



- (51) 국제특허분류:
G10L 17/02 (2013.01) G06F 17/10 (2006.01)
- (21) 국제출원번호: PCT/KR2016/014673
- (22) 국제출원일: 2016년 12월 14일 (14.12.2016)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2015-0183897 2015년 12월 22일 (22.12.2015) KR
- (71) 출원인: 경상대학교산학협력단 (INDUSTRY-ACADEMIC COOPERATION FOUNDATION GYEONGSANG NATIONAL UNIVERSITY) [KR/KR]; 52828 경상남도 진주시 진주대로 501, Gyeongsangnam-do (KR).
- (72) 발명자: 정상배 (JEONG, Sang Bae); 52735 경상남도 진주시 초장로 14 번길 27, 109 동 804 호, Gyeongsangnam-do (KR). 강지훈 (KANG, Jihoon); 63292 제주도 제주시 진남로 2 길 11, 202 호, Jeju-do (KR). 김영일 (KIM, Youngil); 52716 경상남도 진주시 진주대로 859, 3-1201, Gyeongsangnam-do (KR).
- (74) 대리인: 권혁수 (KWON, Hyuk-Soo) 등; 06226 서울시 강남구 언주로 329, 4 층, Seoul (KR).

(81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

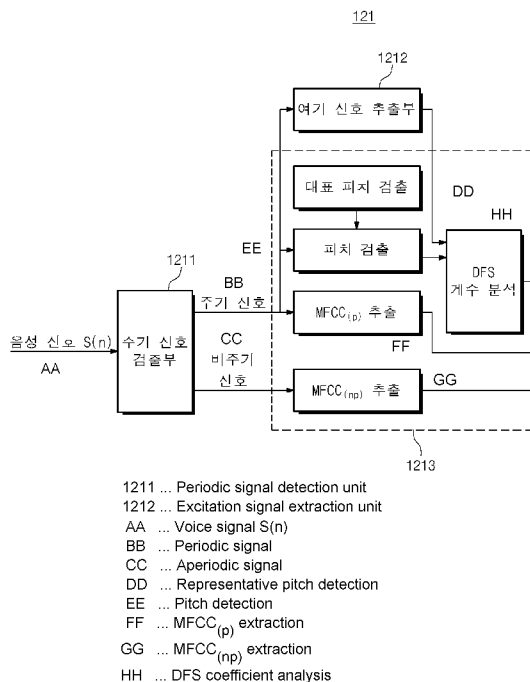
(84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 역내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유럽 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

공개:

— 국제조사보고서와 함께 (조약 제 21 조(3))

(54) Title: APPARATUS FOR EXTRACTING FEATURE PARAMETERS OF INPUT SIGNAL, AND SPEAKER RECOGNITION APPARATUS USING SAME

(54) 발명의 명칭: 입력 신호의 특징 파라미터 추출 장치 및 그를 이용한 화자 인식 장치



(57) Abstract: The present invention provides a feature parameter extraction apparatus and a speaker recognition apparatus using the same, which improve a speaker recognition rate by extracting an excitation signal in a periodic section of an input signal and using feature parameters calculated therefrom. A feature parameter extraction apparatus according to an embodiment of the present invention may comprise: a periodic signal detection unit for detecting a periodic section of an input signal; an excitation signal extraction unit for extracting an excitation signal in the periodic section of the input signal; and a feature parameter calculation unit for calculating feature parameters characterizing the input signal, on the basis of the frequency response spectrum of the excitation signal.

(57) 요약서: 본 발명은 입력 신호의 주기적인 구간에서 여기 신호를 추출하여 그로부터 산출된 특징 파라미터를 이용함으로써 화자 인식을 높인 특징 파라미터 추출 장치 및 그를 이용한 화자 인식 장치를 제공한다. 본 발명의 일 실시 예에 따른 특징 파라미터 추출 장치는 입력 신호의 주기적인 구간을 검출하는 주기 신호 검출부; 상기 입력 신호의 주기적인 구간에서 여기 신호(excitation signal)를 추출하는 여기 신호 추출부; 및 상기 여기 신호의 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부를 포함할 수 있다.

명세서

발명의 명칭: 입력 신호의 특징 파라미터 추출 장치 및 그를 이용한 화자 인식 장치

기술분야

- [1] 본 발명은 입력 신호의 특징 파라미터를 추출하는 장치 및 그를 이용하여 화자를 인식하는 장치에 관한 것이다.

배경기술

- [2] 인간이 발성하는 음성을 컴퓨터가 처리하여 이해하는 음성 처리 기술은 각종 분야에서 활용될 수 있는 전도유망한 기술이다. 특히, 입력받은 음성을 기초로 화자를 식별하는 화자 인식 기술은 보안 시스템에서 본인 확인을 위해 사용되거나 지능형 로봇에서 사용자 식별을 위해 적용될 수 있다.
- [3] 일반적으로 화자 인식을 포함한 음성 인식은 음성의 입력 신호로부터 특징 벡터를 추출하고 이를 사전에 저장된 데이터와 비교하여 정보를 인식한다. 그러나 현재 기술의 음성 인식률은 다양한 분야에 걸쳐 상용화되기에는 한계가 있으며, 그 중에서도 화자 인식 기술은 그 인식률이 높지 않아 지속적인 연구개발이 필요하다.

발명의 상세한 설명

기술적 과제

- [4] 본 발명은 화자 인식률을 개선할 수 있는 특징 파라미터 추출 장치 및 그를 이용한 화자 인식 장치를 제공하는 것을 목적으로 한다.

과제 해결 수단

- [5] 본 발명의 일 실시 예에 따른 특징 파라미터 추출 장치는, 입력 신호의 주기적인 구간을 검출하는 주기 신호 검출부; 상기 입력 신호의 주기적인 구간에서 여기 신호(excitation signal)를 추출하는 여기 신호 추출부; 및 상기 여기 신호의 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부를 포함할 수 있다.
- [6] 상기 주기 신호 검출부는, 상기 입력 신호에 대한 자기 상관 함수(auto-correlation function)의 결과값을 기반으로 주기적인 구간을 검출할 수 있다.
- [7] 상기 주기 신호 검출부는, 상기 입력 신호가 $x(n)$ 이고, 상기 입력 신호에 대한 $R(T)/R(0)$ 가 기 설정된 임계치 이상일 경우 주기가 T 인 주기적인 신호로 판단할 수 있으며, 여기서 $R(T)=\sum x(n)x(n+T)$ 일 수 있다.
- [8] 상기 특징 파라미터 추출 장치는, 상기 입력 신호의 피치(pitch) 값 중 대표 피치 값을 결정하는 대표 피치 검출부; 및 상기 입력 신호의 주기적인 구간에서의 피치 값 중, 상기 대표 피치 값을 기반으로 적정 피치 값을 결정하는 적정 피치 검출부를 더 포함하며, 상기 특징 파라미터 추출부는 상기 적정 피치 값을

