



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0032390
(43) 공개일자 2015년03월26일

(51) 국제특허분류(Int. Cl.)

G10L 25/93 (2013.01)

(21) 출원번호 10-2013-0111424

(22) 출원일자 2013년09월16일

심사청구일자 없음

(71) 출원인

삼성전자주식회사

경기도 수원시 영통구 삼성로 129 (매탄동)

연세대학교 원주산학협력단

강원도 원주시 흥업면 연세대길 1

(72) 발명자

손준일

경기도 용인시 수지구 죽전2동 현인마을e-편한세
상아파트 202동 1204번지

구윤서

서울 서초구 잠원로14길 42, 329동 607호 (잠원동, 신반포아파트)

(뒷면에 계속)

(74) 대리인

특허법인 무한

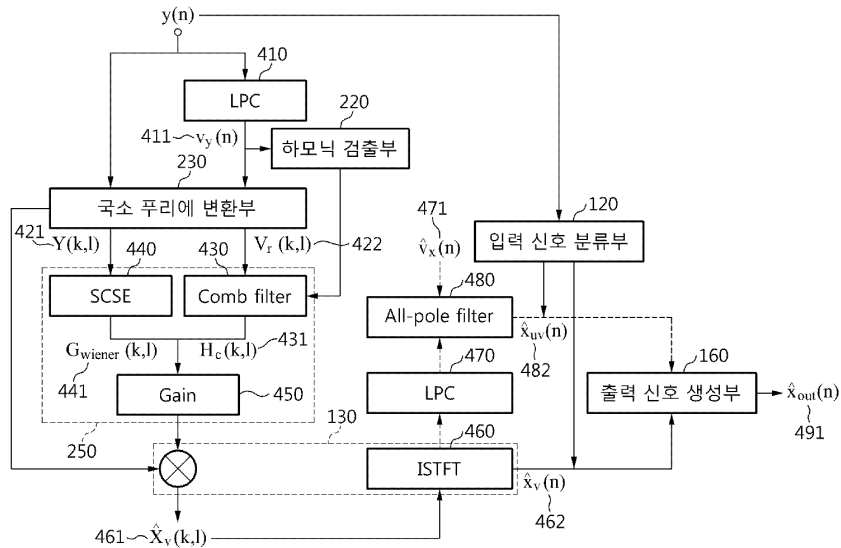
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 음성 명료도 향상을 위한 음성 신호 처리 장치 및 방법

(57) 요약

음성 명료도 향상을 위한 음성 신호 처리 장치 및 방법이 개시된다. 음성 신호 처리 장치는 유성음의 하모닉 특성에 기초한 콤(comb) 필터를 이용하여 입력 신호의 이득(gain)을 결정하는 입력 신호 이득 결정부; 상기 입력 신호에 상기 이득을 적용하여 하모닉 성분을 보존한 유성음을 출력하는 유성음 출력부; 상기 유성음에 기초하여 선형 예측 계수를 결정하는 선형 예측 계수 결정부; 및 상기 선형 예측 계수에 기초한 올 폴 필터(all-pole filter)를 이용하여 상기 입력 신호의 무성음을 보존하는 무성음 보존부를 포함할 수 있다.

대표도



(72) 발명자

김동욱

서울 서초구 동광로46길 13-9, 401호 (반포동, 그
린파크)

박영철

강원 원주시 관부면 시청로 264, 105동 401호 (원
주더샵아파트)

특허청구의 범위

청구항 1

유성음의 하모닉 특성에 기초한 콦(comb) 필터를 이용하여 입력 신호의 이득(gain)을 결정하는 입력 신호 이득 결정부;

상기 입력 신호에 상기 이득을 적용하여 하모닉 성분을 보존한 유성음을 출력하는 유성음 출력부;

상기 유성음에 기초하여 선형 예측 계수를 결정하는 선형 예측 계수 결정부; 및

상기 선형 예측 계수에 기초한 올 폴 필터(all-pole filter)를 이용하여 상기 입력 신호의 무성음을 보존하는 무성음 보존부

를 포함하는 음성 신호 처리 장치.

청구항 2

제1항에 있어서,

상기 입력 신호 이득 결정부는,

선형 예측기를 이용하여 상기 입력 신호의 잔여 신호를 결정하는 잔여 신호 결정부;

상기 잔여 신호의 스펙트럼 도메인에서 하모닉 성분을 검출하는 하모닉 검출부;

검출한 하모닉 성분에 기초하여 콦(comb) 필터를 설계하는 콦 필터 설계부; 및

위너 필터를 이용하여 상기 입력 신호를 필터링한 결과 및 상기 콦 필터를 이용하여 상기 입력 신호를 필터링한 결과에 기초하여 상기 이득을 결정하는 이득 결정부

를 포함하는 음성 신호 처리 장치.

청구항 3

제2항에 있어서,

상기 하모닉 검출부는,

상기 잔여 신호의 스펙트럼 도메인에서, 상기 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 추정하는 잔여 스펙트럼 추정부;

최고점 검출 알고리즘을 이용하여 추정한 잔여 스펙트럼에서 복수의 최고점들을 검출하는 최고점 검출부; 및

검출한 최고점들 간의 간격에 기초하여 하모닉 성분을 검출하는 하모닉 성분 검출부

를 포함하는 음성 신호 처리 장치.

청구항 4

제1항에 있어서,

상기 콦 필터는,

일정한 간격으로 스파이크가 반복되는 주파수 응답(frequency response)을 가지는 함수인 음성 신호 처리 장치.

청구항 5

제1항에 있어서,

상기 유성음 출력부는,

상기 입력 신호에 상기 이득을 적용하여 중간 출력 신호를 생성하고, 중간 출력 신호를 역 국소 푸리에 변환(ISTFT: Inverse Short-time Fourier transform)하거나, 역 고속 푸리에 변환(IFTT: Inverse fast Fourier

transform)하여 상기 유성음을 출력하는 음성 신호 처리 장치.

청구항 6

제1항에 있어서,

상기 선형 예측 계수 결정부는,

상기 유성음을 계수들의 선형 조합과 잔여 신호로 구분하고, 계수들의 선형 조합에 기초하여 선형 예측 계수를 결정하는 음성 신호 처리 장치.

청구항 7

제1항에 있어서,

상기 올 폴 필터는,

상기 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 상기 올 폴 필터에 입력되는 여기 신호 정보로 이용하는 음성 신호 처리 장치.

청구항 8

제1항에 있어서,

상기 유성음과 보존한 무성음을 기초로 음성 출력 신호를 생성하는 출력 신호 생성부

를 더 포함하는 음성 신호 처리 장치.

청구항 9

제8항에 있어서,

상기 출력 신호 생성부는,

상기 입력 신호의 역치 값이 임계값 미만인 구간은 상기 유성음에 기초하여 음성 출력 신호를 생성하고, 상기 입력 신호의 역치 값이 임계값 이상인 구간은 상기 보존한 무성음에 기초하여 음성 출력 신호를 생성하는 음성 신호 처리 장치.

청구항 10

유성음의 하모닉 특성에 기초한 콦(comb) 필터를 이용하여 입력 신호의 이득(gain)을 결정하는 단계;

상기 입력 신호에 상기 이득을 적용하여 하모닉 성분을 보존한 유성음을 출력하는 단계;

상기 유성음에 기초하여 선형 예측 계수를 결정하는 단계; 및

상기 선형 예측 계수에 기초한 올 폴 필터(all-pole filter)를 이용하여 상기 입력 신호의 무성음을 보존하는 단계

를 포함하는 음성 신호 처리 방법.

청구항 11

제10항에 있어서,

상기 입력 신호의 이득을 결정하는 단계는,

선형 예측기를 이용하여 상기 입력 신호의 잔여 신호를 결정하는 단계;

상기 잔여 신호의 스펙트럼 도메인에서 하모닉 성분을 검출하는 단계;

검출한 하모닉 성분에 기초하여 콦(comb) 필터를 설계하는 단계; 및

위너 필터를 이용하여 상기 입력 신호를 필터링한 결과 및 상기 콦 필터를 이용하여 상기 입력 신호를 필터링한 결과에 기초하여 상기 이득을 결정하는 단계

를 포함하는 음성 신호 처리 방법.

