



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년11월07일
(11) 등록번호 10-1673221
(24) 등록일자 2016년11월01일

(51) 국제특허분류(Int. Cl.)

G10L 17/02 (2013.01)

(52) CPC특허분류

G10L 17/02 (2013.01)

(21) 출원번호 10-2015-0183988

(22) 출원일자 2015년12월22일

심사청구일자 2015년12월22일

(56) 선행기술조사문헌

KR1019970012285 A*

JP2006189799 A*

강지훈, 정상배, ‘유/무성음 구분 및 이중적 특징파라미터 결합을 이용한 화자인식 성능 개선’, 한국정보통신학회논문지, Vol.18, No.6, pp.1294~1301, June 2014.*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

경상대학교 산학협력단

경상남도 진주시 진주대로 501 (가좌동)

(72) 발명자

정상배

경상남도 진주시 초창로14번길 27 초전1차푸르지 오아파트 109동 804호

김영일

경상남도 진주시 진주대로 859 (주약동, 한주럭키 아파트) 3-1201

강지훈

제주특별자치도 제주시 진남로2길 11 (화북일동, 대주아파트) 202호

(74) 대리인

권혁수, 송윤호

전체 청구항 수 : 총 12 항

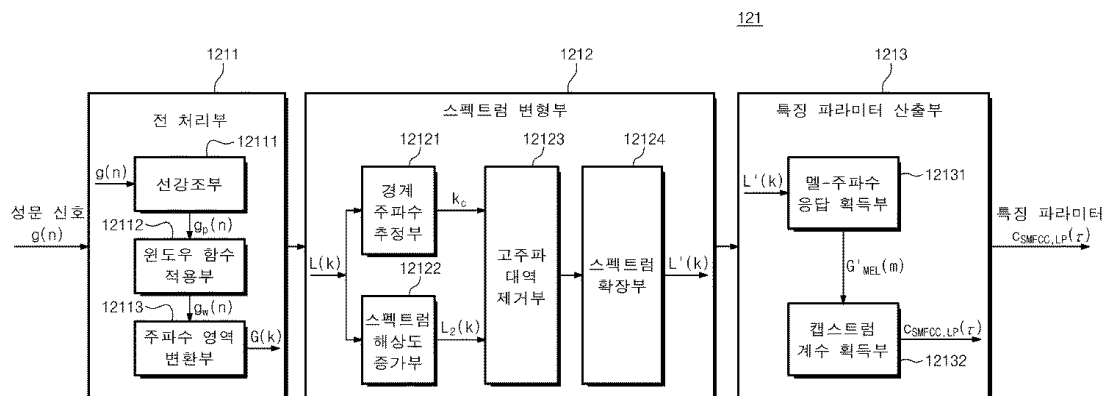
심사관 : 정성윤

(54) 발명의 명칭 화자 인식을 위한 성문과 특징 추출 장치

(57) 요약

본 발명은 입력 신호의 특징 파라미터 추출 장치 및 그를 이용한 화자 인식 장치에 관한 것이다. 본 발명의 일 실시예에 따른 특징 파라미터 추출 장치는, 입력 신호를 전 처리하는 전 처리부; 상기 입력 신호의 주파수 응답 스펙트럼에서 경계 주파수보다 높은 고주파 대역의 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형하는 스펙트럼 변형부; 및 상기 변형된 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부;를 포함할 수 있다.

대표도



이 발명을 지원한 국가연구개발사업

과제고유번호 1345233405

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 이공학개인지초연구지원

연구과제명 스마트 TV용 음성 인터페이스를 위한 원거리 음성개선 기술 개발(5차년도)

기 여 율 1/1

주관기관 경상대학교산학협력단

연구기간 2015.05.01 ~ 2016.04.30

명세서

청구범위

청구항 1

입력 신호를 전 처리하는 전 처리부;

상기 입력 신호의 주파수 응답 스펙트럼에서 상기 입력 신호를 기반으로 결정된 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형하는 스펙트럼 변형부; 및

상기 변형된 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부;

를 포함하고,

상기 스펙트럼 변형부는:

상기 입력 신호에 대한 상기 경계 주파수를 추정하는 경계 주파수 추정부; 및

상기 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수보다 높은 주파수 대역을 제거하는 고주파 대역 제거부;

를 포함하고,

상기 경계 주파수 추정부는:

상기 주파수 응답 스펙트럼의 로그 값을 지수 함수로 모델링하고,

상기 모델링한 지수 함수에서 기 설정된 임계 로그 값에 대응하는 주파수를 상기 경계 주파수로 결정하는 특징 파라미터 추출 장치.

청구항 2

제 1 항에 있어서,

상기 입력 신호는 음성 신호로부터 얻은 성문(glottal) 신호를 포함하는 특징 파라미터 추출 장치.

청구항 3

제 1 항에 있어서,

상기 전 처리부는:

상기 입력 신호의 생성 과정에서 손실된 고주파 성분을 보상하는 선강조부를 포함하는 특징 파라미터 추출 장치.

청구항 4

제 1 항에 있어서,

상기 전 처리부는:

상기 입력 신호에 기 설정된 윈도우 함수를 적용하는 윈도우 함수 적용부를 포함하는 특징 파라미터 추출 장치.

청구항 5

제 1 항에 있어서,

상기 전 처리부는:

시간 영역의 상기 입력 신호를 주파수 영역의 상기 주파수 응답 스펙트럼으로 변환하는 주파수 영역 변환부를 포함하는 특징 파라미터 추출 장치.

청구항 6

삭제

청구항 7

삭제

청구항 8

제 1 항에 있어서,

상기 경계 주파수 추정부는:

상기 주파수 응답 스펙트럼의 로그 값과 계수 및 지수에 의해 정의되는 지수 함수 모델 간의 차분에 기초한 비용이 최소가 되는 상기 계수 및 상기 지수를 산정하고,

상기 입력 로그 값을 상기 산정된 계수로 나눈 값의 로그 값을 상기 산정된 지수로 나눈 값을 상기 경계 주파수로 결정하는 특징 파라미터 추출 장치.

청구항 9

입력 신호를 전 처리하는 전 처리부;

상기 입력 신호의 주파수 응답 스펙트럼에서 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형하는 스펙트럼 변형부; 및

상기 변형된 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부를 포함하며,

상기 스펙트럼 변형부는:

상기 입력 신호에 대한 상기 경계 주파수를 추정하는 경계 주파수 추정부; 및

상기 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수보다 높은 주파수 대역을 제거하는 고주파 대역 제거부;

를 포함하고,

상기 스펙트럼 변형부는:

상기 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시키는 스펙트럼 해상도 증가부를 더 포함하며,

상기 고주파 대역 제거부는:

상기 주파수 영역 해상도가 증가된 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수보다 높은 주파수 대역을 제거하는 특징 파라미터 추출 장치.