- 기반으로 상기 특징 파라미터를 산출할 수 있다.
- [9] 상기 대표 피치 검출부는, 상기 입력 신호의 피치 값을 크기 순으로 나열하여 중간값을 상기 대표 피치 값으로 결정할 수 있다.
- [10] 상기 적정 피치 검출부는, 상기 주기적인 구간의 입력 신호에 대한 자기 상관 함수의 결과를 기반으로 선택된 피치 값들 중, 상기 대표 피치 값에 가장 근사한 값을 적정 피치 값으로 결정할 수 있다.
- [11] 상기 여기 신호 추출부는, 상기 주기적인 구간의 입력 신호를 전처리하여, 상기 입력 신호의 생성 과정에서 손실된 고주파 성분이 보상된 전처리 신호를 출력하는 선강조부를 포함할 수 있다.
- [12] 상기 여기 신호 추출부는, 상기 전처리 신호에 대한 자기 상관 함수의 결과값을 출력하는 자기 상관 함수 추정부; 및 상기 자기 상관 함수 추정부의 출력을 입력받아, 레빈슨 더빈(Levinson-Durbin) 알고리즘을 기반으로 추측 계수를 출력하는 추측 계수 산출부; 및 상기 전처리 신호 및 상기 추측 계수를 기반으로 역 필터링(inverse filtering)을 수행하여 상기 여기 신호를 출력하는 역 필터링부를 포함할 수 있다.
- [13] 상기 특징 파라미터 산출부는, 상기 여기 신호를 상기 적정 피치 값을 기반으로 이산 푸리에 변환(Discrete Fourier Transform)하여 주파수 영역의 이산 푸리에 스펙트럼으로 변환하는 주파수 영역 변환부를 포함할 수 있다.
- [14] 상기 특징 파라미터 산출부는, 상기 이산 푸리에 스펙트럼의 차수별 크기의 로그 값을 상기 특징 파라미터로 산출할 수 있다.
- [15] 상기 특징 파라미터 산출부는, 상기 입력 신호의 주기적인 구간 및 비주기적인 구간에 대하여, 상기 입력 신호의 주파수 응답 스펙트럼에 멜-주파수(Mel-Frequency) 필터를 적용하여 멜 주파수 응답을 획득하는 멜-주파수 응답 획득부; 및 상기 멜-주파수 응답을 역 이산 코사인 변환하여 캡스트럼 계수(cepstral coefficient)를 획득하는 캡스트럼 계수 획득부를 포함할 수 있다.
- [16] 상기 특징 파라미터 산출부는, 상기 특징 파라미터 산출부와 상기 캡스트럼 계수 획득부의 출력에 각각 미리 설정된 가중치를 곱하여 합한 값을 특징 파라미터로 산출할 수 있다.
- [17] 본 발명의 일 실시 예에 따른 화자 인식 장치는, 화자의 음성을 수집하는 음성 수집부; 상기 수집된 음성을 처리하여 기 등록된 사용자의 음성과 일치 여부를 판별하는 음성 처리부; 및 상기 사용자의 음성에 관한 정보를 저장하는 저장부를 포함할 수 있다.
- [18] 상기 음성 처리부는, 입력 신호의 신호가 주기적인 구간을 검출하는 주기 신호 검출부; 상기 입력 신호의 주기적인 구간에서 여기 신호(excitation signal)를 추출하는 여기 신호 추출부; 및 상기 여기 신호의 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부를 포함할 수 있다.

발명의 효과

- [19] 본 발명의 일 실시 예에 따르면, 음성 신호 처리에 있어서 음성 인식률을 높은 특징 파라미터 추출 장치를 얻을 수 있으며, 이를 이용하여 화자 인식률이 향상된 화자 인식 장치를 얻을 수 있다.

도면의 간단한 설명

- [20] 도 1은 본 발명의 일 실시 예에 따른 화자 인식 장치의 예시적인 블록도이다.
- [21] 도 2는 본 발명의 일 실시 예에 따른 특징 파라미터 추출부의 예시적인 블록도이다.
- [22] 도 3은 도 2의 특징 파라미터 추출부에서 대표 피치(pitch)를 검출하는 방법의 예시적인 순서도이다.
- [23] 도 4 및 5는 본 발명의 일 실시 예에 따라 입력 신호에서 대표 피치를 검출하는 방법을 설명하기 위한 그래프이다.
- [24] 도 6는 도 2의 특징 파라미터 추출부의 주기 신호 검출부에서 검출된 주기 신호에 대한 자기 상관 함수의 분석 그래프이다.
- [25] 도 7은 도 2의 특징 파라미터 추출부에서 적정 피치를 검출하는 방법의 예시적인 순서도이다.
- [26] 도 8은 본 발명의 일 실시 예에 따라 입력 신호에서 적정 피치를 검출하는 방법을 설명하기 위한 그래프이다.
- [27] 도 9은 도 2의 특징 파라미터 산출부에서 여기 신호를 추출하는 방법의 예시적인 순서도이다.
- [28] 도 10는 본 발명의 일 실시 예에 따라 화자 인식률이 개선된 것을 보여주는 표이다.

발명의 실시를 위한 형태

- [29] 이하, 본 명세서에 첨부된 도면을 참조하여 본 발명의 실시 예를 상세하게 설명한다.
- [30] 도 1은 본 발명의 일 실시 예에 따른 화자 인식 장치(10)의 예시적인 블록도이다.
- [31] 도 1에 도시된 바와 같이, 본 발명의 일 실시 예에 따른 화자 인식 장치(10)는 음성 수집부(110), 음성 처리부(120), 및 저장부(130)를 포함한다.
- [32] 상기 음성 수집부(110)는 화자의 음성을 수집한다. 본 발명의 일 실시 예에 따르면, 상기 음성 수집부(110)는 화자가 발성한 음성을 전기적인 신호로 변환하는 마이크를 포함할 수 있다. 그러나, 상기 음성 수집부(110)는 화자의 음성을 화자로부터 직접 수집하는 마이크로 제한되지 않고, 화자의 음성과 관련된 신호를 다양한 방법(예컨대, 네트워크를 통한 데이터 통신)으로 입수하는 임의의 장치를 포함한다.
- [33] 상기 음성 처리부(120)는 수집된 음성을 처리하여 기 등록된 사용자의 음성과의 일치 여부를 판별한다. 본 발명의 일 실시 예에 따르면, 상기 음성 처리부(120)는

음성과 관련된 전기 신호(이하, 음성 신호)를 소정의 알고리즘에 따라 처리하는 프로세서를 포함하며, 일 예로 CPU 등을 포함할 수 있으나 이에 제한되지는 않는다. 이 경우, 상기 음성 처리부(120)는 저장부(130)에 저장된 프로그램을 불러와 실행하여 그에 따라 음성 신호를 처리할 수 있으며, 이러한 처리를 통해 얻은 데이터는 상기 저장부(130)에 저장될 수 있다.

[34] 상기 저장부(130)는 사용자의 음성에 관한 정보를 저장한다. 본 발명의 일 실시 예에 따르면, 상기 저장부(130)는 데이터 또는 각종 프로그램을 저장할 수 있는 장치로서, 일 예로 HDD, SSD 등과 같은 대용량의 저장 장치뿐만 아니라 RAM, ROM 등의 각종 메모리, 레지스터, 캐쉬 등을 포함할 수 있다.

[35] 본 발명의 실시 예에 따르면, 상기 음성 처리부(120)는 화자의 음성이 기 등록된 사용자의 음성으로 일치하는지 여부를 판별하기 위해 화자의 음성으로부터 그 음성을 특징짓는 특징 파라미터를 추출하는 특징 파라미터 추출부(121)를 포함한다.

[36] 도 2는 본 발명의 일 실시 예에 따른 특징 파라미터 추출부의 예시적인 블록도이다.

[37] 도 2를 참조하면, 상기 특징 파라미터 추출부(121)는 주기 신호 검출부(1211), 여기 신호 추출부(1212), 및 특징 파라미터 산출부(1213)를 포함할 수 있다.

[38] 상기 특징 파라미터 산출부(1213)는, 여기 신호 추출 및 특징 파라미터 산출에 필요한 대표 피치(pitch) 값을 결정하는 대표 피치 검출부(12131)를 포함할 수 있다.

[39] 도 3은 도 2의 특징 파라미터 추출부(121)에 포함된 대표 피치 검출부(12131)에서 대표 피치(pitch)를 검출하는 방법의 예시적인 순서도이다.

[40] 도 3을 참조하면, 대표 피치 검출부(12131)는 자기 상관 함수의 분석을 통해 피치를 검출한다. 자기 상관 함수는 어떤 임의의 신호가 서로 다른 두 시각에 취하는 값의 상관을 나타내는 함수이다. 예를 들어, t 시각과 $t+\tau$ 시각으로부터 일정 시간 지연된 $t+\tau$ 시각에 취하는 값의 상관도를 나타낼 수 있다. 피치는 음성 신호의 자기 상관 함수가 최대가 되는 위치에서 검출할 수 있다. 구체적으로, 아래 수학적식을 통해 검출할 수 있다.