청구항 12

제11항에 있어서,

상기 하모닉 성분을 검출하는 단계는,

상기 잔여 신호의 스펙트럼 도메인에서, 상기 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 추정하는 단계;

최고점 검출 알고리즘을 이용하여 추정한 잔여 스펙트럼에서 복수의 최고점들을 검출하는 단계; 및

검출한 최고점들 간의 간격에 기초하여 하모닉 성분을 검출하는 단계

를 포함하는 음성 신호 처리 방법.

청구항 13

제10항에 있어서,

상기 콤 필터는,

일정한 간격으로 스파이크가 반복되는 주파수 응답(frequency response)을 가지는 함수인 음성 신호 처리 방법.

청구항 14

제10항에 있어서,

상기 유성음을 출력하는 단계는,

상기 입력 신호에 상기 이득을 적용하여 중간 출력 신호를 생성하고, 중간 출력 신호를 역 국소 푸리에 변환(ISTFT: Inverse Short-time Fourier transform)하거나, 역 고속 푸리에 변환(IFTT: Inverse fast Fourier transform)하여 상기 유성음을 출력하는 음성 신호 처리 방법.

청구항 15

제10항에 있어서,

상기 선형 예측 계수를 결정하는 단계는

상기 유성음을 계수들의 선형 조합과 잔여 신호로 구분하고, 계수들의 선형 조합에 기초하여 선형 예측 계수를 결정하는 음성 신호 처리 방법.

청구항 16

제10항에 있어서,

상기 올 폴 필터는,

상기 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 상기 올 폴 필터에 입력되는 여기 신호 정보로 이용하는 음성 신호 처리 방법.

청구항 17

제10항에 있어서,

상기 유성음과 보존한 무성음을 기초로 음성 출력 신호를 생성하는 단계

를 더 포함하는 음성 신호 처리 방법.

청구항 18

제17항에 있어서,

상기 음성 출력 신호를 생성하는 단계는,

상기 입력 신호의 역치 값이 임계값 미만인 구간은 상기 유성음에 기초하여 음성 출력 신호를 생성하고, 상기 입력 신호의 역치 값이 임계값 이상인 구간은 상기 보존한 무성음에 기초하여 음성 출력 신호를 생성하는 음성 신호 처리 방법.

청구항 19

제10항 내지 제18항 중 어느 한 항의 방법을 실행하기 위한 프로그램이 기록된 컴퓨터에서 판독할 수 있는 기록 매체.

명세서

기술분야

[0001] 이하의 일실시예들은 음성 명료도 향상을 위한 음성 신호 처리 장치 및 방법에 관한 것이다.

배경기술

[0002] 음질 개선 알고리즘은 보청기, 또는 음성 신호를 재현하는 오디오의 출력 신호의 품질을 높이기 위해서 사용되고 있다.

[0003] 그러나, 종래의 배경 잡음 추정 기반의 음질 개선 알고리즘들은 이득 값을 형성하는 조건에서 잔여 배경잡음의 크기와 음성의 왜곡 사이에 trade-off 관계가 존재할 수 있다. 따라서, 입력 신호에서 배경 잡음을 많이 제거할수록 음성의 왜곡이 심해져서 음성의 명료도가 낮아질 수 있다.

[0004] 따라서, 배경 잡음을 제거하면서도 음성 왜곡을 최소화하여 출력 신호의 음성의 명료도를 높이는 방법이 요구된다.

발명의 내용

과제의 해결 수단

[0005] 일실시예에 따른 음성 신호 처리 장치는 유성음의 하모닉 특성에 기초한 콤(comb) 필터를 이용하여 입력 신호의 이득(gain)을 결정하는 입력 신호 이득 결정부; 상기 입력 신호에 상기 이득을 적용하여 하모닉 성분을 보존한 유성음을 출력하는 유성음 출력부; 상기 유성음에 기초하여 선형 예측 계수를 결정하는 선형 예측 계수 결정부; 및 상기 선형 예측 계수에 기초한 올 폴 필터(all-pole filter)를 이용하여 상기 입력 신호의 무성음을 보존하는 무성음 보존부를 포함할 수 있다.

[0006] 일실시예에 따른 음성 신호 처리 장치의 입력 신호 이득 결정부는 선형 예측기를 이용하여 상기 입력 신호의 잔여 신호를 결정하는 잔여 신호 결정부; 상기 잔여 신호의 스펙트럼 도메인에서 하모닉 성분을 검출하는 하모닉 검출부; 검출한 하모닉 성분에 기초하여 콤(comb) 필터를 설계하는 콤 필터 설계부; 및 위너 필터를 이용하여 상기 입력 신호를 필터링한 결과 및 상기 콤 필터를 이용하여 상기 입력 신호를 필터링한 결과에 기초하여 상기 이득을 결정하는 이득 결정부를 포함할 수 있다.

[0007] 일실시예에 따른 음성 신호 처리 장치의 하모닉 검출부는 상기 잔여 신호의 스펙트럼 도메인에서, 상기 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 추정하는 잔여 스펙트럼 추정부; 최고점 검출 알고리즘을 이용하여 추정된 잔여 스펙트럼에서 복수의 최고점들을 검출하는 최고점 검출부; 및 검출한 최고점들 간의 간격에 기초하여 하모닉 성분을 검출하는 하모닉 성분 검출부를 포함할 수 있다.

[0008] 일실시예에 따른 음성 신호 처리 장치의 콤 필터는 일정한 간격으로 스파이크가 반복되는 주파수 응답(frequency response)을 가지는 함수일 수 있다.

[0009] 일실시예에 따른 음성 신호 처리 장치의 유성음 출력부는 상기 입력 신호에 상기 이득을 적용하여 중간 출력 신호를 생성하고, 중간 출력 신호를 역 국소 푸리에 변환(ISTFT: Inverse Short-time Fourier transform)하거나, 역 고속 푸리에 변환(IFFT: Inverse fast Fourier transform)하여 상기 유성음을 출력할 수 있다.

[0010] 일실시예에 따른 음성 신호 처리 장치의 선형 예측 계수 결정부는 상기 유성음을 계수들의 선형 조합과 잔여 신호로 구분하고, 계수들의 선형 조합에 기초하여 선형 예측 계수를 결정할 수 있다.

- [0011] 일실시예에 따른 음성 신호 처리 장치의 올 폴 필터는 상기 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 상기 올 폴 필터에 입력되는 여기 신호 정보로 이용할 수 있다.
- [0012] 일실시예에 따른 음성 신호 처리 장치는 상기 유성음과 보존한 무성음을 기초로 음성 출력 신호를 생성하는 출력 신호 생성부를 더 포함할 수 있다.
- [0013] 일실시예에 따른 음성 신호 처리 장치의 출력 신호 생성부는 상기 입력 신호의 역치 값이 임계값 미만인 구간은 상기 유성음에 기초하여 음성 출력 신호를 생성하고, 상기 입력 신호의 역치 값이 임계값 이상인 구간은 상기 보존한 무성음에 기초하여 음성 출력 신호를 생성할 수 있다.
- [0014] 일실시예에 따른 음성 신호 처리 방법은 유성음의 하모닉 특성에 기초한 콤(comb) 필터를 이용하여 입력 신호의 이득(gain)을 결정하는 단계; 상기 입력 신호에 상기 이득을 적용하여 하모닉 성분을 보존한 유성음을 출력하는 단계; 상기 유성음에 기초하여 선형 예측 계수를 결정하는 단계; 및 상기 선형 예측 계수에 기초한 올 폴 필터(all-pole filter)를 이용하여 상기 입력 신호의 무성음을 보존하는 단계를 포함할 수 있다.

