

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl. ⁶ G10L 7/08	(11) 공개번호 특 1999-0038741	(43) 공개일자 1999년 06월 05일
(21) 출원번호 10-1997-0058578	(22) 출원일자 1997년 11월 06일	
(71) 출원인 한국전자통신연구원 정선중	대전광역시 유성구 가정동 161번지 안영목	
(72) 발명자	대전광역시 유성구 송강동 송강그린아파트 304-202 김회린	
(74) 대리인	대전광역시 유성구 전민동 청구나라아파트 105-1001 박해천, 원석희	
심사청구 : 있음		

(54) 연속 피치 정보를 이용한 음성 검출 방법

요약

1. 청구범위에 기재된 발명이 속한 기술분야
본 발명은 연속 피치 정보를 이용한 음성 검출 방법에 관한 것임.
2. 발명이 해결하려고 하는 기술적 과제
본 발명은, 외부로부터 입력된 신호에 대하여 에너지와 영교차율을 이용하여 음성인지 여부를 판단한 후에 음성 신호라고 판단되면 더 나아가서 연속적인 피치 정보 중 피치의 존재 범위 영역, 인접피치 사이의 주파수 변화량과 피치 발생수를 이용하여 음성을 검출하는 음성 검출 방법을 제공하고자 함.
3. 발명의 해결방법의 요지
본 발명은, 외부로부터 입력되는 음성 신호의 피치 정보를 검출하는 단계, 검출된 피치가 소정의 기준 피치 영역에 존재하는지를 확인하는 단계, 주파수 변화량을 계산한 후에 계산된 주파수 변화량이 소정의 제 1 기준값보다 작은지를 확인하는 단계 및 각 확인 결과에 따라 음성을 검출하는 단계를 포함한다.
4. 발명의 중요한 용도
본 발명은 음성 인식 시스템에 이용됨.

대표도

도 3

명세서

도면의 간단한 설명

도 1 은 종래의 음성 검출 방법에 대한 흐름도.
도 2 는 본 발명이 적용되는 음성 인식 시스템의 일실시에 구성도.
도 3 은 본 발명에 따른 연속 피치 정보를 이용한 음성 검출 방법에 대한 일실시예를 나타내는 흐름도.

* 도면의 주요 부분에 대한 부호의 설명

- 21: 음성 입력 장치 22: 아날로그/디지털 변환 장치
- 23: 기억 장치 24: 중앙 처리 장치
- 25: 인식 결과 출력 장치

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 피치 정보를 이용한 음성 검출 방법에 관한 것으로서, 특히 음성 신호가 지닌 연속된 피치의 특성을 이용하여 음성 검출 장치가 비음성과 사용자의 음성을 구별하도록 한 음성 검출 방법에 관한 것이다.

본 발명에 사용되는 용어를 정의하면 다음과 같다.

피치(Pitch)는 음향 감각의 하나의 속성으로, 주로 음향 자극의 주파수에 관계하나, 음압이나 파형도 다소 관계한다.

비음성은 음성이 아닌 소리로서 전화벨 소리, 문닫는 소리, 책상 두드리는 소리, 마른 기침 소리 및 박수 소리 등을 말한다.

음성 인식 기술이란 사용자가 발성한 음성을 분석하여 그 발성 내용이 무엇인지를 알아내는 기술이다. 기존의 음성 인식 시스템은 음성 검출부, 음성 특징 추출부, 기준 패턴 비교부, 인식 결과 검증부 및 인식 결과 출력부로 구성되어 있다. 음성 인식 시스템에서 음성 검출부는 사용자의 음성을 검출해 내는 일을 담당하고 있다.

기존의 음성 인식 시스템은 음성 검출부에서 음성이 검출되면 음성 특징 추출 과정, 기준 패턴 비교 과정, 인식 결과 검증 과정 및 인식 결과 출력 과정을 수행한다. 따라서, 음성 검출부에서 사용자의 음성 신호가 아닌 소리를 음성으로 잘못 검출할 경우에 전술한 불필요한 여러 과정을 수행하게 되며, 음성 인식 시스템은 사용자가 원하지 않는 동작을 하게 된다.

기존의 음성 검출 방법은 입력 신호의 에너지 및 영교차율(Zero Crossing Ratio)을 이용하여 사용자의 음성을 검출한다. 에너지 및 영교차율을 이용한 기존의 음성 검출 방법은 사용자의 음성뿐만 아니라 주변의 잡음, 예를 들면 전화벨 소리, 책상 두드리는 소리, 문닫는 소리, 마른 기침 소리 및 박수 소리 등을 음성 신호로 잘못 검출하는 경우가 많다.