[41]

$$r_x(\tau) = \sum x(n)x(n+\tau), (\tau = nT_0, r_x(nT_0) = r_x(0))$$

[42] n 은 정수이며, 주기 신호의 기본주기는 T_0 로 정의된다.

[43] 다시 도 3을 참조하면, 대표 피치 검출부(12131)는 입력 신호의 피치 값을 크기 순으로 정렬하여, 중간값을 대표 피치 값으로 결정할 수 있다.

[44] 도 4 및 5는 본 발명의 일 실시 예에 따라 입력 신호에서 대표 피치를 검출하는 방법을 설명하기 위한 그래프이다.

[45] 도 4는 예시적인 입력 신호로부터 검출된 피치 값의 분포이다. 도 4에 도시된 바와 같이, 피치 값이 갑자기 증가하거나(pitch doubling), 갑자기 감소하는

현상(pitch havening)이 발생할 수 있으며, 이러한 피치 값을 이용하여 음성 인식을 수행할 경우 음성 인식을 및 화자 인식이 저하될 수 있다. 따라서 이러한 피치 추출 오류를 보정하기 위해, 대표 피치 검출부(12131)은 도 5에 도시된 바와 같이 피치 값을 크기 순으로 정렬하여, 중간 값을 대표 피치 값으로 결정할 수 있다.

- [46] 대표 피치 검출부(12131)는, 임의의 화자에 대하여 해당 화자의 대표 피치를 결정하여, 추후의 음성 인식 및 화자 인식에 있어서 사용되는 피치 값을 상기 대표 피치 값을 기반으로 보정할 수 있다. 즉, 대표 피치 값을 이용하여 분석 대상이 되는 음성 프레임에서의 피치 추출 오류를 보정할 수 있다. 결정된 대표 피치 값은 주기적인 신호의 분석에 있어서 적정 피치 값을 결정하는 데 이용된다. 이와 관련된 내용은 도 7 및 7을 참조하여 구체적으로 설명될 것이다.
- [47] 다시 도 2를 참조하면, 주기 신호 검출부(1211)는 입력 신호의 주기적인 구간과 비주기적인 구간을 구분한다.
- [48] 입력 신호로 음성 신호가 입력될 수 있다. 음성은 성대에 진동을 일으키는 유성음과 성대의 진동을 일으키지 않는 무성음으로 구분되며, 유성음 구간에서 주기적인 여기 신호의 검출이 가능하다. 본 발명의 일 실시 예에 따른 특징 파라미터 추출 장치는 여기 신호로부터 특징 파라미터를 검출하여 음성 인식률을 향상시키기 위한 보조적인 파라미터로 사용한다. 따라서 입력 신호에서 유성음 구간, 즉 주기적인 신호의 구간을 구분하는 과정이 필요하다.
- [49] 본 발명의 일 실시 예에 따르면, 주기 신호 검출부(1211)는 입력 신호에 대한 자기 상관 함수(auto-correlation fuction)의 결과값을 기반으로 주기적인 구간을 검출한다. 주기 신호 검출부(1211)는 자기 상관 함수의 결과값이 기 설정된 임계치 이상일 경우 주기적인 신호로 판단할 수 있다.
- [50] 도 6는 도 2의 특징 파라미터 추출부의 주기 신호 검출부에서 검출된 주기 신호에 대한 자기 상관 함수의 분석 그래프이다. 주기 신호 검출부(1211)는 입력 신호에 대한 자기 상관 함수의 그래프 윤곽을 분석하여 신호의 주기성을 판단할 수 있다. 도 6를 참조하면, T에 대한 $R(T)/R(0)$ 값이 임계치(R_{th}) 이상인 구간에서 입력 신호가 주기적인 신호인 것으로 판단할 수 있다. 즉, 주기 신호 검출부(1211)는 입력 신호가 상기 입력 신호가 $x(n)$ 이고, 상기 입력 신호에 대한 $R(T)/R(0)$ 가 기 설정된 임계치 이상일 경우 주기가 T인 주기적인 신호로 판단할 수 있다. 여기서 $R(T)=\sum x(n)x(n+T)$ 일 수 있다.
- [51] 다시 도 2를 참조하면, 특징 파라미터 산출부(1213)는 입력 신호의 주기적인 구간에서의 피치 값 중 적정 피치 값을 결정하는 적정 피치 검출부(12132)를 포함한다. 도 7은 적정 피치 검출부(12132)에서 적정 피치를 검출하는 방법의 예시적인 순서도이다.
- [52] 도 7을 참조하면, 적정 피치 검출부(12132)는 입력 신호에 대한 자기 상관 함수를 추정하고 자기 상관 함수 그래프의 윤곽을 분석하여 적정 피치 값을 결정할 수 있다. 도 8은 본 발명의 일 실시 예에 따라 입력 신호에서 적정 피치를

검출하는 방법을 설명하기 위한 그래프이다. 적정 피치 검출부(12132)는 자기 상관 함수를 기반으로 후보 피치 값들을 선출할 수 있다. 도 2에 도시된 바와 같이, 적정 피치 검출부(12132)는 후보 피치 값들 중 대 피치 값을 기반으로 적정 피치 값을 결정할 수 있다. 적정 피치 검출부(12132)는 후보 피치 값들 중 대표 피치 값에 가장 근사한 값을 적정 피치 값으로 결정할 수 있다. 대표 피치 값에 가장 근사한 값을 적정 피치 값으로 이용함으로써, 상술한 바와 같이 pitch doubling 및 pitch havening 현상이 일어날 가능성을 줄일 수 있다.

[53] 도 9은 도 2의 여기 신호 추출부(1212)에서 여기 신호를 추출하는 방법의 예시적인 순서도이다.

[54] 도 9을 참조하면, 여기 신호 추출부(1212)는 선강조부를 포함할 수 있다. 상기 선강조부는 주기적인 구간의 입력 신호 $s(n)$ 을 전처리하여, 상기 입력 신호의 생성 과정에서 손실된 고주파 성분이 보상된 전처리 신호 $s_{pre}(n)$ 을 출력할 수 있다. 화자가 음성을 발성할 때 입술에서 자유 공간으로 공기를 방사하는 과정에서 고주파 성분이 손실된다. 상기 선강조부는 입력 신호에 선강조 필터를 적용하여 손실된 고주파 성분을 보상한다. 일 예로, 상기 선강조 필터는 아래의 수학적식과 같은 전달함수를 갖도록 구현될 수 있다.

[55]

$$H(z)=1-az^{-1}, (0.9 \leq a \leq 1)$$

[56] 통상적으로 $a=0.97$ 일 수 있다.

[57] 다시 도 9을 참조하면, 여기 신호 추출부(1212)는 상기 전처리 신호에 대한 자기 상관 함수를 추정하여 결과값을 출력하는 자기 상관 함수 출력부, 상기 자기 상관 함수 추정부의 출력을 입력받아 레빈슨 더빈(Levinson-Durbin) 알고리즘을 기반으로 추측 계수(a_k)를 출력하는 추측 계수 산출부, 및 상기 전처리 신호와 상기 추측 계수를 기반으로 역 필터링(inverse filtering)을 수행하여 여기 신호 $e(n)$ 을 출력하는 역 필터링부를 포함할 수 있다.

[58] 상기 역 필터링부는 아래의 수학적식과 같은 전달함수를 갖도록 구현될 수 있다.

[59]

$$A(z)=1+\sum_{k=1}^P a_k z^{-k}$$

[60] 여기서 P 는 적정 피치 검출부(12132)에서 검출된 적정 피치 값일 수 있다. 적정 피치 검출부(12132)에서 검출된 적정 피치 값이 정수가 아닌 경우, 반올림 등의 과정을 거쳐 정수 P 값을 산출할 수 있다.

