



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년07월30일
(11) 등록번호 10-1290997
(24) 등록일자 2013년07월23일

(51) 국제특허분류(Int. Cl.)
G10L 21/02 (2006.01) G10L 19/00 (2006.01)
(21) 출원번호 10-2012-0087259
(22) 출원일자 2012년08월09일
심사청구일자 2012년08월09일
(30) 우선권주장
1020120030582 2012년03월26일 대한민국(KR)
(56) 선행기술조사문헌
KR1020090104559 A
EP0545386 A2

(73) 특허권자
세종대학교산학협력단
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
(72) 발명자
김무영
서울특별시 강남구 도곡2동 도곡렉슬아파트 210동 402호
황인우
서울특별시 광진구 군자동 88-18 203호
(74) 대리인
특허법인태백

전체 청구항 수 : 총 22 항

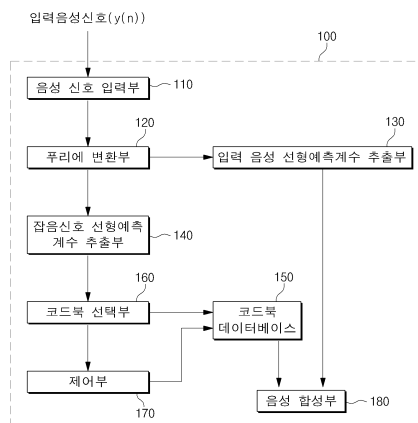
심사관 : 권영학

(54) 발명의 명칭 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법 및 그 장치

(57) 요약

본 발명은 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법 및 그 장치에 관한 것이다. 본 발명에 따른 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법은, 클린 음성 신호와 잡음 신호가 혼합된 입력 음성 신호를 프레임 단위로 입력받는 단계, 상기 입력 음성 신호를 푸리에 변환시켜 입력 음성 파워 스펙트럼을 생성하는 단계, 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계, 상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하고, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계, 기 저장된 복수의 잡음 코드북들 중에서, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값과 가장 유사한 잡음 코드북을 가지고 있는 잡음 코드북을 선택하는 단계, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N개를 적응형 코드벡터로 설정하고, 상기 N개의 적응형 코드벡터를 상기 선택된 잡음 코드북에 저장시켜 슈퍼 코드북을 생성하는 단계, 그리고 상기 슈퍼 코드북 내의 잡음 코드벡터들과 클린 코드북 내의 클린 코드벡터들을 조합한 통합 코드벡터와 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 이용하여 보정된 음성 신호를 추출하는 단계를 포함한다.

대표도 - 도1



특허청구의 범위

청구항 1

클린 음성 신호와 잡음 신호가 혼합된 입력 음성 신호를 프레임 단위로 입력받는 단계,
 상기 입력 음성 신호를 푸리에 변환시켜 입력 음성 파워 스펙트럼을 생성하는 단계,
 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계,
 상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하고, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계,
 기 저장된 복수의 잡음 코드북들 중에서, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값과 가장 유사한 잡음 코드백터를 가지고 있는 잡음 코드북을 선택하는 단계,
 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N개를 적응형 코드백터로 설정하고, 상기 N개의 적응형 코드백터를 상기 선택된 잡음 코드북에 저장시켜 슈퍼 코드북을 생성하는 단계, 그리고
 상기 슈퍼 코드북 내의 잡음 코드백터들과 클린 코드북 내의 클린 코드백터들을 조합한 통합 코드백터와 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 이용하여 보정된 음성 신호를 추출하는 단계를 포함하는 적응형 코드백터를 이용한 코드북 기반 음성 향상 방법.

청구항 2

제1항에 있어서,
 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계는,
 상기 입력 음성 파워 스펙트럼을 역푸리에 변환시켜 상기 입력 음성 신호의 자기상관 값을 생성하는 단계,
 상기 입력 음성 신호의 자기상관 값을 레빈슨 알고리즘(Levinson algorithm)에 적용하여, 상기 입력 음성 신호에 대한 선형예측계수를 추출하는 단계, 그리고
 상기 입력 음성 신호에 대한 선형예측계수의 주파수 응답 값을 이용하여 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계를 포함하는 적응형 코드백터를 이용한 코드북 기반 음성 향상 방법.