이것은 주변 잡음의 에너지 및 영교차율이 음성 검출부에서 미리 정해 놓은 음성에 대한 에너지 및 영교차율의 기준 값을 넘어서는 경우가 발생되기 때문이다.

피치(Pitch) 정보를 이용한 기존의 음성 검출 방법은 우선 에너지와 영교차율 정보를 이용하여 일차적인 음성 영역을 검출하고, 검출된 영역의 최대 에너지 구간에는 유성음이 존재한다고 가정한 후, 최대 에너지 구간 주변에서 피치 정보를 구한다. 그리고 이 구간에서 추정된 피치가 사람의 피치 존재 범위를 나타내는 특정계수 범위내에 있는가를 판단하며 사람의 피치 존재 범위내에 있는 것으로 판단되면 입력 신호에 음성이 존재한다고 판단한다. 이 방법의 경우에 분석 구간 안에서 피치 발생 여부에 초점을 맞추고 있다. 다시 말하면 입력 신호에서 검출된 피치가 특정 범위 내에 존재할 경우에 음성으로 간주한다. 그런데 주변 잡음의 신호원은 그 종류가 매우 다양하기 때문에 모든 주파수 대역에 존재할 수 있다. 따라서, 주변 잡음이 위의 조건을 만족시키는 경우가 종종 발생되므로 잘못된 음성 검출이 발생할 수 있다.

결론적으로 종래의 음성 인식 방법은 음성 검출 과정, 음성 특징 추출 과정, 기준 패턴 비교 과정, 인식 결과 검증 과정 및 인식 결과 출력 과정을 순차적으로 거쳐 수행되며, 특히 음성 검출 과정에 있어서 에너지 및 영교차율과 피치 정보를 이용한 방법이 사용됨으로 인하여 음성 명령어가 아닌 비음성이 음성으로 잘못 검출되는 문제점이 있었다.

이를 구체적으로 살펴보면 다음과 같다.

도 1 은 종래의 음성 검출 방법에 대한 흐름도이다.

먼저, 사용자의 음성 신호를 입력받아(10) 아날로그 음성신호를 디지털 신호로 변환한 후에(11) 분석 크기 만큼의 음성 데이터에 대한 에너지 및 영교차율을 계산한다(12). 이후, 에너지 및 영교차율의 계산 결과가 이미 설정된 기준값보다 크지를 판단하여(13) 기준값보다 크면 다음 단계인 피치 검출과정으로 넘어가고, 기준값보다 크지 않으면 절차를 종료한다. 피치 검출 과정에서는 검출된 영역의 최대 에너지 구간에는 유성음이 존재한다고 가정한 후 최대에너지 구간 주변에서 피치 정보를 구한다(14). 이후에 이 구간에서 추정된 피치가 사람의 피치 존재 범위를 나타내는 특정계수 범위내에 있는가를 판단하여(15) 사람의 피치 존재 범위내에 있는 것으로 판단되면 음성을 검출한 후에(16) 종료하고 기준값을 넘어서지 않으면 바로 종료한다.

이러한 종래 기술은 사용자의 음성 뿐만아니라 주변의 잡음, 예를 들면 전화벨 소리, 책상 두드리는 소리, 문닫는 소리 및 마른 기침 소리 등이 미리 정해 놓은 음성에 대한 에너지 및 영교차율의 기준값을 넘어서는 경우와 입력 신호로부터 검출된 피치가 특정 범위내에 존재하는 경우에 비음성을 음성신호로 인식하여 불필요한 동작을 실행하는 경우가 종종 발생한다는 문제점이 있었다.

발명이 이루고자하는 기술적 과제

상기한 바와 같은 문제점을 해결하기 위하여 안출된 본 발명은, 외부로부터 입력된 신호에 대하여 에너지와 영교차율을 이용하여 음성인지 여부를 판단한 후에 음성 신호라고 판단되면 더 나아가서 연속적인 피치 정보를 이용하여 피치가 기준 피치의 범위내인지, 피치의 주파수 변화량이 기준값의 범위내인지 및 피치 발생수가 기준값을 넘었는지의 여부를 판단하여 상기의 조건을 만족하면 음성 신호로 판단하여 음성을 검출하는 음성 검출 방법을 제공하는데 그 목적이 있다.