[61] 본 발명의 일 실시 예에 따라, 특징 파라미터 산출부(1213)는 상기 여기 신호 $e(n)$ 을 상기 적정 피치 값 P 를 기반으로 이산 푸리에 변환(Discrete Fourier Transform)하여 주파수 영역의 이산 푸리에 스펙트럼(Discrete Fourier Spectrum, DFS)으로 변환하는 주파수 영역 변환부를 포함할 수 있다.

- [62] 특징 파라미터 산출부(1213)는, 여기 신호의 이산 푸리에 스펙트럼 계수(DFS 계수)를 추출하여 입력 신호를 특징짓는 특징 파라미터로 산출할 수 있다. 여기 신호 추출부(1212)에서 추출된 여기 신호 $e(n)$ 이

$$[63] \quad e(n) = \sum_{k=0}^{p-1} E_k e^{j \frac{2\pi}{p} kn}$$

- [64] 일 때, 여기서 p 는 적정 피치 검출부에서 검출된 피치 값일 수 있다. 적정 피치 검출부에서 검출된 피치 값이 정수가 아닌 경우에는 반올림한 값이 p 가 될 수 있다.

- [65] 상기 여기 신호 $e(n)$ 을 이산 푸리에 변환하여 여기 신호에 대한 이산 푸리에 스펙트럼을 얻을 수 있다.

- [66] 상기 여기 신호 $e(n)$ 은 아래와 같이 변환 가능하다.

$$[67] \quad e(n) = \sum_{k=1}^{p-1} \left(A_k \cos\left(\frac{2\pi kn}{p}\right) + B_k \sin\left(\frac{2\pi kn}{p}\right) \right)$$

$$[68] \quad E_k = \frac{1}{p} \sum_{n=0}^{p-1} e(n) e^{-j \frac{2\pi}{p} kn}, (k=1, \dots, p-1)$$

- [69] 여기서 DFS 계수 A_k 및 B_k 를 이용하여 음성 신호를 특징짓는 특징 파라미터를 산출할 수 있다.

- [70] 일 실시 예에 있어서, 특징 파라미터 산출부(1213)는 DFS 계수를 이용하여 아래와 수학적식과 같이 DFS 크기를 구할 수 있다.

$$[71] \quad E_k = \sqrt{A_k^2 + B_k^2}$$

- [72] 일 실시 예에 따라, 특징 파라미터 산출부(1213)는, 피치 값에 해당하는 주파수의 조화 주파수(harmonic frequency)와 상기 산출된 DFS 계수의 크기에 $\log$ 를 취한 값을 한 쌍으로 구성한 특징 벡터를 산출할 수 있다. 산출된 특징 벡터의 수학적식은 아래와 같다.

$$[73] \quad D_i = \left\{ \left(\frac{1}{p}, \log |E_1| \right), \left(\frac{2}{p}, \log |E_2| \right), \dots, \left(\frac{K}{p}, \log |E_K| \right) \right\}$$

- [74] 여기서 K 는 피치 값의 조화(harmonic) 개수이며, E_k 는 입력 신호의 에너지로

정규화된 여기 신호에서 추출된 K번째 DFS 계수의 크기를 나타낸다.

- [75] 다시 도 2를 참조하면, 본 발명의 일 실시 예에 따른 특징 파라미터 추출부(1213)는 입력 신호로부터 MFCC(Mel Frequency Cepstral Coefficient)를 추출할 수 있다.
- [76] 일 실시 예에 따라, 특징 파라미터 추출부(1213)는 입력 신호의 주기적인 구간 및 비주기적인 구간에 대하여, 입력 신호의 주파수 응답 스펙트럼에 멜-주파수(Mel-Frequency) 필터를 적용하여 멜 주파수 응답을 획득하는 멜-주파수 응답 획득부, 및 상기 멜-주파수 응답을 역 이산 코사인 변환하여 캡스트럼 계수(cepstral coefficient)를 획득하는 캡스트럼 계수 획득부를 포함할 수 있다. 따라서, 상기 특징 파라미터 추출부(1213)는 입력 신호에 대한 MFCC를 추출하여 특징 파라미터로 사용할 수 있다.
- [77] 본 발명의 일 실시 예에 따른 특징 파라미터 추출부(1213)는, 상기 MFCC 뿐만 아니라 여기 신호에 대한 DFS 계수를 추가적으로 특징 파라미터로 추출할 수 있다. 따라서 DFS 계수의 보조적인 특징 파라미터로서의 활용에 따라 음성 인식을 및 화자 인식을 향상시킬 수 있다.
- [78] 다시 도 1을 참조하면, 본 발명의 일 실시 예에 따른 화자 인식 장치(10)에 있어서, 음성 처리부(120)는 수집된 음성을 처리하여 기 등록된 사용자의 음성과의 일치 여부를 판별할 수 있다.
- [79] 일 실시 예에 따라, 상기 음성 처리부(120)는 특징 파라미터 추출부(1213)는 상기 특징 파라미터를 토대로 화자의 스코어를 계산하고, 발성 화자를 인식할 수 있다. 구체적으로, 특정 화자마다 생성된 가우시안 혼합 모델과 입력된 음성 신호로부터 추출한 특징 파라미터의 유사도를 점수화하여, 발성 화자를 인식할 수 있다. 상기 가우시안 혼합 모델은 저장부(130)에 저장되어 있을 수 있다.
- [80] 일 실시 예에 있어서, 스코어는 아래 식과 같이 계산될 수 있다.
- [81]

$$score(i) = \sum_{t=0}^{T-1} \log P(x_t | S_i)$$

- [82] x_t 는 분석된 음성 프레임에서의 특징 벡터를, S_i 는 특정 화자에 대한 가우시안 혼합 모델 파라미터를 각각 나타낸다. 상기 P 는 화자의 음향에서 특징 벡터 x_t 가 발생할 확률을 나타낸다.
- [83] 본 발명의 일 실시 예에 따라, 스코어는 특징 파라미터 추출부(121)에서 추출된 다양한 파라미터를 적용하여 계산될 수 있다. 본 발명의 일 실시 예에 따라 주기적인 신호에서 여기 신호를 추출하여 보조적인 파라미터를 생성함으로써, 스코어의 정확성을 높일 수 있다. 다수의 파라미터를 이용하여 스코어를 산출할 경우, 각 파라미터에 가중치를 부여하여 스코어를 산출할 수 있다. 이 때 가중치는 전수 조사에 의해 부여될 수 있다.

- [84] 도 10는 본 발명의 일 실시 예에 따라 화자 인식률이 개선된 것을 보여주는 표이다. 도 10를 참조하면, 본 발명의 일 실시 예에 따라 스코어를 산출하여 화자 인식을 수행한 경우와, 비교 예에 따라 스코어를 산출하여 화자 인식을 수행한 경우의 화자 인식률이 각각 나타나있다. 실시 예에 의한 경우에는, MFCC 파라미터와 여기 신호의 DFS 계수 파라미터를 모두 사용하였고, 비교 예에 의한 경우에는 MFCC 파라미터만을 사용하였다. 스코어 가중치 α , β , 및 γ 는 각각 주기적인 신호에서 추출된 MFCC, 비주기적인 신호에서 추출된 MFCC, 및 여기 신호에서 추출된 DFS 계수 파라미터로부터 산출된 스코어에 부여된 가중치를 나타낸다.
- [85] 도 10의 표에 나타난 바와 같이, 본 발명의 일 실시 예에 따라 여기 신호로부터 추출한 추가적인 파라미터를 더 이용하였을 경우 화자인식률이 더 향상되었다.
- [86] 이상에서 실시 예를 통해 본 발명을 설명하였으나, 위 실시 예는 단지 본 발명의 사상을 설명하기 위한 것으로 이에 한정되지 않는다. 통상의 기술자는 전술한 실시 예에 다양한 변형이 가해질 수 있음을 이해할 것이다. 본 발명의 범위는 첨부된 특허청구범위의 해석을 통해서만 정해진다.