청구항 3

제2항에 있어서,
 상기 입력 음성 신호로부터 잡음 신호를 추정하고, 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계는,
 Long-term 잡음 추정 알고리즘을 이용하여 상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하는 단계,
 상기 추정된 잡음 신호의 잡음 타입을 이용하여, 상기 기 저장된 복수의 잡음 코드북들 중에서 1개의 잡음 코드북을 선택하는 단계,
 상기 추정된 잡음 신호에 대한 잡음 파워 스펙트럼을 역푸리에 변환시켜 상기 추정된 잡음 신호의 자기 상관 값을 생성하는 단계,
 상기 추정된 잡음 신호의 자기 상관 값을 레빈슨 알고리즘에 적용하여, 상기 추정된 잡음 신호에 대한 선형예측계수를 추출하는 단계, 그리고
 상기 추정된 잡음 신호에 대한 선형예측계수의 주파수 응답 값을 이용하여 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계를 포함하는 적응형 코드백터를 이용한 코드북 기반 음성 향상 방법.

청구항 4

제3항에 있어서,

상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값과 가장 유사한 잡음 코드벡터를 가지고 있는 잡음 코드북을 선택하는 단계는,

다음의 수학적식과 같은 IS-D(Itakura-Saito Distortion) 알고리즘을 이용하여 상기 잡음 코드북을 선택하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법:

$$n' = \arg \max_n \left\{ \max_{1 \leq j \leq \nu_n} p(\hat{\mathbb{D}} | a_d^{n,j}) \right\}, 1 \leq n \leq N_d$$

여기서 n' 는 선택된 잡음 코드북, $\hat{\mathbb{D}} = [d[0], d[1], \dots, d[K-1]]^T$ 는 추정된 잡음신호, N_d 는 코드북 데이터베이스 내 잡음 코드북의 개수, ν_n 은 n 번째 코드북 내 코드벡터의 개수이고, $a_d^{n,j}$ 은 n 번째 잡음 코드북의 j 번째 코드벡터의 선형 예측 계수이다.

청구항 5

제4항에 있어서,

상기 슈퍼 코드북을 생성하는 단계는,

상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N 개를 상기 적응형 코드벡터로 설정하여 상기 선택된 잡음 코드북에 저장시키는 단계, 그리고

상기 선택된 잡음 코드북에 기 저장된 잡음 코드벡터와 상기 N 개의 적응형 코드벡터를 통합하여 상기 선택된 잡음 코드북을 상기 슈퍼 코드북으로 변경시키는 단계를 포함하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법.

청구항 6

제5항에 있어서,

상기 보정된 음성 신호를 추출하는 단계는,

상기 슈퍼 코드북에 포함되는 잡음 코드벡터와 상기 클린 코드북에 포함되는 클린 코드벡터를 조합하여 상기 통합 코드벡터를 생성하는 단계,

상기 통합 코드벡터에 대하여 클린 음성 신호의 게인 값과 상기 추정된 잡음 신호의 게인 값을 적용하여 각각 게인 조합 값을 생성하는 단계,

상기 입력 음성 신호의 스펙트럴 엔벨로프 값과 상기 게인 조합 값들 사이의 각각의 유사도를 구하는 단계, 그리고

상기 유사도와 게인 조합 값을 ML(Maximum Likelihood) 또는 베이시안 추정 알고리즘(MMSE)에 적용하여 상기 클린 음성 신호의 스펙트럴 엔벨로프 값과 상기 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계를 포함하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법.

청구항 7

제6항에 있어서,

상기 보정된 음성 신호를 생성하는 단계는,

위너 필터(Wiener filter)를 통해 상기 클린 음성 신호의 스펙트럴 엔벨로프 값과 상기 잡음 신호의 스펙트럴 엔벨로프 값으로부터 보정된 음성 파워 스펙트럼을 생성하는 단계, 그리고

상기 보정된 음성 파워 스펙트럼을 역푸리에 변환시켜 상기 보정된 음성 신호를 추출하는 단계를 더 포함하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법.

청구항 8

제6항에 있어서,

상기 클린 음성 신호의 계인 값과 상기 추정된 잡음 신호의 계인 값은 음성 부존재 확률(SAP) 알고리즘을 이용하여 획득하는 적응형 코드백터를 이용한 코드북 기반 음성 향상 방법.

