 데이터의 에너지 중 음성 인식에 중요한, 상대적으로 높은 에너지 구간에서는 에너지 차감법을 이용하여 깨끗한 환경의 음성 에너지로 일관성 있게 변환시킨다. 그 결과 다양한 환경 및 다양한 신호 대 잡음비에 강인한 에너지 보상이 가능하다. 또한 낮은 에너지 구간에서는 ERN 변환 후에도 여전히 발생하는 훈련 에너지 특징과 인식 에너지 특징을 불일치를 줄이며, 역변환 ERN 을 통해 불일치를 더욱 줄여 잡음에 강인한 음성인식을 구현한다.

도면의 간단한 설명

도 1은 본 발명에 따른 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 장치의 구성을 블록도로 도시한 것이다.

도 2는 본 발명에 따른 잡음환경에 강인한 음성인식을 위한 에너지 특징 보상 방법을 흐름도로 도시한 것이다.

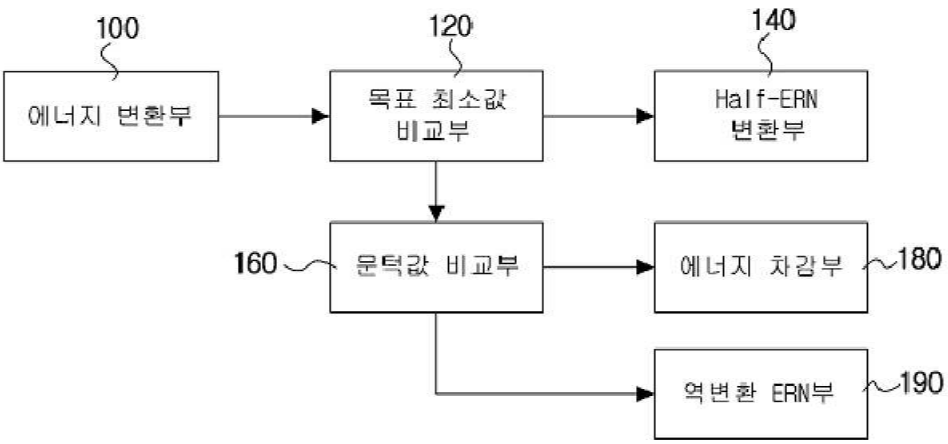
도 3은 잡음이 없는 깨끗한 환경과 10dB 자동차 잡음 환경에 의한 시간에 따른 음성 에너지의 변화를 도시한 것이다.

도 4는 수학식 1을 도식적으로 그린 것이다.

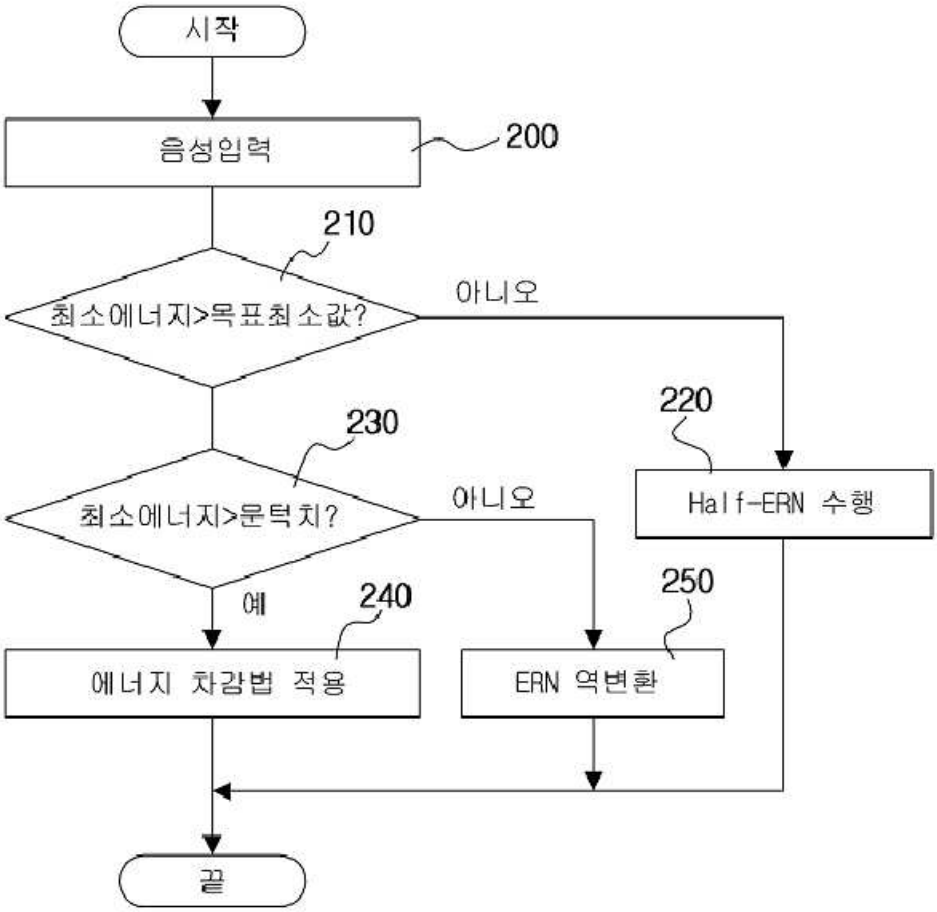
도 5는 기존의 ERN 방법과 본 발명에 의한 방법의 결과를 비교한 것을 도시한 것이다.