청구범위

- [청구항 1] 입력 신호를 입력받아 신호가 주기적인 구간을 검출하는 주기 신호 검출부;
 상기 입력 신호의 주기적인 구간에서 여기 신호(excitation signal)를 추출하는 여기 신호 추출부; 및
 상기 여기 신호의 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부를 포함하는 특징 파라미터 추출 장치.
- [청구항 2] 제1 항에 있어서,
 상기 주기 신호 검출부는,
 상기 입력 신호에 대한 자기 상관 함수(auto-correlation function)의 결과값을 기반으로 주기적인 구간을 검출하는 특징 파라미터 추출 장치.
- [청구항 3] 제2 항에 있어서,
 상기 주기 신호 검출부는,
 상기 입력 신호가 $x(n)$ 이고, 상기 입력 신호에 대한 $R(T)/R(0)$ 가 기 설정된 임계치 이상일 경우 주기가 T 인 주기적인 신호로 판단하며,
 여기서 $R(T)=\sum x(n)x(n+T)$ 인 특징 파라미터 추출 장치.
- [청구항 4] 제1 항에 있어서,
 상기 특징 파라미터 추출 장치는,
 상기 입력 신호의 피치(pitch) 값 중 대표 피치 값을 결정하는 대표 피치 검출부; 및
 상기 입력 신호의 주기적인 구간에서의 피치 값 중, 상기 대표 피치 값을 기반으로 적정 피치 값을 결정하는 적정 피치 검출부를 더 포함하며,
 상기 특징 파라미터 추출부는 상기 적정 피치 값을 기반으로 상기 특징 파라미터를 산출하는 특징 파라미터 추출 장치.
- [청구항 5] 제4 항에 있어서,
 상기 대표 피치 검출부는,
 상기 입력 신호의 피치 값을 크기 순으로 나열하여 중간값을 상기 대표 피치 값으로 결정하는 특징 파라미터 추출 장치.
- [청구항 6] 제4 항에 있어서,
 상기 적정 피치 검출부는,
 상기 주기적인 구간의 입력 신호에 대한 자기 상관 함수의 결과를 기반으로 선택된 피치 값들 중, 상기 대표 피치 값에 가장 근사한 값을 적정 피치 값으로 결정하는 특징 파라미터 추출 장치.
- [청구항 7] 제1 항에 있어서,
 상기 여기 신호 추출부는,
 상기 주기적인 구간의 입력 신호를 전처리하여, 상기 입력 신호의 생성

과정에서 손실된 고주파 성분이 보상된 전처리 신호를 출력하는
선강조부를 포함하는 특징 파라미터 추출 장치.

[청구항 8] 제7 항에 있어서,
상기 여기 신호 추출부는,
상기 전처리 신호에 대한 자기 상관 함수의 결과값을 출력하는 자기 상관
함수 출력부; 및
상기 자기 상관 함수 추정부의 출력을 입력받아, 레빈슨
더빈(Levinson-Durbin) 알고리즘을 기반으로 추측 계수를 출력하는 추측
계수 산출부; 및
상기 전처리 신호 및 상기 추측 계수를 기반으로 역 필터링(inverse
filtering)을 수행하여 상기 여기 신호를 출력하는 역 필터링부를 포함하는
특징 파라미터 추출 장치.

[청구항 9] 제4 항에 있어서,
상기 특징 파라미터 산출부는,
상기 여기 신호를 상기 적정 피치 값을 기반으로 이산 푸리에
변환(Discrete Fourier Transform)하여 주파수 영역의 이산 푸리에
스펙트럼으로 변환하는 주파수 영역 변환부를 포함하는 특징 파라미터
추출 장치.

[청구항 10] 제9 항에 있어서,
상기 특징 파라미터 산출부는,
상기 이산 푸리에 스펙트럼의 차수별 크기의 로그 값을 상기 특징
파라미터로 산출하는 특징 파라미터 추출 장치.

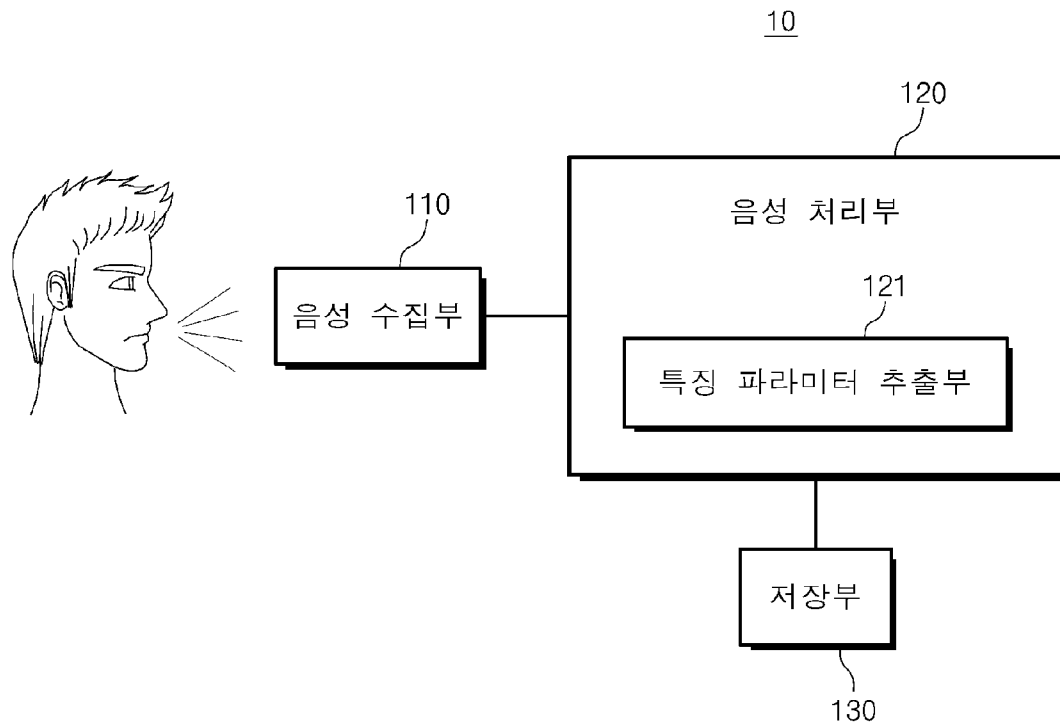
[청구항 11] 제9 항에 있어서,
상기 특징 파라미터 산출부는,
상기 이산 푸리에 스펙트럼의 차수별 에너지에 미리 결정된 가중치를
부여하여 합한 값을 특징 파라미터로 산출하는 특징 파라미터 추출 장치.

[청구항 12] 제1 항에 있어서,
상기 특징 파라미터 추출부는,
상기 입력 신호의 주기적인 구간 및 비주기적인 구간에 대하여, 상기 입력
신호의 주파수 응답 스펙트럼에 멜-주파수(Mel-Frequency) 필터를
적용하여 멜 주파수 응답을 획득하는 멜-주파수 응답 획득부; 및
상기 멜-주파수 응답을 역 이산 코사인 변환하여 캡스트럼 계수(cepstral
coefficient)를 획득하는 캡스트럼 계수 획득부를 포함하는 특징 파라미터
추출 장치.