청구항 9

제8항에 있어서,

상기 클린 음성 신호의 계인 값(c_g)과 상기 추정된 잡음 신호의 계인 값(n_g)은 다음의 수학적식을 통해 연산되는 적응형 코드백터를 이용한 코드북 기반 음성 향상 방법:

$$\begin{bmatrix} \mathbf{C}_g \\ \mathbf{n}_g \end{bmatrix} = \text{SAP} \times \begin{bmatrix} \mathbf{C}_g' \\ \mathbf{n}_g' \end{bmatrix} + (1-\text{SAP}) \times \begin{bmatrix} \mathbf{C}_g'' \\ \mathbf{n}_g'' \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{C}_g' \\ \mathbf{n}_g' \end{bmatrix} = \begin{bmatrix} 0 \\ \| |Y(\omega)|^2 |A_d^n(\omega)|^2 \| \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{C}_g'' \\ \mathbf{n}_g'' \end{bmatrix} = \mathbf{E}^{-1} \cdot \mathbf{F}$$

$$\mathbf{E} = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^4} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_d^n(\omega)|^4} \right\| \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^2 |A_x^m(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^2 |A_d^n(\omega)|^2} \right\| \end{bmatrix}$$

여기서, $\|f(\omega)\| = \int |f(\omega)| d\omega$ 이며, $Y(\omega)$ 는 입력 음성 신호의 주파수 응답 값을 나타내며, $A_d^n(\omega)$ 은 슈퍼 코드북에 저장된 n 번째 코드백터를 나타내고, $A_x^m(\omega)$ 은 클린 코드북에 저장된 m 번째 코드백터를 나타낸다.

청구항 10

제1항에 있어서,

상기 입력 음성 신호에 이어서 입력되는 다음 프레임의 입력 음성 신호로부터 잡음 신호를 추정하고, 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계,

상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 새로운 적응형 코드백터로 설정하는 단계, 그리고

상기 슈퍼 코드북에 저장된 상기 N 개의 적응형 코드백터 중에서 적어도 하나의 적응형 코드백터를 제거하고, 상기 새로운 적응형 코드백터를 상기 슈퍼 코드북에 저장하는 단계를 더 포함하는 적응형 코드백터를 이용한 코드북 기반 음성 향상 방법.

청구항 11

제10항에 있어서,

상기 슈퍼 코드북에 저장된 상기 N개의 적응형 코드벡터 중에서 저장된 시기가 가장 빠른 적응형 코드벡터를 제거하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법.

청구항 12

클린 음성 신호와 잡음 신호가 혼합된 입력 음성 신호를 프레임 단위로 입력받는 음성 신호 입력부,

상기 입력 음성 신호를 푸리에 변환시켜 입력 음성 파워 스펙트럼을 생성하는 푸리에 변환부,

상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 입력 음성 선형예측계수 추출부,

상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하고, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 잡음 신호 선형예측계수 추출부,

복수의 클린 코드벡터를 가지는 클린 음성 코드북과 복수의 잡음 코드벡터를 가지며 서로 다른 타입을 가지는 복수의 잡음 코드북을 포함하는 코드북 데이터베이스,

상기 복수의 잡음 코드북들 중에서, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값과 가장 유사한 잡음 코드벡터를 가지고 있는 잡음 코드북을 선택하는 코드북 선택부,

상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N개를 적응형 코드벡터로 설정하고, 상기 N개의 적응형 코드벡터를 상기 선택된 잡음 코드북에 저장시켜 슈퍼 코드북을 생성하는 제어부, 그리고

상기 슈퍼 코드북 내의 잡음 코드벡터들과 클린 코드북 내의 클린 코드벡터들을 조합한 통합 코드벡터와 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 이용하여 보정된 음성 신호를 추출하는 음성 합성부를 포함하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 13

제12항에 있어서,

상기 입력 음성 선형예측계수 추출부는,

상기 입력 음성 파워 스펙트럼을 역푸리에 변환시켜 상기 입력 음성 신호의 자기상관 값을 생성하고,

상기 입력 음성 신호의 자기상관 값을 레빈슨 알고리즘(Levinson algorithm)에 적용하여 상기 입력 음성 신호에 대한 선형예측계수를 추출한 다음,