발명의 구성 및 작용

상기 목적을 달성하기 위한 본 발명은, 음성 검출 장치에 적용되는 음성 검출 방법에 있어서, 외부로부터

터 입력되는 음성 신호의 피치 정보를 검출하는 제 1 단계; 상기 검출된 피치가 소정의 기준 피치 영역에 존재하는지를 확인하는 제 2 단계; 주파수 변화량을 계산한 후에 계산된 주파수 변화량이 소정의 제 1 기준값보다 작은지를 확인하는 제 3 단계; 및 상기 각 단계에서 주어진 조건을 만족하는 신호를 음성 신호로 판단하여 검출하는 제 4 단계를 포함하여 이루어지는 것을 특징으로 한다.

또한, 본 발명은, 음성 검출 장치에 적용되는 음성 검출 방법에 있어서, 외부로부터 입력되는 음성 신호의 피치 정보를 검출하는 제 1 단계; 상기 검출된 피치가 소정의 기준 피치 영역에 존재하는지를 확인하는 제 2 단계; 상기 피치 정보에서 피치 발생수를 검출하여 소정의 기준값을 넘는지를 확인하는 제 3 단계; 및 상기 각 단계에서 주어진 조건을 만족하는 신호를 음성 신호로 판단하여 검출하는 제 4 단계를 포함하여 이루어지는 것을 특징으로 한다.

이하 첨부된 도 2와 도 3을 참조하여 본 발명에 따른 실시예를 상세히 설명한다.

도 2는 본 발명이 적용되는 음성 인식 시스템의 실시예 구성도이다.

음성 입력 장치(21)는 외부로부터 사용자의 음성 신호를 수신하고, 아날로그/디지털 변환 장치(22)는 음성 입력 장치(21)로부터 출력된 아날로그 신호를 입력받아 디지털 신호로 변환하여 출력하며, 중앙 처리 장치(23)는 아날로그/디지털 변환 장치(22)로부터 디지털 음성 데이터를 입력받아 음성 구간을 검출하고 이 영역에 대한 음성 특징 벡터를 추출하며 이 특징 벡터와 기억장치(24)내에 저장되어 있던 각 어휘들의 기준 패턴과 비교하여 가장 유사한 어휘를 찾은 후에 인식 결과에 대한 검증 과정을 거쳐 인식 결과 출력 장치(25)를 통해서 음성 인식 결과를 출력시킨다.

음성 인식 시스템에서 사용되는 음성 신호는 자음과 모음의 결합으로 이루어진다. 따라서, 음성 신호에는 피치가 존재한다. 일반적으로 남성의 피치는 낮고 여성의 피치는 상대적으로 남성보다 높다. 음성 인식을 위해 사용자가 발성한 음성 신호에는 사람에 따라서 그 영역은 다르지만 고유한 피치 존재 영역을 갖고 있다. 이것은 (수학식 1)로서 표시할 수 있다.

$$Pitch\ Min \leq Pitch(t) \leq Pitch\ Max$$

(수학식 1)에서 Pitch Min은 사용자 음성 신호에 대한 피치의 최소값을 의미하고, Pitch Max는 피치의 최대값을 의미한다. 한편, 비음성은 거의 모든 주파수 대역에 존재한다. 따라서, 단순히 특정 주파수 대역에서 피치가 검출되었다고 해서 입력 신호에 사용자의 음성이 존재한다고 판단하는 것은 성급한 일이다.

피치는 성대의 진동에 의해서 발생된다. 음성 인식 시스템을 사용하는 사람이 평범하게 음성 신호를 발성할 경우에 해당 발화 내에서 피치 변화는 심하지 않다. 특히, 서로 이웃하는 피치의 존재 영역 즉, 주파수 영역은 매우 유사한 특성을 갖는다. 이것은 (수학식 2)로서 표시될 수 있다.

$$| Pitch(t) - Pitch(t+1) | \leq Pitch\ TH$$

(수학식 2)에서 Pitch TH는 이웃하는 피치 사이의 변화량인데 사람에 따라서 다르다. 한편, 대부분의 비음성은 값이 매우 크게 나타난다. 이러한 특성은 음성 신호와 비음성을 구별하는데 도움을 준다.

사용자의 음성 신호는 비음성과는 달리 항상 위의 두 가지 조건을 동시에 만족한다. 또한 위의 두 가지 조건을 만족하는 피치의 발생 수는 N개 이상이 된다. N은 음성 신호의 종류 및 사용자의 발성 방법에 따라서 달라지므로 음성 인식 시스템의 사용 조건에 따라서 음성 검출 장치에서 사용되는 N의 값은 바뀔 수 있다.