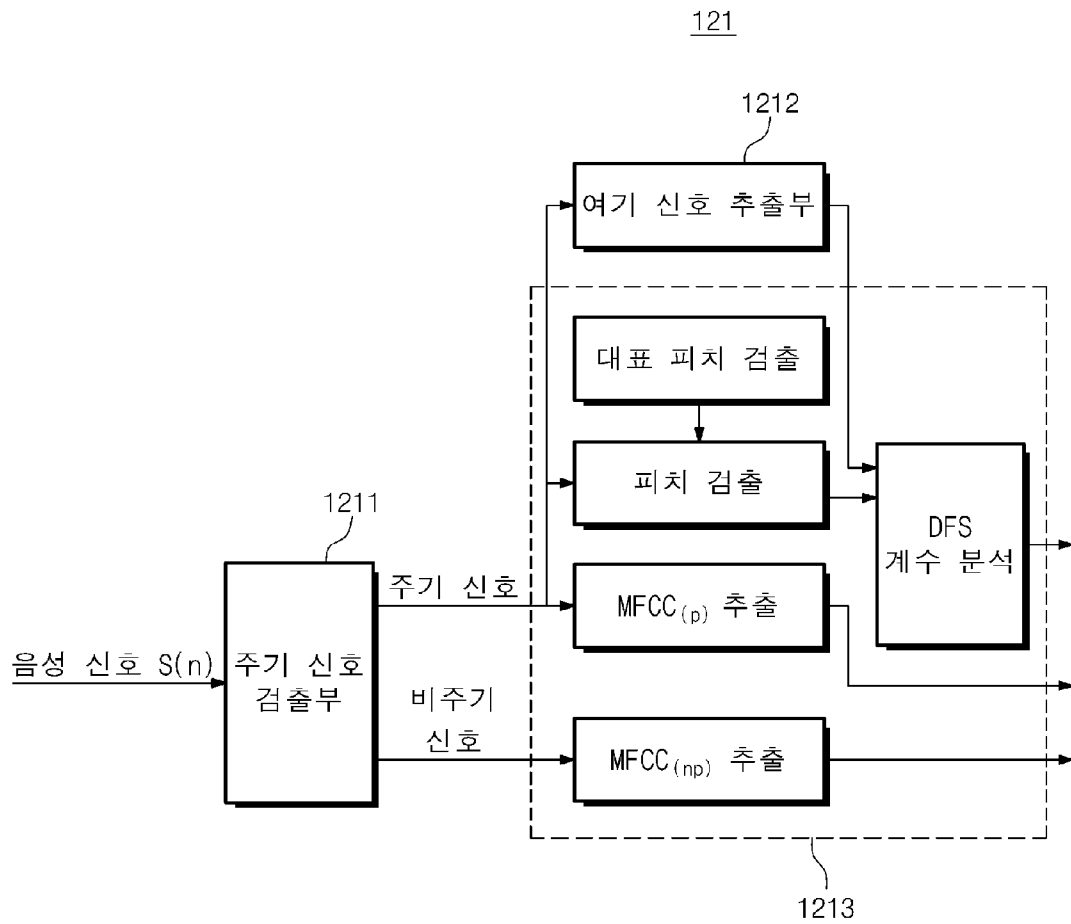
[청구항 13] 화자의 음성을 수집하는 음성 수집부;
상기 수집된 음성을 처리하여 기 등록된 사용자의 음성과 일치 여부를
판별하는 음성 처리부; 및
상기 사용자의 음성에 관한 정보를 저장하는 저장부를 포함하되,

상기 음성 처리부는:
입력 신호의 주기적인 구간을 검출하는 주기 신호 검출부;
상기 입력 신호의 주기적인 구간에서 여기 신호(excitation signal)를
추출하는 여기 신호 추출부; 및
상기 여기 신호의 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를
특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부를 포함하는
화자 인식 장치.

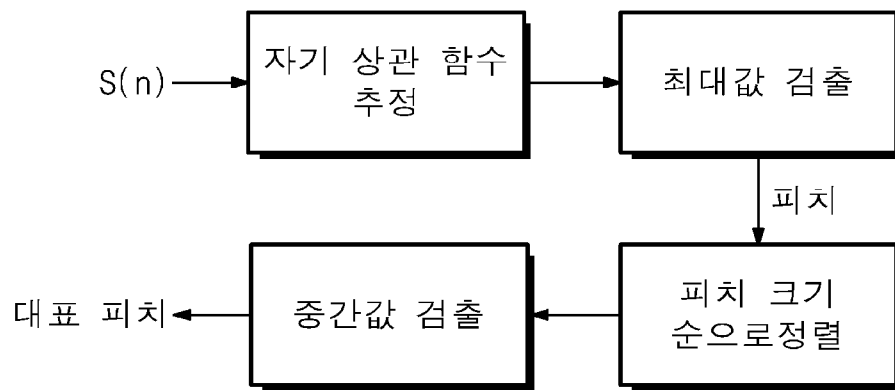
[도1]



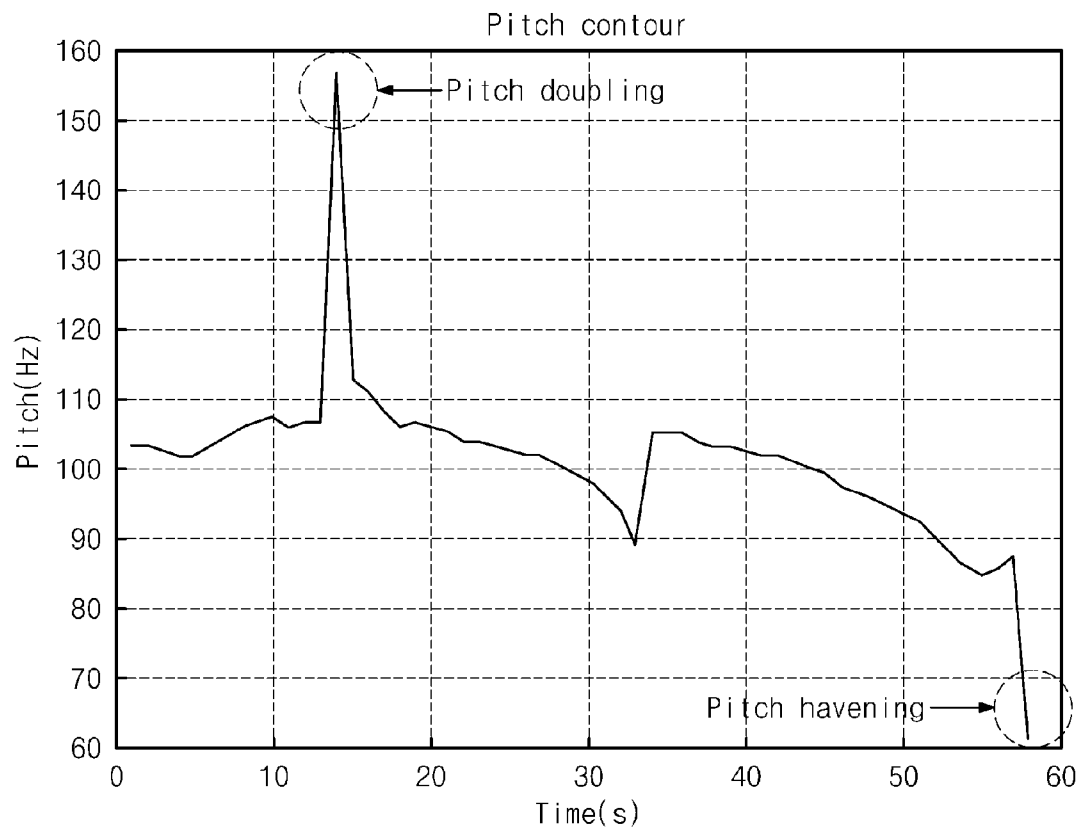
[도2]