상기 입력 음성 신호에 대한 선형예측계수의 주파수 응답 값을 이용하여 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 14

제13항에 있어서,

상기 잡음 신호 선형예측계수 추출부는,

Long-term 잡음 추정 알고리즘을 이용하여 상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하고, 상기 추정된 잡음 신호의 잡음 타입을 이용하여, 상기 복수의 잡음 코드북들 중에서 1개의 잡음 코드북을 선택하며, 상기 추정된 잡음 신호에 대한 잡음 파워 스펙트럼을 역푸리에 변환시켜 추정된 잡음 신호의 자기 상관 값을 생성한 다음,

상기 추정된 잡음 신호의 자기 상관 값을 레빈슨 알고리즘에 적용하여, 상기 추정된 잡음 신호에 대한 선형예측계수를 추출하고, 상기 추정된 잡음 신호에 대한 선형예측계수의 주파수 응답 값을 이용하여 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 15

제14항에 있어서,

상기 코드북 선택부는,

다음의 수식식과 같은 IS-D(Itakura-Saito Distortion) 알고리즘을 이용하여 상기 잡음 코드북을 선택하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치:

$$n' = \arg \max_n \left\{ \max_{1 \leq j \leq v_n} p(\hat{\mathbb{D}} | a_d^{n,j}) \right\}, 1 \leq n \leq N_d$$

여기서 n' 는 선택된 잡음 코드북, $\hat{\mathbb{D}} = [d[0], d[1], \dots, d[K-1]]^T$ 는 추정된 잡음신호, N_d 는 코드북 데이터베이스 내 잡음 코드북의 개수, v_n 은 n 번째 코드북 내 코드벡터의 개수이고, $a_d^{n,j}$ 은 n 번째 잡음 코드북의 j 번째 코드벡터의 선형 예측 계수이다.

청구항 16

제15항에 있어서,

상기 제어부는,

상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N 개를 상기 적응형 코드벡터로 설정하여 상기 선택된 잡음 코드북에 저장시키고,

상기 선택된 잡음 코드북에 기 저장된 잡음 코드벡터와 상기 N 개의 적응형 코드벡터를 통합하여 상기 선택된 잡음 코드북을 상기 슈퍼 코드북으로 변경시키는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 17

제16항에 있어서,

상기 음성 합성부는,

상기 슈퍼 코드북에 포함되는 잡음 코드벡터와 상기 클린 코드북에 포함되는 클린 코드벡터를 조합하여 상기 통합 코드벡터를 생성하고, 상기 통합 코드벡터에 대하여 클린 음성 신호의 게인 값과 상기 추정된 잡음 신호의 게인 값을 적용하여 각각 게인 조합 값을 생성한 다음,

상기 입력 음성 신호의 스펙트럴 엔벨로프 값과 상기 게인 조합 값들 사이의 각각의 유사도를 구한 뒤, 상기 유사도와 게인 조합 값을 ML(Maximum Likelihood) 또는 베이시안 추정 알고리즘(MMSE)에 적용하여 상기 클린 음성 신호의 스펙트럴 엔벨로프 값과 상기 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 18

제17항에 있어서,

상기 음성 합성부는,

위너 필터(Wiener filter)를 통해 상기 클린 음성 신호의 스펙트럴 엔벨로프 값과 상기 잡음 신호의 스펙트럴 엔벨로프 값으로부터 보정된 음성 파워 스펙트럼을 생성하고, 상기 보정된 음성 파워 스펙트럼을 역푸리에 변환시켜 상기 보정된 음성 신호를 추출하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 19

제17항에 있어서,

상기 클린 음성 신호의 게인 값과 상기 추정된 잡음 신호의 게인 값은 음성 부존재 확률(SAP) 알고리즘을 이용하여 획득하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 20