도 3은 본 발명에 따른 연속 피치 정보를 이용한 음성 검출 방법에 대한 실시예를 나타내는 흐름도이다.

먼저, 사용자의 음성 신호를 입력받아(10) 아날로그 음성신호를 디지털 신호로 변환한 후에(11) 분석 크기 만큼의 음성 데이터에 대한 에너지 및 영교차율을 계산한다(12). 이후, 에너지 및 영교차율의 계산 결과가 이미 설정된 기준값보다 크지를 판단하여(13) 기준값보다 크면 다음 단계인 피치 검출 과정으로 넘어가고, 기준값보다 크지 않으면 절차를 종료한다. 피치 검출 과정에서는 검출된 영역의 최대 에너지 구간에는 유성음이 존재한다고 가정된 후에 최대에너지 구간 주변에서 피치 정보를 구한다(14). 이후에 이 구간에서 추정된 피치가 사람의 피치 존재 범위를 나타내는 특정계수 범위내에 있는가를 판단하여(15) 특정 계수 범위내에 있는 것으로 판단되면 주파수 변화량 차이 비교 과정으로 넘어가고, 특정 계수 범위내에 있지 않다고 판단되면 절차를 종료한다. 주파수 변화량 차이 비교 과정에서는 서로 이웃하는 피치값 사이의 주파수 변화량을 계산하여(30) 설정된 기준값을 넘어서는지 여부를 비교하여(31) 설정된 기준값을 넘어서지 않으면 다음 과정인 피치 발생수 판단 과정으로 넘어가고, 그렇지 않으면 종료한다. 피치 발생수 판단 과정에서는 음성 신호의 발생 피치수를 검출한 후에(32), 검출한 피치 발생수가 기준값을 넘어서는지 여부를 판단하여(33) 기준값을 넘어서면 음성을 검출한 후에(16) 절차를 종료하고 그렇지 않으면 바로 종료한다.

이상에서 설명한 본 발명에서 주파수 변화량 차이 비교 과정(31)과 피치 발생수 판단 과정(33) 중 한 과정을 실시하지 않아도 상기의 목적에 맞는 효과를 얻을 수 있다.

또한, 본 발명에서 피치 존재 범위 판단 과정(15), 주파수 변화량 차이 비교 과정(31) 및 피치 발생수 판단 과정(33)은 반드시 순차적으로 실시되어야 하는 것은 아니며 임의적인 순서로 실시되어도 동일한 효과를 얻을 수 있다.

또한, 본 발명은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 있어 본 발명의 기술적 사상을 벗어나지 않는 범위내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시예 및 첨부된 도면에 한정되는 것이 아니다.

발명의 효과

상기와 같은 본 발명은, 음성 명령어가 아닌 비음성 신호의 입력을 사전에 막을 수 있어 음성 인식 시스템의 성능을 향상시킬 수 있으며, 비음성을 음성으로 인식하여 사용자가 원하지 않는 불필요한 동작을 수행하는 것을 방지할 수 있어 비음성 신호에 대한 잘못된 반응을 방지할 수 있는 효과가 있다.

(57) 청구의 범위**청구항 1**

음성 검출 장치에 적용되는 음성 검출 방법에 있어서,
외부로부터 입력되는 음성 신호의 피치 정보를 검출하는 제 1 단계;
상기 검출된 피치가 소정의 기준 피치 영역에 존재하는지를 확인하는 제 2 단계;
주파수 변화량을 계산한 후에 계산된 주파수 변화량이 소정의 제 1 기준값보다 작은지를 확인하는 제 3 단계; 및
상기 각 단계에서 주어진 조건을 만족하는 신호를 음성 신호로 판단하여 검출하는 제 4 단계를 포함하여 이루어진 연속 피치 정보를 이용한 음성 검출 방법.

청구항 2

제 1 항에 있어서,
상기 제 4 단계 수행전에,
상기 피치 정보에서 피치 발생수를 검출하여 피치 발생수가 소정의 제 2 기준값을 넘는지를 확인하는 제 5 단계를 더 포함하여 이루어진 연속 피치 정보를 이용한 음성 검출 방법.

청구항 3

제 2 항에 있어서,
상기 제 5 단계는,
상기 피치 정보에서 피치 발생수를 검출하는 제 6 단계;
상기 검출된 피치 발생수가 상기 소정의 제 2 기준값을 넘는지를 비교하는 제 7 단계; 및
상기 제 7 단계의 비교 결과, 피치 발생수가 상기 소정의 제 2 기준값을 넘으면 다음 단계를 수행하고, 피치 발생수가 상기 제 2 기준값을 넘지 않으면 종료하는 제 8 단계를 포함하여 이루어진 연속 피치 정보를 이용한 음성 검출 방법.