도면의 간단한 설명

- [0015] 도 1은 일실시예에 따른 음성 신호 처리 장치의 구조를 도시한 도면이다.
- 도 2는 일실시예에 따른 입력 신호 이득 결정부의 구조를 도시한 도면이다.
- 도 3은 일실시예에 따른 하모닉 검출부의 구조를 도시한 도면이다.
- 도 4는 일실시예에 따른 음성 신호 처리 과정에서 정보들의 흐름을 도시한 도면이다.
- 도 5는 일실시예에 따른 하모닉 검출 결과의 일례이다.
- 도 6은 일실시예에 따른 콤 필터로 필터링한 결과인 콤 필터 이득의 일례이다.
- 도 7은 일실시예에 따른 음성 신호 처리 방법을 도시한 도면이다.
- 도 8은 일실시예에 따른 입력 신호 이득 결정 과정을 도시한 도면이다.
- 도 9는 일실시예에 따른 하모닉 검출 과정을 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0016] 이하, 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다.
- [0017] 도 1은 일실시예에 따른 음성 신호 처리 장치의 구조를 도시한 도면이다.
- [0018] 도 1을 참고하면, 음성 신호 처리 장치(100)는 입력 신호 이득 결정부(110), 입력 신호 분류부(120), 유성음 출력부(130), 선형 예측 계수 결정부(140), 무성음 보존부(150), 및 출력 신호 생성부(160)를 포함할 수 있다.
- [0019] 음성 신호 처리 장치(100)는 난청자들의 청력을 보상하기 위한 난청 보상 장치에 포함될 수 있다. 그리고, 음성 신호 처리 장치(100)는 난청 보상 장치의 마이크가 수집한 음성 신호를 처리할 수 있다.
- [0020] 또한, 음성 신호 처리 장치(100)는 음성 신호를 재현하는 오디오 장치에 포함될 수도 있다.
- [0021] 입력 신호 이득 결정부(110)는 유성음의 하모닉 특성에 기초한 콤(comb) 필터를 이용하여 입력 신호의 이득(gain)을 결정할 수 있다. 이때, 콤 필터는 일정한 간격으로 스파이크가 반복되는 주파수 응답(frequency response)을 가지는 함수일 수 있다.
- [0022] 입력 신호 이득 결정부(110)의 상세 구성 및 동작은 이하 도 2를 참조하여 상세히 설명한다.
- [0023] 입력 신호 분류부(120)는 입력 신호를 유성음과 무성음으로 분류할 수 있다.
- [0024] 예를 들어, 입력 신호 분류부(120)는 유무성음 판별기와 VAD(voiced activity detector)를 이용하여 입력 신호의 현재 프레임이 잡음 구간인지 여부를 판단하고, 잡음 구간이 아닌 경우, 현재 프레임에 포함된 음성을 유성음, 또는 무성음으로 분류할 수 있다.
- [0025] 입력 신호는 수학식 1로 나타낼 수 있다.

수학식 1

$$y(n) = x(n) + w(n)$$

이때, $y(n)$ 는 잡음과 음성이 섞여 있는 입력 신호이고, $x(n)$ 는 목적 음성 신호이며, $w(n)$ 는 잡음 신호일 수 있다.

또한, 입력 신호는 선형 예측을 통해 선형 계수들의 조합과 잔여 신호 $v_y(n)$ 로 나누어 질 수 있다. 자기 상관 함수 계산에 선형 계수들을 이용할 경우, 현재 프레임에서의 음성의 피치를 계산할 수 있다.

그리고, 잔여 신호는 STFT(Short-time Fourier transform)을 통해 잔여 스펙트럼 도메인으로 변환될 수 있다. 이때, 입력 신호 분류부(120)가 수학식 2와 같이 입력 스펙트럼 $Y(k, l)$ 과 잔여 신호 스펙트럼 $V_y(k, l)$ 의 비율인 $\gamma(k, l)$ 을 dB 스케일로 그러면 스펙트럼 평탄도(spectral flatness)를 나타내는 값이 될 수 있다.

수학식 2

$$\gamma(k, l) = \sum_k |Y(k, l)|^2 / \sum_k |V_y(k, l)|^2$$

그리고, 입력 신호 분류부(120)는 스펙트럼 평탄도를 나타내는 값을 이용하여 현재 프레임이 잡음 구간인지 아니면, 음성이 존재하는 음성 구간인지를 판단할 수 있다.

그리고, 입력 신호 분류부(120)는 스펙트럼 평탄도를 나타내는 값을 이용하여 현재 프레임이 잡음 구간인지 아니면, 음성이 존재하는 음성 구간인지를 판단할 수 있다.

구체적으로, 입력 신호 분류부(120)는 스펙트럼 평탄도를 나타내는 값이 임계값, 또는 스펙트럼 평탄도를 나타내는 값들의 평균값 미만인 경우, 현재 프레임을 잡음 구간으로 판단할 수 있다. 또한, 입력 신호 분류부(120)는 스펙트럼 평탄도를 나타내는 값이 임계값, 또는 스펙트럼 평탄도를 나타내는 값들의 평균값 이상인 경우, 현재 프레임을 음성 구간으로 판단할 수 있다. 예를 들어, 입력 신호 분류부(120)는 현재 프레임이 다른 프레임에 비하여 스펙트럼 평탄도를 나타내는 값이 높은 경우, 음성 구간으로 판단할 수 있다.

또한, 음성은 성대의 진동 유무에 따라 유성음과 무성음으로 나누어 지며, 두 음성의 특성이 서로 다르다.

현재 프레임이 음성 구간인 경우, 입력 신호 분류부(120)는 현재 프레임이 유성음인지, 아니면 무성음인지를 판단할 수 있다. 이때, 입력 신호 분류부(120)는 음성 에너지와 제로 크로싱 비율(ZCR: zero-crossing rate)을 기초로 현재 프레임을 유성음, 또는 무성음으로 판단할 수 있다.

예를 들어, 백색 잡음의 특성을 가지고 있는 무성음은 낮은 음성 에너지와 높은 ZCR을 가지며 주기적 신호인 유성음은 상대적으로 높은 음성 에너지와 낮은 ZCR을 가질 수 있다. 따라서, 현재 프레임의 음성 에너지가 임계값 미만이거나 현재 프레임이 임계값 이상의 ZCR을 가지는 경우, 입력 신호 분류부(120)는 현재 프레임을 무성음으로 판단할 수 있다. 또한, 현재 프레임의 음성 에너지가 임계값 이상이거나 현재 프레임이 임계값 미만이 ZCR을 가지는 경우, 입력 신호 분류부(120)는 현재 프레임을 유성음으로 판단할 수 있다.

유성음 출력부(130)는 입력 신호에 입력 신호 이득 결정부(110)가 결정한 이득을 적용하여 하모닉 성분을 보존한 유성음을 출력할 수 있다. 이때, 하모닉 성분을 보존한 유성음은 입력 신호 분류부(120)가 분류한 입력 신호의 유성음일 수 있다.

구체적으로, 유성음 출력부(130)는 입력 신호에 입력 신호 이득 결정부(110)가 결정한 이득을 적용하여 중간 출력 신호를 생성하고, 중간 출력 신호를 역 국소 푸리에 변환(ISTFT: Inverse Short-time Fourier transform)하거나, 역 고속 푸리에 변환(IFFT: Inverse fast Fourier transform)하여 하모닉 성분을 보존한 유성음 $\hat{x}_v(n)$ 을 출력할 수 있다.

예를 들어, 유성음 출력부(130)는 수학식 3을 이용하여 중간 출력 신호 $\hat{X}_v(k, l)$ 를 생성할 수 있다.

수학식 3

$$\hat{X}_v(k, l) = Y(k, l)H_c(k, l)$$

[0040]

[0041]

이때, $Y(k, l)$ 는 입력 신호를 국소 푸리에 변환한 입력 스펙트럼이고, $H_c(k, l)$ 는 입력 신호 이득 결정부(110)가 결정한 이득, 또는 입력 신호 이득 결정부(110)가 이용하는 콤 필터의 이득 중 하나일 수 있다.