청구항 10

제 9 항에 있어서,

상기 스펙트럼 해상도 증가부는:

상기 주파수 응답 스펙트럼을 주파수 영역에서 기 설정된 배수만큼 업-샘플링하는 특징 파라미터 추출 장치.

청구항 11

입력 신호를 전 처리하는 전 처리부;

상기 입력 신호의 주파수 응답 스펙트럼에서 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형하는 스펙트럼 변형부; 및

상기 변형된 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부를 포함하며,

상기 스펙트럼 변형부는:

상기 입력 신호에 대한 상기 경계 주파수를 추정하는 경계 주파수 추정부; 및

상기 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수보다 높은 주파수 대역을 제거하는 고주파 대역 제거부;

를 포함하고,

상기 스펙트럼 변형부는:

상기 추정된 경계 주파수보다 높은 주파수 대역이 제거된 주파수 응답 스펙트럼의 주파수 대역을 확장시키는 스펙트럼 확장부를 더 포함하는 특징 파라미터 추출 장치.

청구항 12

제 11 항에 있어서,

상기 스펙트럼 확장부는:

상기 추정된 경계 주파수보다 높은 주파수 대역이 제거된 주파수 응답 스펙트럼에 포함된 다수의 샘플 값을 기반으로 보간을 수행하는 특징 파라미터 추출 장치.

청구항 13

제 1 항에 있어서,

상기 특징 파라미터 산출부는:

상기 변형된 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 획득부;

를 포함하는 특징 파라미터 추출 장치.

청구항 14

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 신호를 전 처리하는 전 처리부;

상기 입력 신호의 주파수 응답 스펙트럼에서 상기 입력 신호를 기반으로 결정된 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형하는 스펙트럼 변형부; 및

상기 변형된 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부;

를 포함하며,

상기 스펙트럼 변형부는:

상기 입력 신호에 대한 상기 경계 주파수를 추정하는 경계 주파수 추정부; 및

상기 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수보다 높은 주파수 대역을 제거하는 고주파 대역 제거부;

를 포함하고,

상기 경계 주파수 추정부는:

상기 주파수 응답 스펙트럼의 로그 값을 지수 함수로 모델링하고,

상기 모델링한 지수 함수에서 기 설정된 임계 로그 값에 대응하는 주파수를 상기 경계 주파수로 결정하는 화자 인식 장치.

발명의 설명

기술 분야

[0001] 본 발명은 입력 신호의 특징 파라미터를 추출하는 장치 및 그를 이용하여 화자를 인식하는 장치에 관한 것이다.

배경 기술

[0002] 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.

[0003] 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 내용

해결하려는 과제

[0004] 본 발명의 실시예는 화자 인식률을 개선시킬 수 있는 특징 파라미터 추출 장치 및 그를 이용한 화자 인식 장치를 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0005] 본 발명의 일 실시예에 따른 특징 파라미터 추출 장치는, 입력 신호를 전 처리하는 전 처리부; 상기 입력 신호의 주파수 응답 스펙트럼에서 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형하는 스펙트럼 변형부; 및 상기 변형된 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부;를 포함할 수 있다.

[0006] 상기 입력 신호는 음성 신호로부터 얻은 성문(glottal) 신호를 포함할 수 있다.

[0007] 상기 전 처리부는: 상기 입력 신호의 생성 과정에서 손실된 고주파 성분을 보상하는 선강조부를 포함할 수 있다.

[0008] 상기 전 처리부는: 상기 입력 신호에 기 설정된 윈도우 함수를 적용하는 윈도우 함수 적용부를 포함할 수 있다.

[0009] 상기 전 처리부는: 시간 영역의 상기 입력 신호를 주파수 영역의 상기 주파수 응답 스펙트럼으로 변환하는 주파수 영역 변환부를 포함할 수 있다.

[0010] 상기 스펙트럼 변형부는: 상기 입력 신호에 대한 상기 경계 주파수를 추정하는 경계 주파수 추정부; 및 상기 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수보다 높은 주파수 대역을 제거하는 고주파 대역 제거부;를 포함할 수 있다.

[0011] 상기 경계 주파수 추정부는: 상기 주파수 응답 스펙트럼의 로그 값을 지수 함수로 모델링하고, 상기 모델링한 지수 함수에서 기 설정된 임계 로그 값에 대응하는 주파수를 상기 경계 주파수로 결정할 수 있다.

[0012] 상기 경계 주파수 추정부는: 상기 주파수 응답 스펙트럼의 로그 값과 계수 및 지수에 의해 정의되는 지수 함수 모델 간의 차분에 기초한 비용이 최소가 되는 상기 계수 및 상기 지수를 산정하고, 상기 임계 로그 값을 상기 산정된 계수로 나눈 값의 로그 값을 상기 산정된 지수로 나눈 값을 상기 경계 주파수로 결정할 수 있다.

[0013] 상기 스펙트럼 변형부는: 상기 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시키는 스펙트럼 해상도 증가부를 더 포함하며, 상기 고주파 대역 제거부는: 상기 주파수 영역 해상도가 증가된 주파수 응답 스펙트럼에서

상기 추정된 경계 주파수보다 높은 주파수 대역을 제거할 수 있다.

- [0014] 상기 스펙트럼 해상도 증가부는: 상기 주파수 응답 스펙트럼을 주파수 영역에서 기 설정된 배수만큼 업-샘플링할 수 있다.
- [0015] 상기 스펙트럼 변형부는: 상기 추정된 경계 주파수보다 높은 주파수 대역이 제거된 주파수 응답 스펙트럼의 주파수 대역을 확장시키는 스펙트럼 확장부를 더 포함할 수 있다.
- [0016] 상기 스펙트럼 확장부는: 상기 추정된 경계 주파수보다 높은 주파수 대역이 제거된 주파수 응답 스펙트럼에 포함된 다수의 샘플 값을 기반으로 보간을 수행할 수 있다.
- [0017] 상기 특징 파라미터 산출부는: 상기 변형된 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 획득부;를 포함할 수 있다.
- [0018] 본 발명의 일 실시예에 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 포함하되, 상기 음성 처리부는: 입력 신호를 전 처리하는 전 처리부; 상기 입력 신호의 주파수 응답 스펙트럼에서 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형하는 스펙트럼 변형부; 및 상기 변형된 주파수 응답 스펙트럼을 기반으로 상기 입력 신호를 특징짓는 특징 파라미터를 산출하는 특징 파라미터 산출부;를 포함할 수 있다.

발명의 효과

- [0019] 본 발명의 실시예에 따르면, 음성 처리에 있어서 음성 인식률을 높일 수 있으며, 특히 화자 인식률을 향상시킬 수 있다.