제19항에 있어서,

상기 클린 음성 신호의 게인 값(c_g)과 상기 추정된 잡음 신호의 게인 값(n_g)은 다음의 수학식을 통해 연산되는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치:

$$\begin{bmatrix} \mathbf{C}_g \\ \mathbf{n}_g \end{bmatrix} = \text{SAP} \times \begin{bmatrix} \mathbf{C}_g' \\ \mathbf{n}_g' \end{bmatrix} + (1-\text{SAP}) \times \begin{bmatrix} \mathbf{C}_g'' \\ \mathbf{n}_g'' \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{C}_g' \\ \mathbf{n}_g' \end{bmatrix} = \begin{bmatrix} 0 \\ \| |Y(\omega)|^2 |A_d^n(\omega)|^2 \| \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{C}_g'' \\ \mathbf{n}_g'' \end{bmatrix} = \mathbf{E}^{-1} \cdot \mathbf{F}$$

$$\mathbf{F} = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^4} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_d^n(\omega)|^4} \right\| \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^2 |A_x^m(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^2 |A_d^n(\omega)|^2} \right\| \end{bmatrix}$$

여기서, $\|f(\omega)\| = \int |f(\omega)| d\omega$ 이며, $Y(\omega)$ 는 입력 음성 신호의 주파수 응답 값을 나타내며, $A_d^n(\omega)$ 은 슈퍼 코드북에 저장된 n 번째 코드벡터를 나타내고, $A_x^m(\omega)$ 은 클린 코드북에 저장된 m 번째 코드벡터를 나타낸다.

청구항 21

제12항에 있어서,

상기 입력 음성 신호에 이어서 입력되는 다음 프레임의 입력 음성 신호가 입력되면,

상기 푸리에 변환부는 상기 다음 프레임의 입력 음성 신호를 푸리에 변환시켜 입력 음성 파워 스펙트럼을 생성하고,

상기 잡음 신호 선형예측계수 추출부는 상기 입력 음성 신호로부터 잡음 신호를 추정하고, 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하며,

상기 제어부는,

상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 새로운 적응형 코드벡터로 설정하며, 상기 슈퍼 코드북에 저장된 상기 N 개의 적응형 코드벡터 중에서 적어도 하나의 적응형 코드벡터를 제거하고, 상기 새로운 적응형 코드벡터를 상기 슈퍼 코드북에 저장하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

청구항 22

제21항에 있어서,

상기 제어부는,

상기 슈퍼 코드북에 저장된 상기 N개의 적응형 코드벡터 중에서 저장된 시기가 가장 빠른 적응형 코드벡터를 제거하는 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치.

명세서

기술분야

[0001] 본 발명은 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법 및 그 장치에 관한 것이다. 구체적으로는 적응형 코드벡터를 이용하여 학습한 잡음과 학습하지 않은 잡음, 모든 경우에 대해서 성능을 보장할 수 있는 코드북 기반 음성 향상 방법 및 그 장치에 관한 것이다.

배경기술

[0002] 음성향상 기법은 휴대기기를 이용한 통신 및 음악정보 처리 분야는 물론이고 로봇제어를 위한 음성인식 등 다양한 분야에서 필요로 하는 기술이며, 특히 음성향상 기법의 성능을 향상시키기 위해서는 잡음을 제거하는 기술이 중요하다.

[0003] 잡음을 제거하는 기술에는 잡음으로 오염된 음성이 깨끗한 음성과 배경 잡음의 합이라는 가정에서 시작된 스펙트럴 제거(spectral subtraction; SS) 방법, 잡음의 변화가 주파수 밴드 별로 독립적이라는 가정을 이용하여 성능을 개선한 MBSS(multi-band spectral subtraction) 방법 등이 있다. 또한 잡음 추정 알고리즘은 음성 인식(voice activity detection; VAD)을 이용한 방법이나 최소 통계(Minimum statistics; MS)를 이용한 방법 등 다양한 알고리즘이 연구되어 있다.

[0004] 또한 단기 예측계수 코드북 기법(Codebook Driven Short-Term Predictor parameter estimation; CDSTP)과 같은 방법은 음성과 잡음의 선형 예측 계수를 데이터베이스로 이용하고, 입력된 음성신호와의 최대 유사도(Maximum Likelihood; ML) 또는 최소 평균 제곱 오류(Minimum Mean Square Error; MMSE) 추정을 통하여 잡음을 제거하고 향상된 음성 신호를 얻는다. 이 기법은 코드북에 포함된 다양한 선형 예측 계수를 이용하기 때문에 정상(stationary) 배경 잡음 환경뿐만 아니라 비정상(non-stationary) 배경 잡음 환경에서도 뛰어난 성능을 보인다.