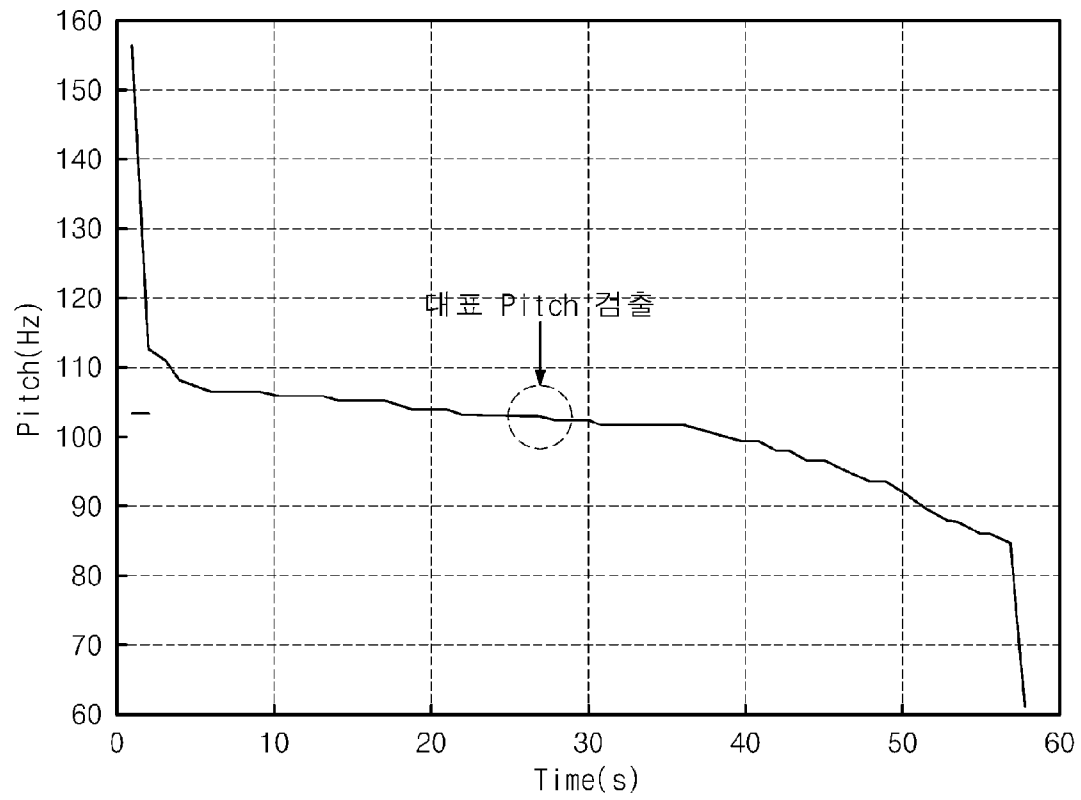
[도3]



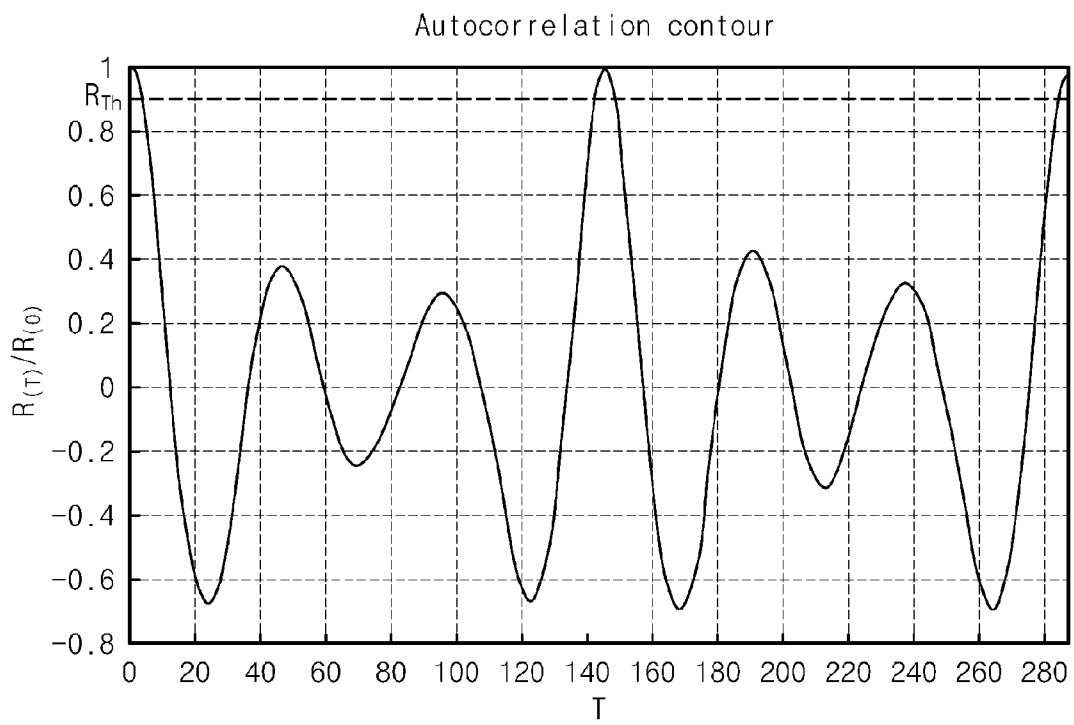
[도4]



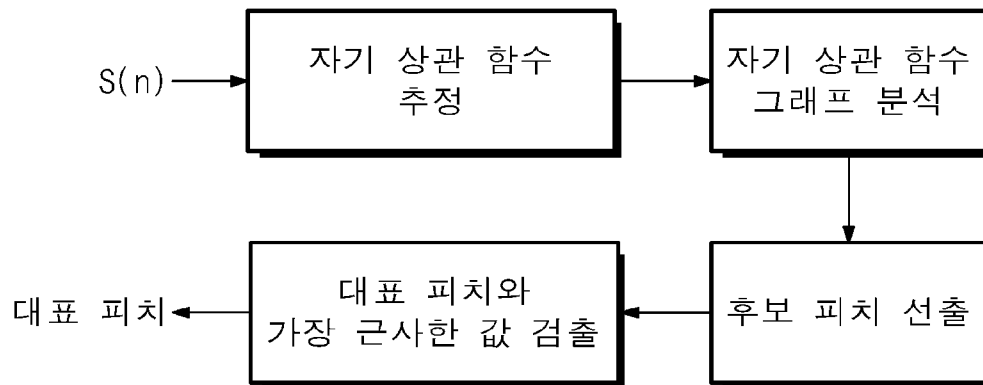
[도5]



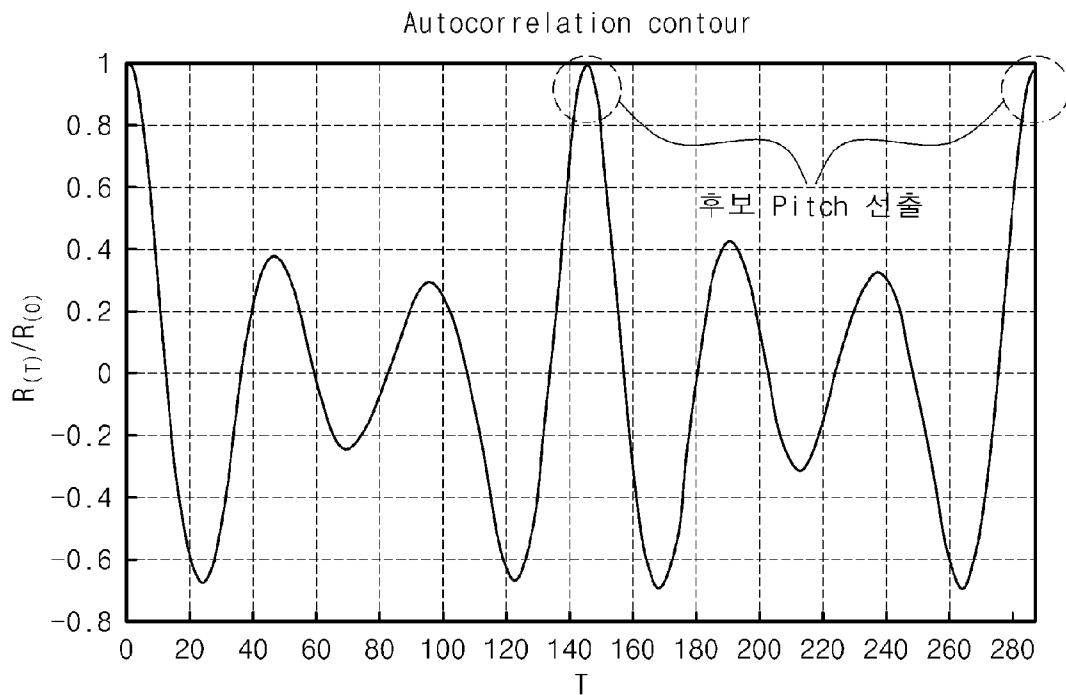
[도6]