도면의 간단한 설명

- [0020] 도 1은 본 발명의 일 실시예에 따른 화자 인식 장치의 예시적인 블록도이다.
- 도 2는 본 발명의 일 실시예에 따른 특징 파라미터 추출부의 예시적인 블록도이다.
- 도 3은 본 발명의 일 실시예에 따라 경계 주파수를 기준으로 고주파 대역이 제거되는 입력 신호의 주파수 응답 스펙트럼 및 그 지수 함수 모델의 예시적인 그래프이다.
- 도 4는 본 발명의 일 실시예에 따라 지수 함수 모델의 계수 및 지수를 산정하는 과정을 설명하기 위한 예시적인 그래프이다.
- 도 5는 본 발명의 일 실시예에 따라 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시키는 과정을 설명하기 위한 예시적인 그래프이다.
- 도 6은 본 발명의 일 실시예에 따라 주파수 응답 스펙트럼의 주파수 대역을 확장시키는 과정을 설명하기 위한 예시적인 그래프이다.
- 도 7은 본 발명의 일 실시예에 따른 특징 파라미터 추출 방법의 예시적인 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0021] 이하, 본 명세서에 첨부된 도면을 참조하여 본 발명의 실시예를 상세하게 설명한다.
- [0022] 도 1은 본 발명의 일 실시예에 따른 화자 인식 장치(10)의 예시적인 블록도이다.
- [0023] 도 1에 도시된 바와 같이, 상기 화자 인식 장치(10)는 음성 수집부(110), 음성 처리부(120) 및 저장부(130)를 포함한다.
- [0024] 상기 음성 수집부(110)는 화자의 음성을 수집한다. 본 발명의 일 실시예에 따르면, 상기 음성 수집부(110)는 화자가 발생한 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.
- [0025] 상기 음성 처리부(120)는 수집된 음성을 처리하여 기 등록된 사용자의 음성과 일치 여부를 판별한다. 본 발명

의 일 실시예에 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.

[0026] 상기 저장부(130)는 사용자의 음성에 관한 정보를 저장한다. 본 발명의 일 실시예에 따르면, 상기 저장부(130)는 데이터 또는 각종 프로그램을 저장할 수 있는 저장 장치로서, 일 예로 HDD, SSD 등과 같은 대용량의 저장 장치뿐만 아니라 RAM, ROM 등의 각종 메모리, 레지스터, 캐쉬 등을 포함할 수 있다.

[0027] 본 발명의 실시예에 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.

[0028] 도 2는 본 발명의 일 실시예에 따른 특징 파라미터 추출부(121)의 예시적인 블록도이다.

[0029] 도 2를 참조하면, 상기 특징 파라미터 추출부(121)는 전 처리부(1211), 스펙트럼 변형부(1212) 및 특징 파라미터 산출부(1213)를 포함할 수 있다.

[0030] 상기 전 처리부(1211)는 입력 신호를 전 처리한다. 상기 스펙트럼 변형부(1212)는 입력 신호의 주파수 응답 스펙트럼에서 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼을 변형한다. 상기 특징 파라미터 산출부(1213)는 변형된 주파수 응답 스펙트럼을 기반으로 입력 신호를 특징짓는 특징 파라미터를 산출한다.

[0031] 본 발명의 실시예에 따르면, 상기 특징 파라미터 추출부(121)에 의해 처리되어 특징 파라미터가 획득되는 입력 신호는 음성 신호로부터 얻은 성문(glottal) 신호를 포함할 수 있다. 본 발명의 실시예와 같이 성문 신호를 처리하여 특징 파라미터를 얻는 경우, 특징 파라미터에 화자의 형태학적 특성을 반영할 수 있어 인식률 향상에 기여할 수 있다.

[0032] 도 2를 참조하면, 상기 전 처리부(1211)는 선강조부(12111), 윈도우 함수 적용부(12112) 및 주파수 영역 변환부(12113)를 포함할 수 있다.

[0033] 상기 선강조부(12111)는 입력 신호(예컨대, 도 2에서 성문 신호 $g(n)$)의 생성 과정에서 손실된 고주파 성분을 보상한다. 화자가 음성을 발생할 때 입술에서 자유 공간으로 공기를 방사하는 과정에서 고주파 성분이 손실된다. 상기 선강조부(12111)는 입력 신호에 선강조 필터를 적용하여 손실된 고주파 성분을 보상한다. 일 예로, 상기 선강조 필터는 아래의 수학적식과 같은 전달함수를 갖도록 구현될 수 있다.

수학적식 1

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

[0034]

[0035] 상기 윈도우 함수 적용부(12112)는 입력 신호에 기 설정된 윈도우 함수를 적용한다. 상기 윈도우 함수 적용부(12112)는 입력 신호의 프레임 간 불연속성으로 인한 왜곡을 최소화하기 위한 것으로, 입력 신호에 소정의 윈도우 함수를 적용함으로써 단구간 데이터의 활용에 따른 주파수 응답의 해상도 저하를 완화시킨다. 일 예로, 상기 윈도우 함수 적용부(12112)는 아래의 수학적식과 같이 입력 신호 $g_p(n)$ 에 윈도우 함수 $w(n)$ 를 적용하여 출력 신호 $g_w(n)$ 를 생성할 수 있다.

수학식 2

$$g_w(n) = g_p(n) \cdot w(n), (0 \leq n < N_w)$$

[0036]

[0037]

여기서, N_w 는 단구간 입력 신호의 프레임 크기이다.

[0038]

본 발명의 일 실시예에 따르면, 상기 윈도우 함수 적용부(12112)는 입력 신호 $g_p(n)$ 에 해밍 윈도우 함수를 적용할 수 있다. 해밍 윈도우 함수는 아래의 수학적식과 같이 나타낼 수 있다.

수학식 3

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N_w - 1}\right), (0 \leq n < N_w)$$

[0039]

[0040]

상기 주파수 영역 변환부(12113)는 시간 영역의 입력 신호를 주파수 영역의 주파수 응답 스펙트럼으로 변환한다. 일 예로, 상기 주파수 영역 변환부(12113)는 아래의 수학적식과 같이 시간 영역의 입력 신호 $g_w(n)$ 를 이산 푸리에 변환(Discrete Fourier Transform, DFT)하여 주파수 영역의 주파수 응답 스펙트럼 $G(k)$ 으로 변환할 수 있다.

수학식 4

$$G(k) = \sum_{n=0}^{N-1} g_w(n) e^{-j \frac{2\pi}{N} kn}, (0 \leq k < N)$$

[0041]

[0042]

여기서, N 은 이산 푸리에 변환의 크기이다.

[0043]

앞서 설명한 바와 같이, 상기 전 처리부(1211)가 입력 신호 $g(n)$ 를 전 처리하여 출력 신호 $G(k)$ 를 출력하면, 상기 스펙트럼 변형부(1212)는 입력 신호의 주파수 응답 스펙트럼 $G(k)$ 에서 경계 주파수보다 높은 고주파 평탄 대역을 제거하여 상기 주파수 응답 스펙트럼 $G(k)$ 을 변형한다.