[0042]

그리고, 유성음 출력부(130)는 하모닉 성분을 보존한 유성음 $\hat{x}_v(n)$ 을 선형 예측 계수 결정부(140)에 전송할 수 있다.

[0043]

선형 예측 계수 결정부(140)는 하모닉 성분을 보존한 유성음 $\hat{x}_v(n)$ 에 기초하여 무성음 보존부(150)가 이용할 선형 예측 계수를 결정할 수 있다. 예를 들어, 선형 예측 계수 결정부(140)는 선형 예측 부호화(LPC: Linear predictive coding)를 수행하는 선형 예측기(Linear prediction)일 수 있다.

[0044]

이때, 선형 예측 계수 결정부(140)는 유성음 출력부(130)로부터 하모닉 성분을 보존한 유성음 $\hat{x}_v(n)$ 을 수신할 수 있다.

[0045]

다음으로, 선형 예측 계수 결정부(140)는 수신한 유성음 $\hat{x}_v(n)$ 을 수학식 4와 같이 계수들의 선형 조합과 잔여 신호로 구분하고, 계수들의 선형 조합에 기초하여 선형 예측 계수를 결정할 수 있다.

수학식 4

$$\hat{x}_v(n) = - \sum_{i=1}^p a_i^c \hat{x}_v(n-i) + v_{\hat{x}_v}(n)$$

[0046]

[0047]

이때, $\hat{x}_v(n)$ 는 중간 출력 신호 $\hat{X}_v(k, l)$ 를 역 고속 푸리에 변환한 $IFFT[\hat{X}_v(k, l)]$ 이며, 중간 출력 신호 $\hat{X}_v(k, l)$ 의 시간 도메인 신호일 수 있다. 또한, $v_{\hat{x}_v}(n)$ 는 잔여 신호이고, a_i^c 는 선형 예측 계수일 수 있다.

[0048]

무성음 보존부(150)는 선형 예측 계수 결정부(140)가 결정한 선형 예측 계수에 기초하여 올 폴 필터(all-pole filter)를 구성하고, 구성한 올 폴 필터를 이용하여 입력 신호의 무성음을 보존할 수 있다. 이때, 올 폴 필터는 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 올 폴 필터에 입력되는 여기 신호 정보로 이용할 수 있다.

[0049]

무성음은 유성음에 비해 에너지가 낮고 백색 잡음과 같은 특성을 가지고 있다. 또한, 유성음이 저주파 대역에 높은 에너지를 갖는데 비해 무성음은 상대적으로 고주파 대역에 에너지가 집중되어 있다. 그리고, 무성음은 비주기적 신호 이므로 콤 필터에 의한 음질 개선 효과가 적을 수 있다.

[0050]

따라서, 무성음 보존부(150)는 콤 필터를 이용하여 결정된 이득을 기초로 결정된 선형 예측 계수를 이용한 올 폴 필터 합성을 이용하여 목적 음성의 무성음 성분을 추정할 수 있다.

[0051]

이때, 무성음 보존부(150)는 수학식 5와 같이 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼 $\hat{v}_x(n)$ 을 올 폴 필터 G에 입력되는 여기 신호(excitation) 정보로 이용하여 입력 신호의 무성음 $\hat{x}_{uv}(n)$ 을 출력할 수 있다. 이때, 잔여 스펙트럼은 잔여 도메인에서 추정된 목적 음성의 잔여 신호일 수 있다.

수학식 5

$$\hat{x}_{uv}(n) = G \hat{v}_x(n)$$

[0052]

이때, 올 폴 필터 G는 수학식 6과 같이 선형 예측 계수 결정부(140)가 결정한 선형 예측 계수 a_i^c 로 구성될 수 있다.

수학식 6

$$G = \frac{1}{1 + \sum_{i=1}^p a_i^c z^{-i}}$$

[0054]

무성음 보존부(150)는 유성음 출력부(130)에서 유성음이 향상된 신호의 선형 예측 계수를 이용하여 입력 신호의 무성음을 처리함으로써, 목적 음성과 더 가까우면서 자연스러운 소리를 얻을 수 있다. 또한, 무성음 보존부(150)는 유성음 출력부(130)에서 유성음이 향상된 신호의 선형 예측 계수를 이용하여 입력 신호의 무성음을 처리함으로써, 종래의 음질 개선 기술들에 비해 신호의 왜곡이 없으므로 에너지가 약한 무성음 성분들도 보존할 수 있다.

[0056]

출력 신호 생성부(160)는 유성음 출력부(130)가 출력한 유성음과 무성음 보존부(150)가 출력한 무성음을 기초로 음성 출력 신호를 생성할 수 있다.

[0057]

이때, 출력 신호 생성부(160)는 입력 신호의 역치 값이 임계값 미만인 구간은 하모닉 성분을 보존한 유성음에 기초하여 음성 출력 신호를 생성하고, 입력 신호의 역치 값이 임계값 이상인 구간은 보존한 무성음에 기초하여 음성 출력 신호를 생성할 수 있다.

[0058]

예를 들어, 출력 신호 생성부(160)는 수학식 7을 이용하여 음성 출력 신호를 생성할 수 있다.

수학식 7

$$\hat{x}_{out}(n) = \begin{cases} \hat{x}_v(n) & \text{if zero crossing rate} < \sigma_v \\ \hat{x}_{uv}(n) & \text{if zero crossing rate} \geq \sigma_v \end{cases}$$

[0059]

이때, σ_v 는 유성음과 무성음을 결정하는 역치 값이고, $\hat{x}_v(n)$ 는 유성음 출력부(130)가 출력한 유성음이며, $\hat{x}_{uv}(n)$ 는 무성음 보존부(150)가 보존한 무성음일 수 있다.

[0060]

일실시예에 따른 음성 신호 처리 장치(100)는 유성음과 무성음의 서로 다른 특성을 고려하여 음성 신호를 처리함으로써, 배경잡음을 효과적으로 줄이면서도 유성음의 하모닉 성분과 백색잡음의 특성을 갖는 무성음 성분들을 효과적으로 보존하여 음성의 명료도를 높일 수 있다.

[0061]

도 2는 일실시예에 따른 입력 신호 이득 결정부의 구조를 도시한 도면이다.

[0062]

도 2를 참고하면, 입력 신호 이득 결정부(110)는 잔여 신호 결정부(210), 하모닉 검출부(220), 국소 푸리에 변환부(230), 콤 필터 설계부(240), 및 이득 결정부(250)를 포함할 수 있다.

[0063]

잔여 신호 결정부(210)는 선형 예측기(Linear prediction)를 이용하여 입력 신호의 잔여 신호(residual signal)를 결정할 수 있다.

[0064]

- [0065] 하모닉 검출부(220)는 잔여 신호 결정부(210)가 결정한 잔여 신호의 스펙트럼 도메인에서 하모닉 성분을 검출할 수 있다.
- [0066] 하모닉 검출부(220)의 상세 구성 및 동작은 이하 도 3을 참조하여 상세히 설명한다.
- [0067] 국소 푸리에 변환부(230)는 입력 신호와 잔여 신호를 각각 국소 푸리에 변환하여 입력 스펙트럼과 잔여 신호 스펙트럼을 출력할 수 있다.
- [0068] 콤 필터 설계부(240)는 하모닉 검출부(220)가 검출한 하모닉 성분에 기초하여 콤(comb) 필터를 설계할 수 있다.
- [0069] 예를 들어, 콤 필터 설계부(240)는 수학적 식 8과 같은 콤 필터 이득 $H_c(k)$ 를 출력하는 콤 필터를 설계할 수 있다.

수학적 식 8

[0070]
$$H_c(k) = \begin{cases} B_c e^{-\frac{2(k-k_c)^2}{\sigma}} & k \in [k_c - k_0/2, k_c + k_0/2] \\ B_k & \text{otherwise} \end{cases}$$

- [0071] 이때, k_c 는 하모닉 검출부(220)가 검출한 하모닉 성분이고, k_0 는 입력 신호에서 현재 프레임의 기본 주파수(fundamental frequency)일 수 있다.
- [0072] 또한, $B_c(k)$ 는 필터의 가중치 값이고, $B_k(k)$ 는 위너 필터를 이용해 설계된 이득 값일 수 있다. 이때, $B_k(k)$ 는 하모닉 성분이 아닌 다른 구간에 적용될 수 있다. 그리고, $B_c(k)$ 와 $B_k(k)$ 는 각각 수학적 식 9와 수학적 식 10으로 표현될 수 있다.