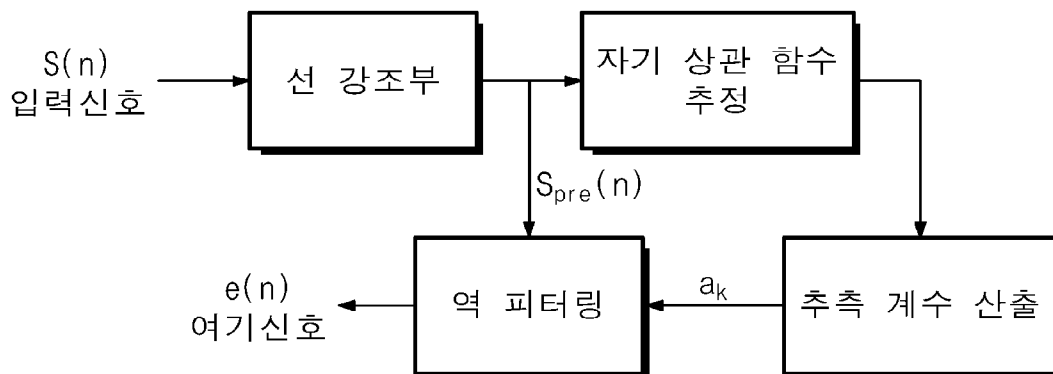
[도7]



[도8]



[도9]



[도10]

스코어 가중치			화자인식률 (%)	
α	β	γ	실시예	비교예
0.46	0.47	0.07	91.51	74.08
0.45	0.47	0.08	93.90	87.98
0.52	0.39	0.09	95.88	93.38
0.47	0.41	0.12	98.27	97.39
0.58	0.30	0.13	99.23	98.46

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2016/014673**A. CLASSIFICATION OF SUBJECT MATTER****G10L 17/02(2013.01)i, G06F 17/10(2006.01)i**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L 17/02; G10L 17/00; G10L 15/01; G10L 17/04; G10L 15/08; G10L 15/28; G06F 17/10

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Utility models and applications for Utility models: IPC as above

Japanese Utility models and applications for Utility models: IPC as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

eKOMPASS (KIPO internal) & Keywords: voice recognition, narrator recognition, parameter extraction, cycle, excitation signal, frequency response spectrum, pitch detection.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	KANG, Ji Hoon et al., "Speaker Recognition Performance Improvement by Voiced/Unvoiced Classification and Heterogeneous Feature Combination", Journal of the Korea Institute of Information and Communication Engineering, vol. 18, no. 6, pp. 1294-1301, June 2014 See page 1298 and figure 4.	1-3,7-8,12-13
A		4-6,9-11
Y	KR 10-2005-0048214 A (YONSEI UNIVERSITY) 24 May 2005 See claims 1-6 and figures 1-2.	1-3,7-8,12-13
A	KR 10-2010-0036893 A (SAMSUNG ELECTRONICS CO., LTD.) 08 April 2010 See paragraphs [0031]-[0040]; claims 1-5; and figures 1-3.	1-13
A	KR 10-1153129 B1 (MICROSOFT CORPORATION) 04 June 2012 See paragraphs [0071]-[0074] and claim 1.	1-13
A	KR 10-2009-0043203 A (INDUSTRY-ACADEMIC COOPERATION FOUNDATION, YONSEI UNIVERSITY) 06 May 2009 See paragraphs [0020]-[0042]; claims 1-3; and figure 3.	1-13



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

29 MARCH 2017 (29.03.2017)

Date of mailing of the international search report

29 MARCH 2017 (29.03.2017)

Name and mailing address of the ISA/KR

Korean Intellectual Property Office
Government Complex-Daejeon, 189 Seonsa-ro, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2016/014673

Patent document cited in search report	Publication date	Patent family member	Publication date
KR 10-2005-0048214 A	24/05/2005	KR 10-0526110 B1	08/11/2005
KR 10-2010-0036893 A	08/04/2010	US 2010-0082341 A1	01/04/2010
		US 2010-0083834 A1	08/04/2010
		US 8317905 B2	27/11/2012
		WO 2010-039271 A1	08/04/2010
KR 10-1153129 B1	04/06/2012	CN 1760972 A	19/04/2006
		EP 1647969 A1	19/04/2006
		JP 2006-113569 A	27/04/2006
		JP 4936696 B2	23/05/2012
		KR 10-2006-0050689 A	19/05/2006
		US 2006-0085187 A1	20/04/2006
		US 7684988 B2	23/03/2010
KR 10-2009-0043203 A	06/05/2009	KR 10-0933946 B1	28/12/2009

A. 발명이 속하는 기술분류(국제특허분류(IPC))

G10L 17/02(2013.01)i, G06F 17/10(2006.01)i

B. 조사된 분야

조사된 최소문헌(국제특허분류를 기재)

G10L 17/02; G10L 17/00; G10L 15/01; G10L 17/04; G10L 15/08; G10L 15/28; G06F 17/10

조사된 기술분야에 속하는 최소문헌 이외의 문헌

한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC

국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우))

eKOMPASS(특허청 내부 검색시스템) & 키워드: 음성 인식, 화자 인식, 파라미터 추출, 주기, 여기 신호, 주파수 응답 스펙트럼, 피치 검출.

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
Y	강지훈 등, “유/무성음 구분 및 이중적 특징파라미터 결합을 이용한 화자인식 성능 개선”, 한국정보통신학회논문지, Vol. 18, No. 6, pp.1294-1301, 2014.06 페이지 1298 및 도면 4 참조.	1-3, 7-8, 12-13
A		4-6, 9-11
Y	KR 10-2005-0048214 A (학교법인연세대학교) 2005.05.24 청구항 1-6 및 도면 1-2 참조.	1-3, 7-8, 12-13
A	KR 10-2010-0036893 A (삼성전자주식회사) 2010.04.08 단락 [0031]-[0040]; 청구항 1-5; 및 도면 1-3 참조.	1-13
A	KR 10-1153129 B1 (마이크로소프트 코퍼레이션) 2012.06.04 단락 [0071]-[0074] 및 청구항 1 참조.	1-13
A	KR 10-2009-0043203 A (연세대학교 산학협력단) 2009.05.06 단락 [0020]-[0042]; 청구항 1-3; 및 도면 3 참조.	1-13

☐ 추가 문헌이 C(계속)에 기재되어 있습니다.☒ 대응특허에 관한 별지를 참조하십시오.

* 인용된 문헌의 특별 카테고리:

“A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌

“T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌

“E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌

“X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다.

“L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌

“Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다.

“O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌

“&” 동일한 대응특허문헌에 속하는 문헌

“P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌

국제조사의 실제 완료일

2017년 03월 29일 (29.03.2017)

국제조사보고서 발송일

2017년 03월 29일 (29.03.2017)

ISA/KR의 명칭 및 우편주소

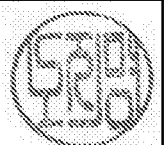
대한민국 특허청
(35208) 대전광역시 서구 청사로 189,
4동 (둔산동, 정부대전청사)

팩스 번호 +82-42-481-8578

심사관

노지명

전화번호 +82-42-481-8528



국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
KR 10-2005-0048214 A	2005/05/24	KR 10-0526110 B1	2005/11/08
KR 10-2010-0036893 A	2010/04/08	US 2010-0082341 A1	2010/04/01
		US 2010-0083834 A1	2010/04/08
		US 8317905 B2	2012/11/27
		WO 2010-039271 A1	2010/04/08
KR 10-1153129 B1	2012/06/04	CN 1760972 A	2006/04/19
		EP 1647969 A1	2006/04/19
		JP 2006-113569 A	2006/04/27
		JP 4936696 B2	2012/05/23
		KR 10-2006-0050689 A	2006/05/19
		US 2006-0085187 A1	2006/04/20
		US 7684988 B2	2010/03/23
KR 10-2009-0043203 A	2009/05/06	KR 10-0933946 B1	2009/12/28