[0044]

도 2를 참조하면, 상기 스펙트럼 변형부(1212)는 경계 주파수 추정부(12121) 및 고주파 대역 제거부(12123)를 포함할 수 있다.

[0045]

상기 경계 주파수 추정부(12121)는 입력 신호에 대한 경계 주파수를 추정한다. 상기 고주파 대역 제거부(12123)는 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수보다 높은 주파수 대역을 제거한다.

[0046]

도 3은 본 발명의 일 실시예에 따라 경계 주파수 k_c 를 기준으로 고주파 대역이 제거되는 입력 신호의 주파수 응답 스펙트럼 및 그 지수 함수 모델의 예시적인 그래프이다.

[0047]

먼저, 상기 스펙트럼 변형부(1212)는 입력 신호의 주파수 응답 스펙트럼 $G(k)$ 을 변형하기 전에 상기 주파수 응답 스펙트럼 $G(k)$ 의 로그 값 $L(k) = \log |G(k)|$ 을 취할 수 있다.

[0048]

그리고, 상기 경계 주파수 추정부(12121)는 주파수 응답 스펙트럼의 로그 값 $L(k)$ 을 지수 함수로 모델링하고, 모델링한 지수 함수에서 기 설정된 임계 로그 값 L_{TH} 에 대응하는 주파수를 경계 주파수 k_c 로 결정할 수 있다.

[0049] 예를 들어, 도 3을 참조하면, 상기 경계 주파수 추정부(12121)는 주파수 인덱스 k 에 걸쳐 요동하는 주파수 응답 스펙트럼의 로그 값 $L(k)$ 을 지수 함수 $Ae^{-\alpha k}$ 로 모델링할 수 있다.

[0050] 본 발명의 일 실시예에 따르면, 상기 경계 주파수 추정부(12121)는 주파수 응답 스펙트럼의 로그 값 $L(k)$ 과 계수 A 및 지수 α 에 의해 정의되는 지수 함수 모델 $Ae^{-\alpha k}$ 간의 차분에 기초한 비용이 최소가 되는 계수 A_{opt} 및 지수 α_{opt} 를 산정할 수 있다. 일 예로, 주파수 응답 스펙트럼의 로그 값 $L(k)$ 과 지수 함수 모델 $Ae^{-\alpha k}$ 간의 차분에 기초한 비용을 계산하기 위한 비용 함수는 아래의 수학적식과 같다.

수학적식 5

$$J = \sum_{k=0}^{N-1} \left(L(k) - Ae^{-\alpha k} \right)^2$$

[0051]

[0052] 도 4는 본 발명의 일 실시예에 따라 지수 함수 모델 $Ae^{-\alpha k}$ 의 계수 A 및 지수 α 를 산정하는 과정을 설명하기 위한 예시적인 그래프이다.

[0053] 본 발명의 일 실시예에 따르면, 상기 경계 주파수 추정부(12121)는 최급강하법(steepest descent method)을 이용하여 비용 함수 J 로부터 최소 비용을 갖는 계수 A_{opt} 및 지수 α_{opt} 를 산정할 수 있다.

[0054] 예를 들어, 도 4를 참조하면, 상기 경계 주파수 추정부(12121)는 계수 A 와 지수 α 를 변수로 갖는 비용 함수 J 에서 비용 함수의 기울기가 가장 큰 방향으로 계수 A 및 지수 α 를 변경해 가면서 비용이 최소가 되는 계수 A_{opt} 및 지수 α_{opt} 를 구할 수 있다.

[0055] 그리고 나서, 상기 경계 주파수 추정부(12121)는 계수 A_{opt} 및 지수 α_{opt} 로 모델링된 지수 함수 $A_{opt}e^{-\alpha_{opt}k}$ 에서 기 설정된 임계 로그 값 L_{TH} 에 대응하는 주파수를 경계 주파수 k_c 로 결정할 수 있다.

[0056] 본 발명의 실시예에 따르면, 상기 경계 주파수 추정부(12121)는 임계 로그 값 L_{TH} 을 상기 산정된 계수 A_{opt} 로 나눈 뒤, 그 로그 값을 취하여 상기 산정된 지수 α_{opt} 로 나누어 경계 주파수 k_c 를 산출할 수 있다. 다시 말해, 경계 주파수 k_c 는 아래의 수학적식과 같이 계산될 수 있다.

수학적식 6

$$k_c = \frac{1}{\alpha_{opt}} \ln\left(\frac{L_{TH}}{A_{opt}}\right)$$

[0057]

[0058] 그리고 나서, 다시 도 3을 참조하면, 상기 고주파 대역 제거부(12123)는 주파수 응답 스펙트럼의 로그 값 $L(k)$ 에서 경계 주파수 k_c 보다 높은 주파수 대역을 제거하고 경계 주파수 k_c 보다 낮은 주파수 대역은 유지시킬 수 있다.

[0059] 본 발명자는 단구간 성문 신호 $g(n)$ 의 주파수 응답에서 고주파 대역의 다소 평탄한 부분이 화자 인식에 있어서 식별력 향상에 크게 도움이 되지 않는 것을 발견하였다. 따라서, 본 발명의 실시예는 주파수 응답 스펙트럼의 로그 값 $L(k)$ 에서 저주파의 경사진 부분과 고주파의 평탄한 부분을 구분하기 위한 경계 주파수 k_c 를 추정하고, 경계 주파수 k_c 를 기준으로 그보다 높은 주파수 대역을 제거하고 그보다 낮은 주파수 대역을 기반으로 입력 신호

의 특징 파라미터를 추출하는 기술을 제안한다. 이와 같이 본 발명의 실시예에 따라 추출된 특징 파라미터를 이용하여 화자 인식을 수행하는 경우, 화자 인식률이 크게 개선될 수 있다.