청구항 4

제 1 항 내지 제 3 항 중 어느 한 항에 있어서,
상기 제 3 단계는,
상기 피치 정보에서 주파수 변화량을 계산하는 제 9 단계;
상기 계산된 주파수 변화량이 상기 소정의 제 1 기준값보다 작은지를 비교하는 제 10 단계; 및
상기 제 10 단계의 비교 결과, 주파수 변화량이 상기 소정의 제 1 기준값보다 작으면 다음 단계를 수행하고, 주파수 변화량이 상기 제 1 기준값보다 작지 않으면 종료하는 제 11 단계를 포함하여 이루어진 연속 피치 정보를 이용한 음성 검출 방법.

청구항 5

음성 검출 장치에 적용되는 음성 검출 방법에 있어서,
외부로부터 입력되는 음성 신호의 피치 정보를 검출하는 제 1 단계;
상기 검출된 피치가 소정의 기준 피치 영역에 존재하는지를 확인하는 제 2 단계;
상기 피치 정보에서 피치 발생수를 검출하여 소정의 기준값을 넘는지를 확인하는 제 3 단계; 및
상기 각 단계에서 주어진 조건을 만족하는 신호를 음성 신호로 판단하여 검출하는 제 4 단계를 포함하여 이루어진 연속 피치 정보를 이용한 음성 검출 방법.

청구항 6

제 5 항에 있어서,
상기 제 3 단계는,

상기 피치 정보에서 피치 발생수를 검출하는 제 5 단계;

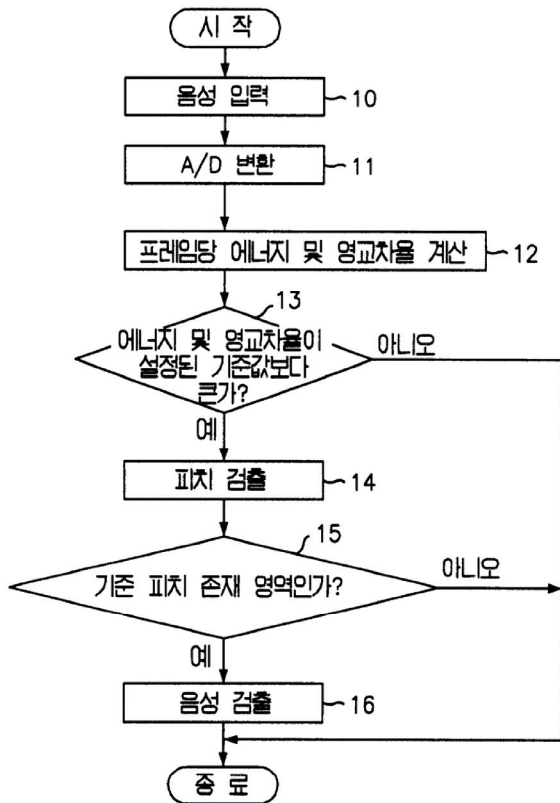
상기 검출된 피치 발생수가 상기 소정의 기준값을 넘는지를 비교하는 제 6 단계; 및

상기 제 6 단계의 비교 결과, 피치 발생수가 상기 소정의 기준값을 넘으면 다음 단계를 수행하고, 피치 발생수가 상기 기준값을 넘지 않으면 종료하는 제 7 단계

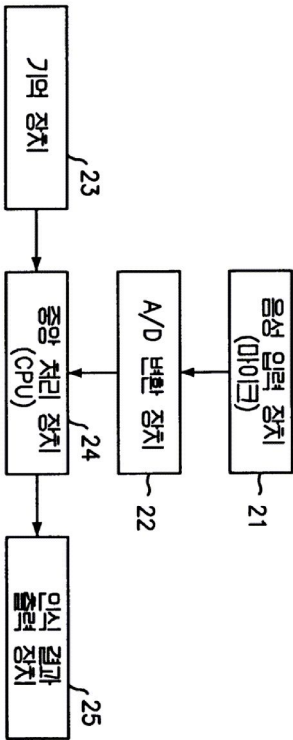
를 포함하여 이루어진 연속 피치 정보를 이용한 음성 검출 방법.

도면

도면1



도면2



도면3