수학적 식 9

[0073]
$$B_c(k) = \frac{E[|\hat{X}(k)|^2]}{E[|Y(k)|^2]}$$

수학적 식 10

[0074]
$$B_k(k) = \frac{\xi(k)}{1+\xi(k)}$$

- [0075] 이때, $\xi(k)$ 는 수학적 식 11로 표현될 수 있다.

수학적 식 11

[0076]
$$\xi(k) = \frac{E[|\hat{X}(k)|^2]}{E[|W(k)|^2]}$$

- [0077] 콤 필터 설계부(240)가 설계한 콤 필터는 일정한 간격으로 스파이크가 반복되는 주파수 응답(frequency response)을 가지는 함수로써 일정한 주기로 반복되는 하모닉 성분들을 살리는데 효과적인 필터일 수 있다. 따라서, 콤 필터 설계부(240)가 설계한 콤 필터는 일반적인 잡음 추정 알고리즘을 통해서 형성된 이득이 에너지가

약한 하모닉 성분들을 없애는 단점을 보완해 줄 수 있다.

- [0078] 이득 결정부(250)는 위너 필터를 이용하여 입력 신호를 필터링한 결과인 위너 필터 이득 및 콤 필터 설계부(240)가 설계한 콤 필터를 이용하여 입력 신호를 필터링한 결과인 콤 필터 이득에 기초하여 입력 신호의 이득을 결정할 수 있다. 이때, 위너 필터 이득은 단채널 음질 개선 알고리즘을 이용하여 획득한 이득일 수 있다.
- [0079] 입력 신호 이득 결정부(110)는 목적 음성 신호의 잔여 스펙트럼에서 하모닉 성분을 검출하여 유성음의 하모닉 특성을 고려한 콤 필터를 설계하고, 설계한 콤 필터를 이용한 이득과, 위너 필터를 이용한 이득을 결합함으로써, 음성의 하모닉 성분들의 왜곡을 최소화 하면서 충분한 배경 잡음을 제거 할 수 있는 이득을 형성할 수 있다.
- [0080] 도 3은 일실시예에 따른 하모닉 검출부의 구조를 도시한 도면이다.
- [0081] 도 3을 참고하면, 하모닉 검출부(220)는 잔여 스펙트럼 추정부(310), 최고점 검출부(320), 및 하모닉 성분 검출부(330)를 포함할 수 있다.
- [0082] 잔여 스펙트럼 추정부(310)는 잔여 신호 결정부(210)가 결정한 잔여 신호의 스펙트럼 도메인에서, 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 추정할 수 있다. 이때, 잔여신호의 스펙트럼은 주파수 편평도(flatness)의 영향으로 신호의 주파수 도메인 신호보다 잡음에 묻혀 있던 하모닉 성분을 더 쉽게 검출할 수 있다.
- [0083] 최고점 검출부(320)는 최고점 검출 알고리즘을 이용하여 잔여 스펙트럼 추정부(310)가 추정한 잔여 스펙트럼에서 복수의 최고점들을 검출할 수 있다.
- [0084] 하모닉 성분 검출부(330)는 최고점 검출부(320)가 검출한 최고점들 간의 간격에 기초하여 하모닉 성분을 검출할 수 있다.
- [0085] 예를 들어, 최고점 검출부(320)가 검출한 최고점 간의 간격이 $0.7k_0$ 보다 작은 경우, 하모닉 성분 검출부(330)는 최고점 검출부(320)가 검출한 최고점들을 잡음에 의한 피크로 간주하여 삭제할 수 있다.
- [0086] 또한, 최고점 검출부(320)가 검출한 최고점 간의 간격이 $1.3k_0$ 보다 큰 경우, 하모닉 성분 검출부(330)는 최고점 검출부(320)가 검출한 최고점들 사이에 사라진 하모닉 성분이 존재한다고 간주하고, 기본 주파수의 정수 배를 이용하여 사라진 하모닉 성분을 검출할 수 있다.
- [0087] 도 4는 일실시예에 따른 음성 신호 처리 과정에서 정보들의 흐름을 도시한 도면이다.
- [0088] 먼저, 입력 신호 이득 결정부(110)의 잔여 신호 결정부(210)는 선형 예측기(Linear prediction)(410)로 입력 신호 $y(n)$ 을 선형 예측 부호화(LPC: Linear predictive coding)하여 입력 신호 $y(n)$ 의 잔여 신호 $v_y(n)$ (411)를 출력할 수 있다.
- [0089] 이때, 하모닉 검출부(220)는 잔여 신호 $v_y(n)$ (411)의 스펙트럼 도메인에서, 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 추정하고, 추정한 잔여 스펙트럼에서 하모닉 성분들을 검출할 수 있다. 또한, 콤 필터 설계부(240)는 하모닉 검출부(220)가 검출한 하모닉 성분에 기초하여 콤 필터(comb filter)(430)를 설계할 수 있다.
- [0090] 다음으로, 국소 푸리에 변환부(230)는 입력 신호 $y(n)$ 과 잔여 신호 $v_y(n)$ (411)를 각각 국소 푸리에 변환하여 입력 스펙트럼 $Y(k, l)$ (421)과 잔여 신호 스펙트럼 $V_y(k, l)$ (422)을 출력할 수 있다.
- [0091] 이때, 하모닉 검출부(220)가 검출한 하모닉 성분에 기초한 콤 필터(430)는 잔여 신호 스펙트럼 $V_y(k, l)$ (422)을 필터링하여 콤 필터 이득 $H_c(k, l)$ (431)을 출력할 수 있다.
- [0092] 또한, 단 채널 위너 필터 중 하나인 SCSE는 입력 스펙트럼 $Y(k, l)$ (421)를 필터링하여 위너 필터 이득 $G_{wiener}(k, l)$ (441)을 출력할 수 있다.
- [0093] 그리고, 이득 결정부(250)는 콤 필터 이득 $H_c(k, l)$ (431)과 위너 필터 이득 $G_{winner}(k, l)$ (441)를 결합하여 입력

신호의 이득(gain)(450)를 결정할 수 있다.

[0094] 또한, 입력 신호 분류부(120)는 입력 신호 $y(n)$ 를 유성음과 무성음으로 분류할 수 있다.

[0095] 그 다음으로, 유성음 출력부(130)는 입력 신호 $y(n)$ 에 이득(450)을 적용하여 중간 출력 신호 $\hat{X}_v(k,l)$ (461)를 생성할 수 있다.

[0096] 그리고, 유성음 출력부(130)는 역 국소 푸리에 변환기(460)로 중간 출력 신호 $\hat{X}_v(k,l)$ (461)를 역 국소 푸리에 변환(ISTFT)하여 입력 신호 분류부(120)가 분류한 유성음 $\hat{x}_v(n)$ (462)을 출력할 수 있다.

[0097] 그리고, 유성음 출력부(130)는 유성음 $\hat{x}_v(n)$ (462)을 선형 예측 계수 결정부(140)에 전송할 수 있다.

[0098] 다음으로, 선형 예측 계수 결정부(140)는 선형 예측기(470)로 유성음 $\hat{x}_v(n)$ (462)을 선형 예측 부호화(LPC: Linear predictive coding)하여 선형 예측 계수 a_i^c 를 결정할 수 있다.

[0099] 구체적으로, 선형 예측 계수 결정부(140)는 수신한 유성음 $\hat{x}_v(n)$ (462)을 수학식 4와 같이 계수들의 선형 조합과 잔여 신호로 구분하고, 계수들의 선형 조합에 기초하여 선형 예측 계수 a_i^c 를 결정할 수 있다.