- [0060] 나아가, 다시 도 2를 참조하면, 상기 스펙트럼 변형부(1212)는 스펙트럼 해상도 증가부(12122)를 더 포함할 수 있다. 상기 스펙트럼 해상도 증가부(12122)는 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시킨다.
- [0061] 앞서 설명한 바와 같이, 경계 주파수 k_c 를 기준으로 그보다 낮은 저주파 대역만을 이용하여 특징 파라미터를 추출하는 경우, 스펙트럼의 주파수 영역 해상도가 낮아져 입력 신호를 바람직하게 특징짓는 특징 파라미터를 추출하기 어려울 수 있다.
- [0062] 따라서, 이 실시예는 상기 스펙트럼 해상도 증가부(12122)를 더 포함함으로써 주파수 응답 스펙트럼에서 고주파 대역을 제거하기 전에 주파수 영역 해상도를 기 설정된 배수만큼 증가시킬 수 있다.
- [0063] 도 5는 본 발명의 일 실시예에 따라 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시키는 과정을 설명하기 위한 예시적인 그래프이다.
- [0064] 본 발명의 일 실시예에 따르면, 상기 스펙트럼 해상도 증가부(12122)는 주파수 응답 스펙트럼을 주파수 영역에서 기 설정된 배수만큼 업-샘플링할 수 있다.
- [0065] 예를 들어, 도 5를 참조하면, 상기 스펙트럼 해상도 증가부(12122)는 로그 주파수 응답 스펙트럼 $L(k)$ 의 응답과 응답 사이에 0을 삽입하고, 저역 통과 필터를 적용하여 주파수 영역의 해상도가 2 배만큼 증가된 주파수 응답 스펙트럼 $L_2(k)$ 를 얻을 수 있다.
- [0066] 이 실시예와 같이 상기 스펙트럼 변형부(1212)에 상기 스펙트럼 해상도 증가부(12122)가 더 포함되는 경우, 상기 고주파 대역 제거부(12123)는 주파수 영역 해상도가 증가된 주파수 응답 스펙트럼 $L_2(k)$ 에서 경계 주파수 k_c 보다 높은 주파수 대역을 제거할 수 있다.
- [0067] 더 나아가, 다시 도 2를 참조하면, 상기 스펙트럼 변형부(1212)는 스펙트럼 확장부(12124)를 더 포함할 수 있다. 상기 스펙트럼 확장부(12124)는 경계 주파수 k_c 보다 높은 주파수 대역이 제거된 주파수 응답 스펙트럼의 주파수 대역을 확장시킨다.
- [0068] 앞서 설명한 바와 같이, 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시킨 뒤 경계 주파수 k_c 를 기준으로 고주파 대역을 제거하여 스펙트럼을 변형시키는 경우, 변형된 스펙트럼은 정수가 아닌 주파수 인덱스 k 에서 주파수 응답값을 가질 수 있다.
- [0069] 따라서, 이 실시예는 상기 스펙트럼 확장부(12124)를 더 포함함으로써 상기 스펙트럼 변형부(1212)에 의해 변형된 주파수 응답 스펙트럼이 최종적으로 정수의 주파수 인덱스 k 에서 주파수 응답값을 가지도록 할 수 있다.
- [0070] 도 6은 본 발명의 일 실시예에 따라 주파수 응답 스펙트럼의 주파수 대역을 확장시키는 과정을 설명하기 위한 예시적인 그래프이다.
- [0071] 본 발명의 일 실시예에 따르면, 상기 스펙트럼 확장부(12124)는 경계 주파수 k_c 를 기준으로 고주파 대역이 제거된 주파수 응답 스펙트럼에 포함된 다수의 샘플 값을 기반으로 보간을 수행할 수 있다.
- [0072] 예를 들어, 도 6에 도시된 바와 같이, 상기 스펙트럼 확장부(12124)는 두 점을 기반으로 아래의 수학적식과 같이 선형 보간을 수행할 수 있다.

수학식 7

$$L'(k)=L_2(k_0)+(k-k_0)(L_2(k_1)-L_2(k_0)),(k=0,1,...,N/2)$$

[0073]

[0074] 위 수학적식에서 사용되는 k_0 및 k_1 은 아래의 수학적식과 같이 정의될 수 있다.

수학식 8

$$k_0 = \lfloor \frac{k_{c,2}}{N/2} k \rfloor, k_1 = k_0 + 1$$

[0075]

[0076] 여기서, $k_{c,2} = 2k_c$ 이다.

[0077] 앞서 설명한 바와 같이, 상기 스펙트럼 변형부(1212)가 입력 신호의 주파수 응답 스펙트럼 $G(k)$ 에서 경계 주파수 k_c 보다 높은 고주파 평탄 대역을 제거하여 주파수 응답 스펙트럼을 변형하면, 상기 특징 파라미터 산출부(1213)는 변형된 주파수 응답 스펙트럼 $L'(k)$ 을 기반으로 입력 신호 $g(n)$ 를 특징짓는 특징 파라미터를 산출한다.

[0078] 다시 도 2를 참조하면, 상기 특징 파라미터 산출부(1213)는 멜-주파수(Mel-Frequency) 응답 획득부(12131) 및 캡스트럼 계수 획득부(12132)를 포함할 수 있다.

[0079] 상기 멜-주파수 응답 획득부(12131)는 변형된 주파수 응답 스펙트럼에 멜-주파수 필터를 적용하여 멜-주파수 응답을 획득한다. 본 발명의 실시예에 따르면, 상기 멜-주파수 응답 획득부(12131)는 멜-주파수 필터를 적용하기 전에 상기 스펙트럼 변형부(1212)에 의해 변형된 주파수 응답 스펙트럼 $L'(k)$ 의 지수 값 $G'(k) = \exp(L'(k))$ 을 취한다.

[0080] 그 뒤, 상기 멜-주파수 응답 획득부(12131)는 아래의 수학식과 같이 변형된 주파수 응답 스펙트럼의 지수 값의 절대값 $|G'(k)|$ 에 기 설정된 멜-스케일 필터 बैं크를 적용할 수 있다.

수학식 9

$$G'_{MEL}(m) = \sum_{k=k_1(m)}^{k_2(m)} FB^{(m)}(k) |G'(k)|, (1 \leq m \leq M)$$

[0081]

[0082] 여기서, $G'_{MEL}(m)$ 은 m 번째 멜-주파수 응답이고, $FB^{(m)}(k)$ 는 m 번째 멜-주파수 필터 बैं크의 k 번째 응답이고, $k_1(m)$ 및 $k_2(m)$ 은 각각 m 번째 멜-주파수 필터 बैं크의 시작 주파수 인덱스 및 끝 주파수 인덱스이다.

[0083] 그리고 나서, 상기 캡스트럼 계수 획득부(12132)는 멜-주파수 응답 $G'_{MEL}(m)$ 을 역 이산 코사인 변환(Inverse Discrete Cosine Transform, IDCT)하여 캡스트럼 계수(cepstral coefficient)를 획득한다. 본 발명의 실시예에 따르면, 상기 캡스트럼 계수 획득부(12132)는 역 이산 코사인 변환을 하기 전에 멜-주파수 응답 $G'_{MEL}(m)$ 의 로그 값 $\log(G'_{MEL}(m))$ 을 취한다.

[0084] 그 뒤, 상기 캡스트럼 계수 획득부(12132)는 아래의 수학식과 같이 멜-주파수 응답의 로그 값 $\log(G'_{MEL}(m))$ 을 역 이산 코사인 변환하여 캡스트럼을 얻고 캡스트럼의 크기를 구하여 캡스트럼 계수 $c_{SMFCC,LP}(\tau)$ 를 산출할 수 있다.