도면

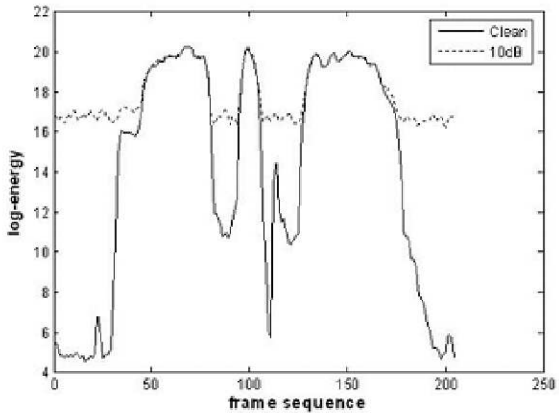
도면1



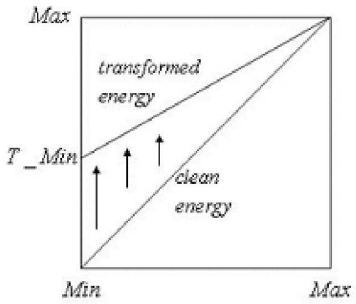
도면2


시작
음성입력
200
210
최소에너지>목표최소값?
아니오
220
Half-ERN 수행
230
최소에너지>문턱치?
아니오
예
240
에너지 차감법 적용
250
ERN 역변환
끝

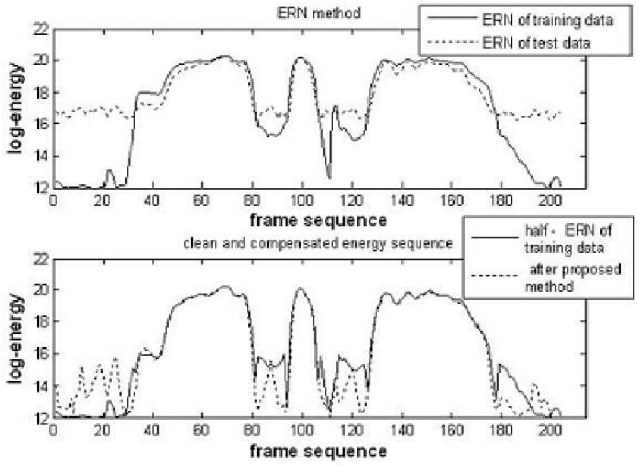
도면3


Clean
10dB
log-energy
frame sequence
22
20
18
16
14
12
10
8
6
4
0
50
100
150
200
250

도면4


Max
transformed
energy
T _Min
clean
energy
Min
Max

도면5


ERN method
ERN of training data
ERN of test data
log-energy
frame sequence
clean and compensated energy sequence
half - ERN of training data
after proposed method
log-energy
frame sequence