[0100] 그 다음으로, 무성음 보존부(150)는 선형 예측 계수 결정부(140)가 결정한 선형 예측 계수 a_i^c 에 기초하여 올 폴 필터(all-pole filter)(480)를 구성하고, 구성한 올 폴 필터를 이용하여 입력 신호의 무성음을 보존할 수 있다.

[0101] 구체적으로, 무성음 보존부(150)는 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼 $\hat{v}_x(n)$ (481)을 올 폴 필터(480)에 입력되는 여기 신호(excitation) 정보로 이용하여 입력 신호의 무성음 $\hat{x}_{uv}(n)$ (482)을 출력할 수 있다.

[0102] 마지막으로, 출력 신호 생성부(160)는 유성음 출력부(130)가 출력한 유성음 $\hat{x}_v(n)$ (462)과 무성음 보존부(150)가 출력한 무성음 $\hat{x}_{uv}(n)$ (482)을 기초로 음성 출력 신호 $\hat{x}_{out}(n)$ (491)를 생성할 수 있다.

[0103] 이때, 출력 신호 생성부(160)는 입력 신호의 역치 값이 임계값 미만인 구간은 유성음 $\hat{x}_v(n)$ (462)을 선택하여 음성 출력 신호 $\hat{x}_{out}(n)$ (491)을 생성하고, 입력 신호의 역치 값이 임계값 이상인 구간은 무성음 $\hat{x}_{uv}(n)$ (482)을 선택하여 음성 출력 신호 $\hat{x}_{out}(n)$ (491)을 생성할 수 있다.

[0104] 도 5는 일실시예에 따른 하모닉 검출 결과의 일례이다.

[0105] 도 5의 케이스 1(Case 1)은 종래 기술에 따라 주파수 도메인 신호에서 하모닉 성분을 검출한 결과이고, 케이스 2(Case 2)는 하모닉 검출부(220)가 잔여 신호의 스펙트럼에서 하모닉 성분을 검출한 결과일 수 있다. 그리고, 케이스 1과 케이스 2는 백색 잡음이 가해진 음성 입력 신호의 5dB SNR에서 동일한 조건의 최고점 검출 알고리즘을 적용한 결과일 수 있다.

[0106] 주파수 도메인 신호(500)는 케이스 1에 도시된 바와 같이 복수의 최고점들을 포함할 수 있다. 그리고, 종래 기술은 주파수 도메인 신호(500)의 최고점들 중에서 하모닉 성분(501)을 검출할 수 있다. 그러나, 케이스 1에서 2kHz~4kHz 대역(510)의 최고점들은 검출된 하모닉 성분(501)보다 에너지가 낮으므로, 하모닉 성분으로 검출될 수 없다.

- [0107] 반면, 잔여 신호의 스펙트럼은 케이스 2에 도시된 바와 같이 최고점들 간의 에너지 차이가 주파수 도메인 신호(500)보다 적을 수 있다. 따라서, 하모닉 검출부(220)는 2kHz~4kHz 대역(520)에 포함된 최고점들도 하모닉 성분으로 검출할 수 있다.
- [0108] 도 6은 일실시예에 따른 콤 필터로 필터링한 결과인 콤 필터 이득의 일례이다.
- [0109] 도 6은 입력 신호에서 유성음들이 포함된 유성음 구간의 스펙트럼(610)과 콤 필터로 필터링한 결과인 콤 필터 이득(620)을 도시하고 있다.
- [0110] 유성음 구간의 스펙트럼(610)은 도 6에 도시된 바와 같이 목표 음성의 스펙트럼(611)에 잡음이 포함된 잡음 음성 스펙트럼(612)일 수 있다. 이때, 잡음 음성 스펙트럼(612)는 잡음에 의하여 최고점(621), 최고점(622)와 같이 목표 음성 스펙트럼(611)의 최고점이 묻혀버릴 가능성이 있다.
- [0111] 반면, 콤 필터 설계부(240)가 설계한 콤 필터는 일정한 주기로 반복되는 하모닉 성분을 살릴 수 있다. 따라서, 콤 필터로 필터링한 결과인 콤 필터 이득(620)은 도 6에 도시된 바와 같이 에너지가 작아 잡음에 의해 묻혔던 최고점(621), 최고점(622)이 잡음으로 판단되어 삭제되는 것을 방지할 수 있다.
- [0112] 도 7은 일실시예에 따른 음성 신호 처리 방법을 도시한 도면이다.
- [0113] 단계(710)에서 입력 신호 이득 결정부(110)는 유성음의 하모닉 특성에 기초한 콤(comb) 필터를 이용하여 입력 신호의 이득(gain)을 결정할 수 있다. 이때, 콤 필터는 일정한 간격으로 스파이크가 반복되는 주파수 응답(frequency response)을 가지는 함수일 수 있다. 또한, 입력 신호는 난청 보상 장치의 마이크가 수집한 음성 신호일 수 있다.
- [0114] 단계(720)에서 입력 신호 분류부(120)는 입력 신호를 유성음과 무성음으로 분류할 수 있다. 예를 들어, 입력 신호 분류부(120)는 유무성음 판별기와 VAD(voiced activity detector)를 이용하여 입력 신호의 현재 프레임이 잡음 구간인지 여부를 판단하고, 잡음 구간이 아닌 경우, 현재 프레임에 포함된 음성을 유성음, 또는 무성음으로 분류할 수 있다.
- [0115] 단계(730)에서 유성음 출력부(130)는 입력 신호에 단계(710)에서 결정한 이득을 적용하여 하모닉 성분을 보존한 유성음을 출력할 수 있다. 이때, 하모닉 성분을 보존한 유성음은 단계(720)에서 분류한 입력 신호의 유성음일 수 있다.
- [0116] 구체적으로, 유성음 출력부(130)는 입력 신호에 입력 신호 이득 결정부(110)가 결정한 이득을 적용하여 중간 출력 신호를 생성하고, 중간 출력 신호를 역 국소 푸리에 변환(ISTFT: Inverse Short-time Fourier transform)하거나, 역 고속 푸리에 변환(IFFT: Inverse fast Fourier transform)하여 하모닉 성분을 보존한 유성음 $\hat{x}_v(n)$ 을 출력할 수 있다.
- [0117] 단계(740)에서 선형 예측 계수 결정부(140)는 단계(730)에서 출력한 유성음에 기초하여 무성음 보존부(150)가 이용할 선형 예측 계수를 결정할 수 있다.
- [0118] 단계(750)에서 무성음 보존부(150)는 단계(740)에서 결정한 선형 예측 계수에 기초하여 올 폴 필터(all-pole filter)를 구성하고, 구성한 올 폴 필터를 이용하여 입력 신호의 무성음을 보존할 수 있다. 이때, 올 폴 필터는 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 올 폴 필터에 입력되는 여기 신호 정보로 이용할 수 있다.
- [0119] 단계(760)에서 출력 신호 생성부(160)는 단계(730)에서 출력한 유성음과 단계(750)에서 보존한 무성음을 기초로 음성 출력 신호를 생성할 수 있다.
- [0120] 이때, 출력 신호 생성부(160)는 입력 신호의 역치 값이 임계값 미만인 구간은 하모닉 성분을 보존한 유성음에 기초하여 음성 출력 신호를 생성하고, 입력 신호의 역치 값이 임계값 이상인 구간은 보존한 무성음에 기초하여 음성 출력 신호를 생성할 수 있다.
- [0121] 일실시예에 따른 음성 신호 처리 방법은 유성음과 무성음의 서로 다른 특성을 고려하여 음성 신호를 처리함으로써, 배경잡음을 효과적으로 줄이면서도 유성음의 하모닉 성분과 백색잡음의 특성을 갖는 무성음 성분들을 효과

적으로 보존하여 음성의 명료도를 높일 수 있다.