수학식 10

$$c_{SMFCC,LP}(\tau) = \sum_{m=1}^M \log(G'_{MEL}(m)) \cos\left(\frac{\pi\tau(m-0.5)}{M}\right), 1 \leq \tau \leq D$$

[0085]

[0086]

여기서, M은 멜-필터 뱅크의 개수이고, D는 캡스트럼의 차수이다.

[0087]

이와 같이 산출된 캡스트럼 계수 $c_{SMFCC,LP}(\tau)$ 는 상기 화자 인식 장치(10)에서 입력 신호의 특징 파라미터로 사용될 수 있다.

[0088]

본 발명의 일 실시예에 따르면, 상기 화자 인식 장치(10)는 기 결정된 확률 분포 모델을 기반으로 한 기계 학습 알고리즘을 통해 입력 신호의 특징 파라미터 $c_{SMFCC,LP}(\tau)$ 를 이용하여 화자를 인식할 수 있다. 일 예로, 화자 인식에 사용되는 확률 분포 모델은 가우시안 혼합 모델(Gaussian Mixture Model, GMM)일 수 있으나 이에 제한되지는 않는다.

[0089]

또한, 본 발명의 일 실시예에 따르면, 상기 화자 인식 장치(10)는 앞서 산출된 특징 파라미터 $c_{SMFCC,LP}(\tau)$ 외에 화자 인식을 위해 사용되는 각종 특징 파라미터를 조합하여 화자 인식을 실시할 수 있다.

[0090]

도 7은 본 발명의 일 실시예에 따른 특징 파라미터 추출 방법(20)의 예시적인 흐름도이다.

[0091]

상기 특징 파라미터 추출 방법(20)은 앞서 설명한 본 발명의 실시예에 따른 특징 파라미터 추출부(121)에 의해 실행될 수 있다.

[0092]

도 7을 참조하면, 상기 특징 파라미터 추출 방법(20)은 입력 신호 $g(n)$ 를 전 처리하는 단계(S210), 상기 입력 신호의 주파수 응답 스펙트럼 $G(k)$ 에서 경계 주파수 k_c 보다 높은 고주파 평탄 대역을 제거하여 주파수 응답 스펙트럼을 변형하는 단계(S220), 및 변형된 주파수 응답 스펙트럼 $L'(k)$ 을 기반으로 입력 신호 $g(n)$ 를 특징짓는 특징 파라미터 $c_{SMFCC,LP}(\tau)$ 를 산출하는 단계(S230)를 포함한다.

[0093]

상기 입력 신호 $g(n)$ 를 전 처리하는 단계(S210)는 입력 신호 $g(n)$ 의 생성 과정에서 손실된 고주파 성분을 보상하는 선강조 단계를 포함할 수 있다.

[0094]

상기 입력 신호 $g(n)$ 를 전 처리하는 단계(S210)는 입력 신호 $g(n)$ 에 기 설정된 윈도우 함수 $w(n)$ 를 적용하는 단계를 포함할 수 있다.

[0095]

상기 입력 신호 $g(n)$ 를 전 처리하는 단계(S210)는 시간 영역의 입력 신호 $g(n)$ 를 주파수 영역의 주파수 응답 스펙트럼 $G(k)$ 으로 변환하는 단계를 포함할 수 있다.

[0096]

상기 주파수 응답 스펙트럼을 변형하는 단계(S220)는 입력 신호에 대한 경계 주파수 k_c 를 추정하는 단계, 및 주파수 응답 스펙트럼에서 상기 추정된 경계 주파수 k_c 보다 높은 주파수 대역을 제거하는 단계를 포함할 수 있다.

[0097]

본 발명의 일 실시예에 따르면, 상기 경계 주파수 k_c 를 추정하는 단계는 주파수 응답 스펙트럼의 로그 값 $L(k)$ 을 지수 함수 Ae^{-ak} 로 모델링하는 단계, 및 모델링한 지수 함수 $A_{opt}e^{-a_{opt}k}$ 에서 기 설정된 임계 로그 값 L_{TH} 에 대응하는 주파수를 경계 주파수 k_c 로 결정하는 단계를 포함할 수 있다.

[0098]

예를 들어, 상기 경계 주파수 k_c 를 추정하는 단계는 주파수 응답 스펙트럼의 로그 값 $L(k)$ 과 계수 A 및 지수 a에 의해 정의되는 지수 함수 모델 Ae^{-ak} 간의 차분에 기초한 비용이 최소가 되는 계수 A_{opt} 및 지수 a_{opt} 를 산정하는 단계, 및 상기 임계 로그 값 L_{TH} 을 상기 산정된 계수 A_{opt} 로 나눈 값의 로그 값을 상기 산정된 지수 a_{opt} 로 나눈 값을 경계 주파수 k_c 로 결정하는 단계를 포함할 수 있다.

[0099]

나아가, 상기 주파수 응답 스펙트럼을 변형하는 단계(S220)는 주파수 응답 스펙트럼에서 경계 주파수 k_c 보다 높은 주파수 대역을 제거하는 단계 전에, 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시키는 단계를 더 포

함할 수 있다. 이 경우, 상기 경계 주파수 k_c 보다 높은 주파수 대역을 제거하는 단계는 주파수 영역 해상도가 증가된 주파수 응답 스펙트럼 $L_2(k)$ 에서 경계 주파수 k_c 보다 높은 주파수 대역을 제거하는 단계를 포함할 수 있다.

[0100] 여기서, 상기 주파수 응답 스펙트럼의 주파수 영역 해상도를 증가시키는 단계는 주파수 응답 스펙트럼을 주파수 영역에서 기 설정된 배수만큼 업-샘플링하는 단계를 포함할 수 있다.

[0101] 더 나아가, 상기 주파수 응답 스펙트럼을 변형하는 단계(S220)는 경계 주파수 k_c 보다 높은 주파수 대역에 제거된 주파수 응답 스펙트럼의 주파수 대역을 확장시키는 단계를 더 포함할 수 있다.

[0102] 여기서, 상기 주파수 대역을 확장시키는 단계는 경계 주파수 k_c 보다 높은 주파수 대역이 제거된 주파수 응답 스펙트럼에 포함된 다수의 샘플 값을 기반으로 보간을 수행하는 단계를 포함할 수 있다.

[0103] 상기 특징 파라미터를 산출하는 단계(S230)는 변형된 주파수 응답 스펙트럼 $L'(k)$ 에 멜-주파수 필터를 적용하여 멜-주파수 응답 $G'_{MEL}(m)$ 을 획득하는 단계, 및 상기 멜-주파수 응답 $G'_{MEL}(m)$ 을 역 이산 코사인 변환하여 캡스트럼 계수 $c_{SMFCC,LP}(\tau)$ 를 획득하는 단계를 포함할 수 있다.

[0104] 상기 특징 파라미터 추출 방법(20)은 컴퓨터에서 실행되기 위한 프로그램으로 제작되어 컴퓨터가 읽을 수 있는 기록매체에 저장될 수 있다. 상기 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 저장장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있다. 또한, 상기 특징 파라미터 추출 방법(20)은 컴퓨터와 결합되어 실행시키기 위하여 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.