- [0122] 도 8은 일실시예에 따른 입력 신호 이득 결정 과정을 도시한 도면이다. 이때, 도 8의 단계(810) 내지 단계(850)는 도 7의 단계(710)에 포함될 수 있다.
- [0123] 단계(810)에서 잔여 신호 결정부(210)는 선형 예측기(Linear prediction)를 이용하여 입력 신호의 잔여 신호(residual signal)를 결정할 수 있다.
- [0124] 단계(820)에서 하모닉 검출부(220)는 단계(810)에서 결정한 잔여 신호의 스펙트럼 도메인에서 하모닉 성분을 검출할 수 있다.
- [0125] 단계(830)에서 국소 푸리에 변환부(230)는 입력 신호와 단계(810)에서 결정한 잔여 신호를 각각 국소 푸리에 변환하여 입력 스펙트럼과 잔여 신호 스펙트럼을 출력할 수 있다.
- [0126] 단계(840)에서 콦 필터 설계부(240)는 단계(820)에서 검출한 하모닉 성분에 기초하여 콦(comb) 필터를 설계할 수 있다. 이때, 콦 필터 설계부(240)가 설계한 콦 필터는 일정한 간격으로 스파이크가 반복되는 주파수 응답(frequency response)을 가지는 함수로써 일정한 주기로 반복되는 하모닉 성분들을 살리는데 효과적인 필터일 수 있다.
- [0127] 단계(850)에서 이득 결정부(250)는 위너 필터를 이용하여 단계(830)에서 출력한 입력 스펙트럼을 필터링한 결과인 위너 필터 이득 및 단계(840)에서 설계한 콦 필터를 이용하여 단계(830)에서 출력한 잔여 신호 스펙트럼을 필터링한 결과인 콦 필터 이득에 기초하여 입력 신호의 이득을 결정할 수 있다. 이때, 위너 필터 이득은 단채널 음질 개선 알고리즘을 이용하여 획득한 이득일 수 있다.
- [0128] 도 9는 일실시예에 따른 하모닉 검출 과정을 도시한 도면이다. 이때, 도 9의 단계(910) 내지 단계(930)는 도 8의 단계(820)에 포함될 수 있다.
- [0129] 단계(910)에서 잔여 스펙트럼 추정부(310)는 도 8의 단계(810)에서 결정한 잔여 신호의 스펙트럼 도메인에서, 입력 신호에 포함된 목적 음성 신호의 잔여 스펙트럼을 추정할 수 있다.
- [0130] 단계(920)에서 최고점 검출부(320)는 최고점 검출 알고리즘을 이용하여 단계(910)에서 추정한 잔여 스펙트럼에서 복수의 최고점들을 검출할 수 있다.
- [0131] 단계(930)에서 하모닉 성분 검출부(330)는 단계(920)에서 검출한 최고점들 간의 간격에 기초하여 하모닉 성분을 검출할 수 있다.
- [0132] 예를 들어, 최고점 검출부(320)가 검출한 최고점 간의 간격이 0.7k0 보다 작은 경우, 하모닉 성분 검출부(330)는 최고점 검출부(320)가 검출한 최고점들을 잡음에 의한 피크로 간주하여 삭제할 수 있다.
- [0133] 또한, 최고점 검출부(320)가 검출한 최고점 간의 간격이 1.3k0 보다 큰 경우, 하모닉 성분 검출부(330)는 최고점 검출부(320)가 검출한 최고점들 사이에 사라진 하모닉이 존재한다고 간주하고, 기본 주파수의 정수 배를 이용하여 사라진 하모닉 성분을 검출할 수 있다.
- [0134] 일실시예에 따른 음성 신호 처리 장치 및 방법은 유성음과 무성음의 서로 다른 특성을 고려하여 음성 신호를 처리함으로써, 배경잡음을 효과적으로 줄이면서도 유성음의 하모닉 성분과 백색잡음의 특성을 갖는 무성음 성분들을 효과적으로 보존하여 음성의 명료도를 높일 수 있다.
- [0135] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체

(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0136]

[0137]

이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0138]

[0139]

그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

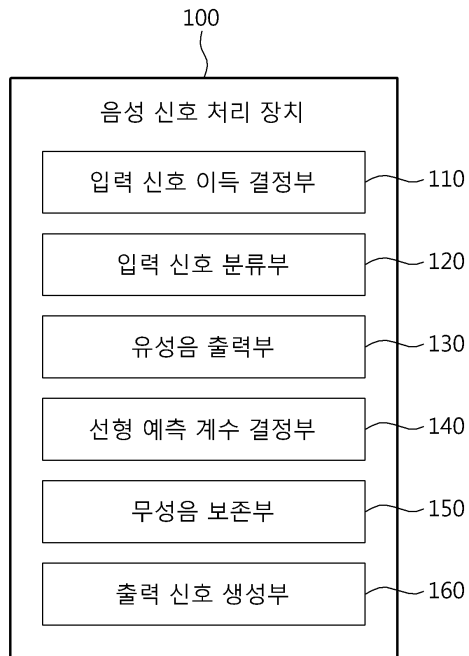
부호의 설명

[0140]

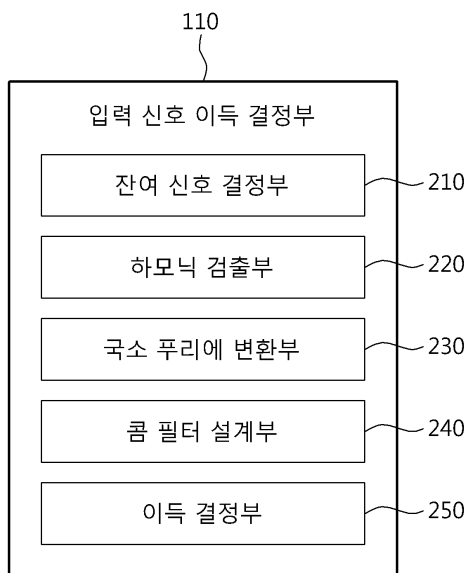
- 100: 음성 신호 처리 장치
- 110: 입력 신호 이득 결정부
- 130: 유성음 출력부
- 140: 선형 예측 계수 결정부
- 150: 무성음 보존부
- 210: 잔여 신호 결정부
- 220: 하모닉 검출부
- 240: 콤 필터 설계부
- 250: 이득 결정부

도면

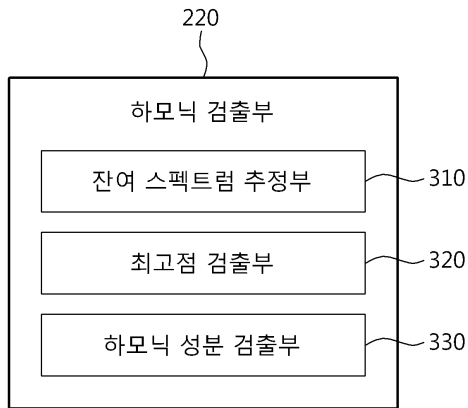
도면1



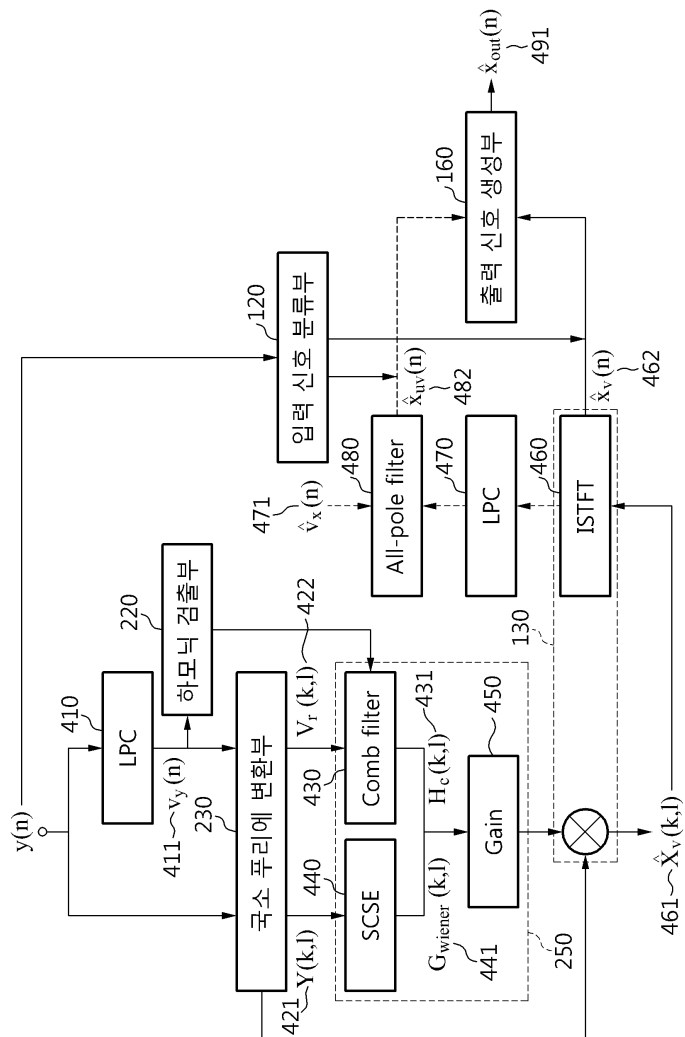
도면2



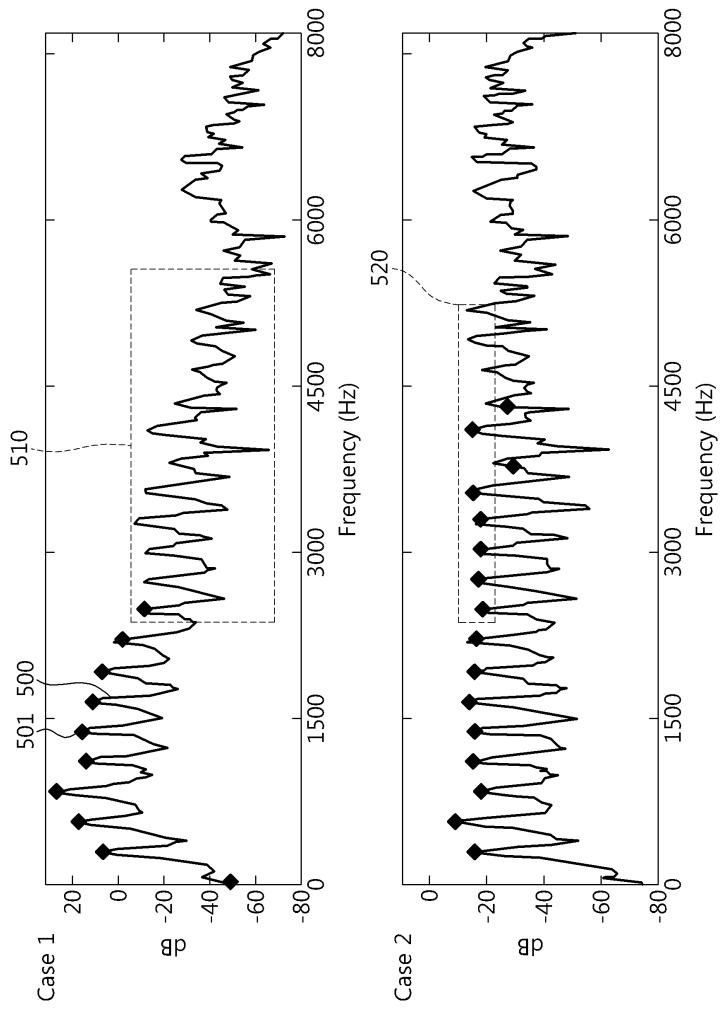
도면3



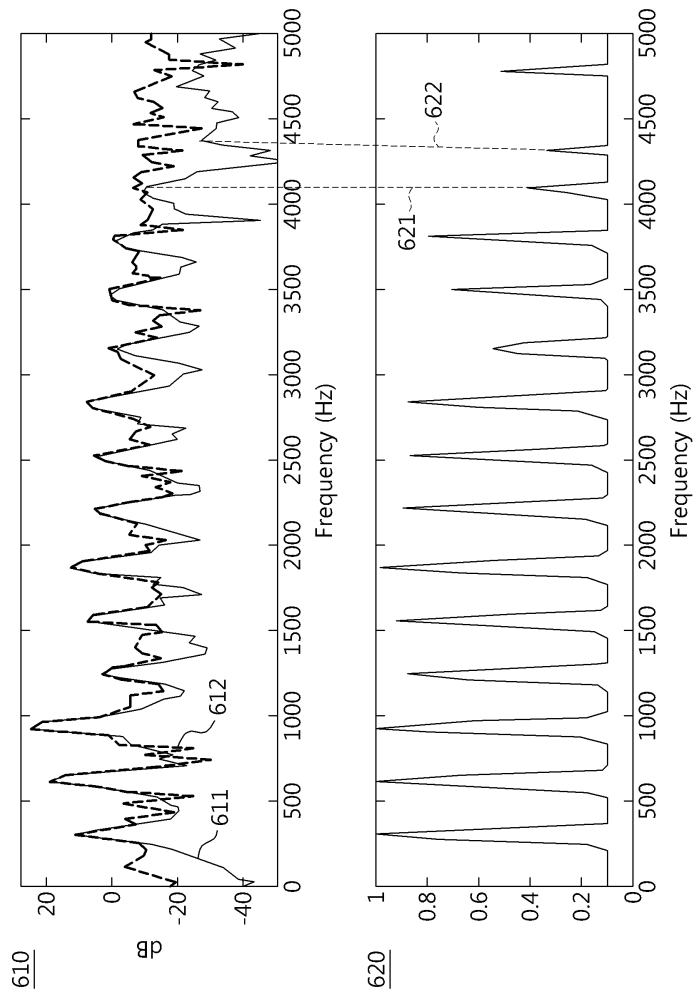
도면4



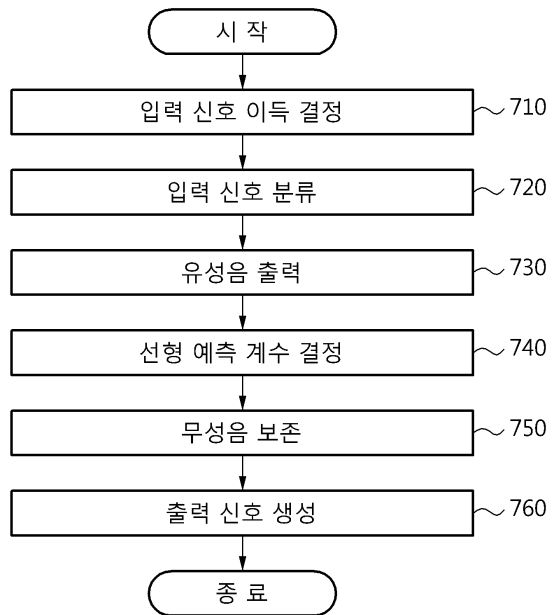
도면5



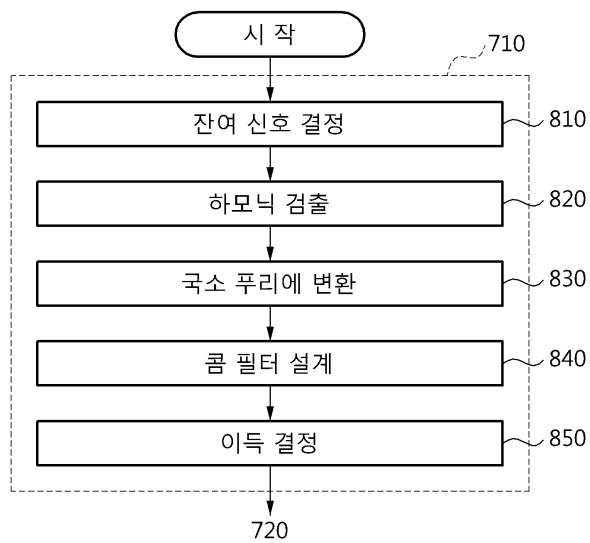
도면6



도면7



도면8



도면9