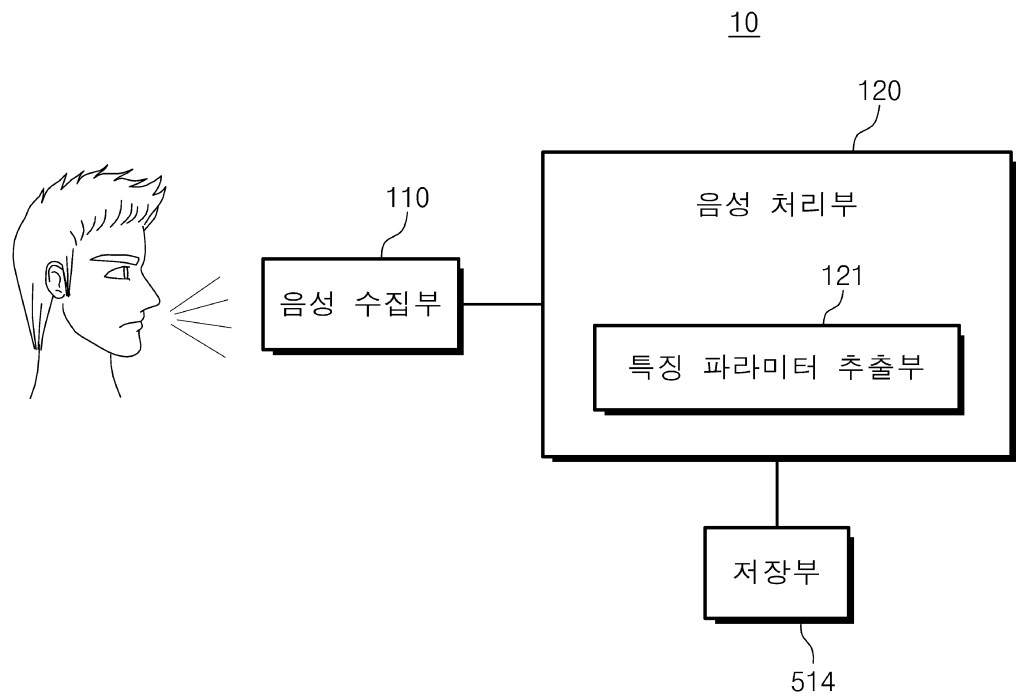
[0105] 이상에서 실시예를 통해 본 발명을 설명하였으나, 위 실시예는 단지 본 발명의 사상을 설명하기 위한 것으로 이에 한정되지 않는다. 통상의 기술자는 전술한 실시예에 다양한 변형이 가해질 수 있음을 이해할 것이다. 본 발명의 범위는 첨부된 특허청구범위의 해석을 통해서만 정해진다.

부호의 설명

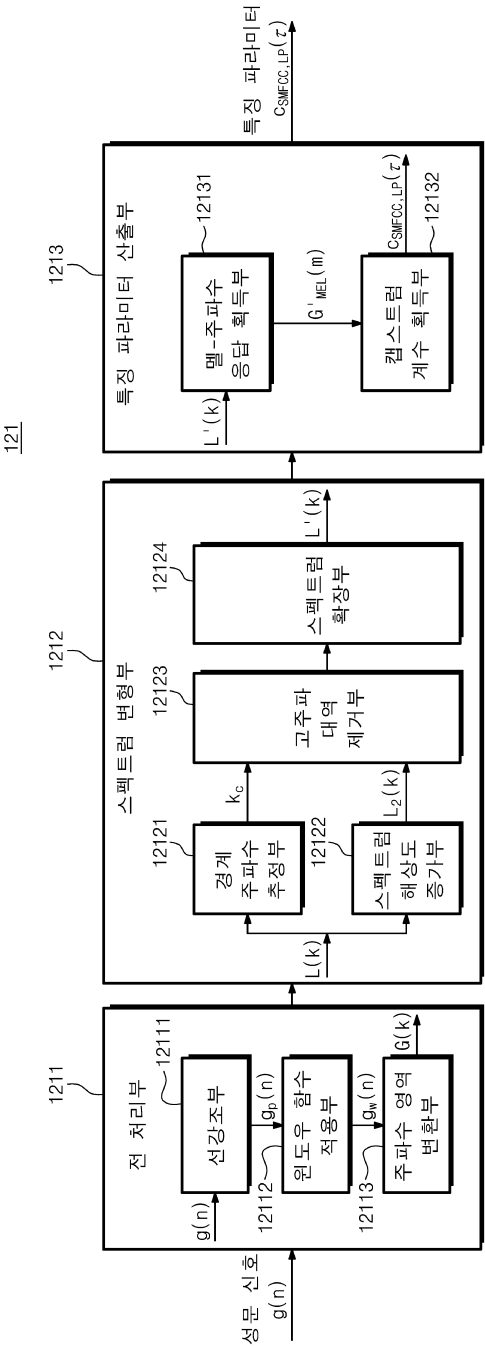
- [0106] 10: 화자 인식 장치
- 110: 음성 수집부
- 120: 음성 처리부
- 121: 특징 파라미터 추출부
- 130: 저장부
- 1211: 전 처리부
- 1212: 스펙트럼 변형부
- 1213: 특징 파라미터 산출부
- 12111: 선강조부
- 12112: 윈도우 함수 적용부
- 12113: 주파수 영역 변환부
- 12121: 경계 주파수 추정부
- 12122: 스펙트럼 해상도 증가부
- 12123: 고주파 대역 제거부
- 12124: 스펙트럼 확장부
- 12131: 멜-주파수 응답 획득부
- 12132: 캡스트럼 계수 획득부

도면

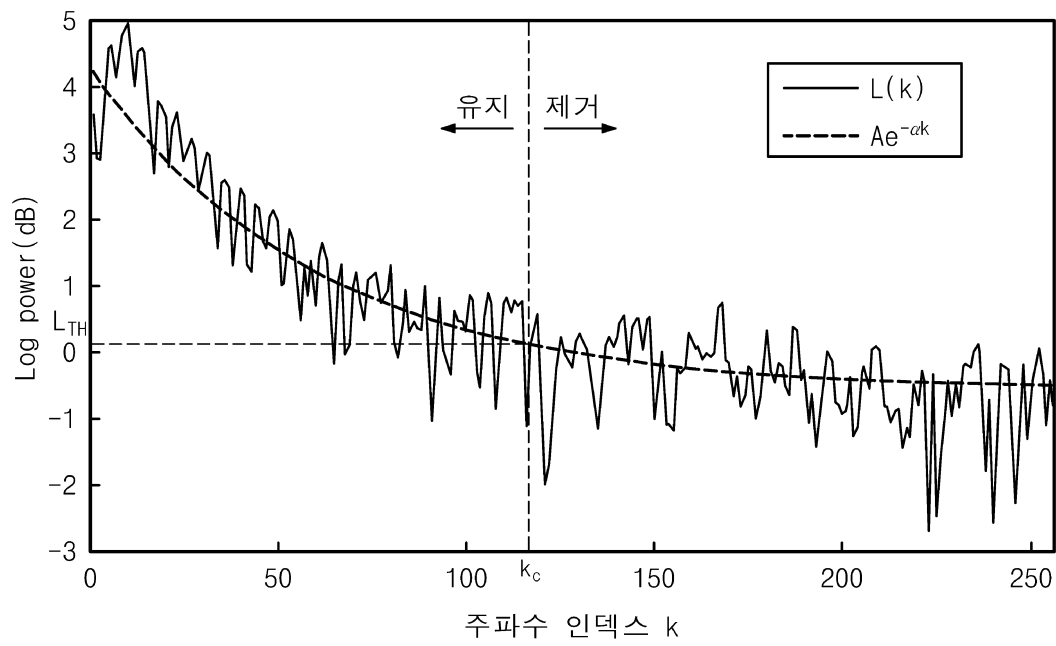
도면1



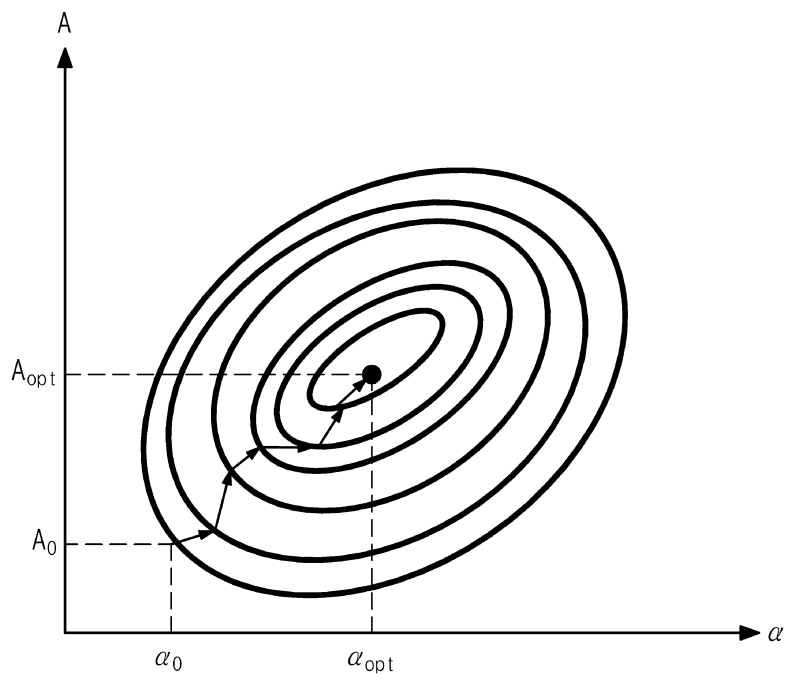
도면2



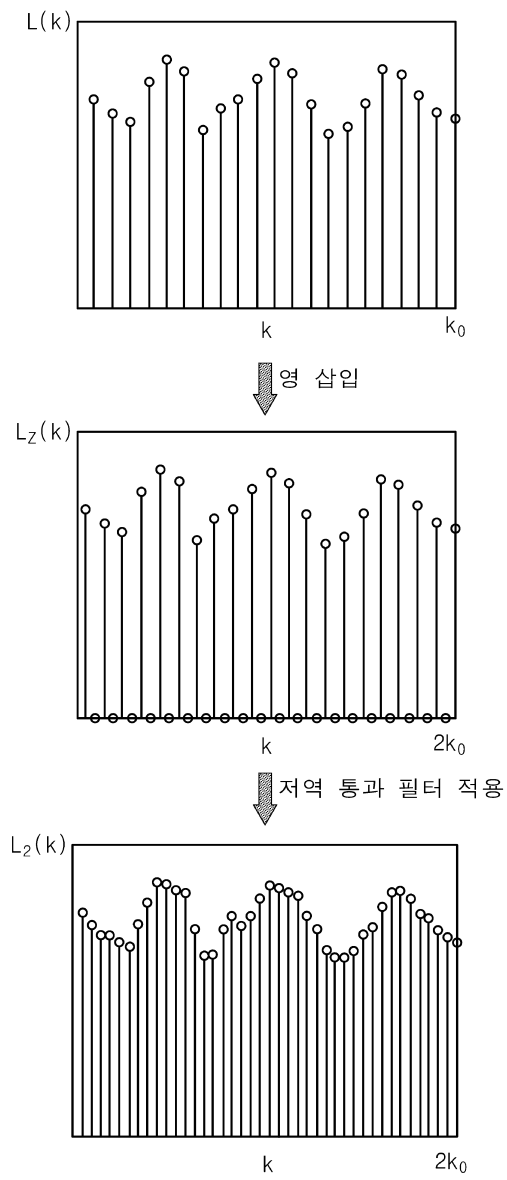
도면3



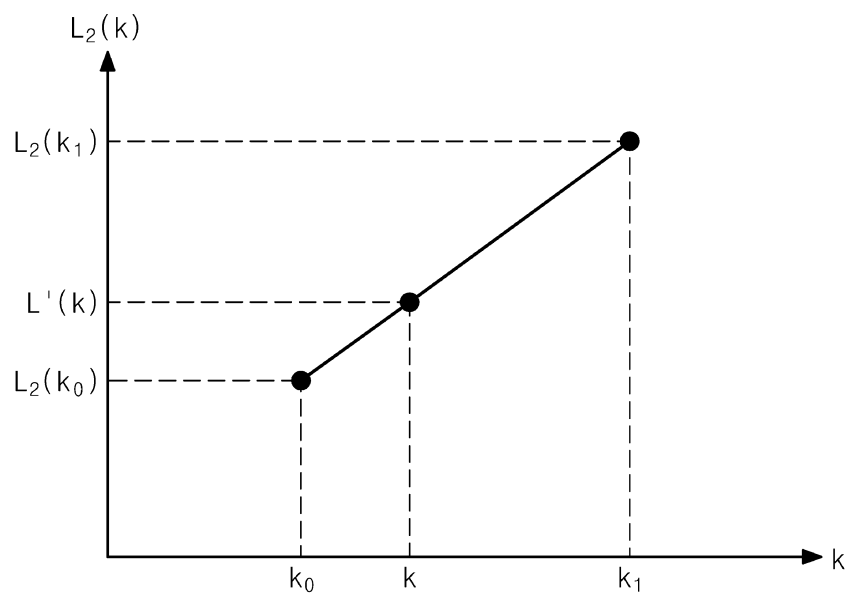
도면4



도면5



도면6



도면7