[0005] 이와 같이 기존 코드북 기반 음성향상 기법은 미리 학습된 음성과 잡음의 선형 예측 계수를 데이터베이스로 이용하고, 입력신호와의 ML 또는 MMSE 추정을 통하여 향상된 음성 신호를 얻는다. 그러나, 이와 같은 기법은 코드북에 포함된 다양한 선형 예측 계수를 이용하기 때문에 기존의 잡음 코드북(fixed codebooks) 내에 해당 잡음이 있는 경우에는 뛰어난 성능을 보이지만 학습되지 않은 잡음에 관해서는 성능을 보장하지 못하는 문제점이 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명의 목적은 학습한 잡음과 학습하지 않은 잡음, 모든 경우에 대해서 성능을 보장할 수 있는 적응형 코드북을 이용한 코드북 기반 음성 향상 방법 및 그 장치를 제공하는 것이다.

과제의 해결 수단

[0007] 상기한 바와 같은 목적을 달성하기 위한 본 발명의 실시예에 따른 적응형 코드벡터를 이용한 코드북 기반 음성 향상 방법은, 클린 음성 신호와 잡음 신호가 혼합된 입력 음성 신호를 프레임 단위로 입력받는 단계, 상기 입력 음성 신호를 푸리에 변환시켜 입력 음성 파워 스펙트럼을 생성하는 단계, 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계, 상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하고, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계, 기 저장된 복수의 잡음 코드북들 중에서, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값과 가장 유사한 잡음 코드벡터를 가지고 있는 잡음 코드북을 선택하는 단계, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N개를 적응형 코드벡터로 설정하고, 상기 N개의 적응형 코드벡터를 상기 선택된 잡음 코드북에 저장시켜 슈퍼 코드북을 생성하는 단계, 그리고 상기 슈퍼 코드북 내의 잡음 코드벡터들과 클린 코드북 내의 클린 코드벡터들을 조합한 통합 코드벡터와 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 이용하여 보정된 음성 신호를 추출하는 단계를 포함한다.

[0008] 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계는, 상기 입력 음성 파워 스펙트럼을 역푸리에 변환시켜 상기 입력 음성 신호의 자기상관 값을 생성하는 단계, 상기 입력 음성 신호의 자기상관 값을 레빈슨 알고리즘(Levinson algorithm)에 적용하여, 상기 입력 음성 신호에 대한 선형예측계수를 추출하는 단계, 그리고 상기 입력 음성 신호에 대한 선형예측계수의 주파수 응답 값을 이용하여 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계를 포함할 수 있다.

[0009] 상기 입력 음성 신호로부터 잡음 신호를 추정하고, 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계는, Long-term 잡음 추정 알고리즘을 이용하여 상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하는 단계, 상기 추정된 잡음 신호의 잡음 타입을 이용하여, 상기 기 저장된 복수의 잡음 코드북들 중에서 1개의 잡음 코드북을 선택하는 단계, 상기 추정된 잡음 신호에 대한 잡음 파워 스펙트럼을 역푸리에 변환시켜 상기 추정된 잡음 신호의 자기 상관 값을 생성하는 단계, 상기 추정된 잡음 신호의 자기 상관 값을 레빈슨 알고리즘에 적용하여, 상기 추정된 잡음 신호에 대한 선형예측계수를 추출하는 단계, 그리고 상기 추정된 잡음 신호에 대한 선형예측계수의 주파수 응답 값을 이용하여 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 단계를 포함할 수 있다.

[0010] 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값과 가장 유사한 잡음 코드백터를 가지고 있는 잡음 코드북을 선택하는 단계는, 다음의 수학적식과 같은 IS-D(Itakura-Saito Distortion) 알고리즘을 이용하여 상기 잡음 코드북을 선택할 수 있다.

$$n' = \arg \max_n \left\{ \max_{1 \leq j \leq \nu_n} p(\hat{\mathbb{D}} | a_d^{n,j}) \right\}, 1 \leq n \leq N_d$$

[0011] 여기서 n' 는 선택된 잡음 코드북, $\hat{\mathbb{D}} = [d[0], d[1], \dots, d[K-1]]^T$ 는 추정된 잡음신호, N_d 는 코드북 데이터베이스 내 잡음 코드북의 개수, ν_n 은 n 번째 코드북 내 코드백터의 개수이고, $a_d^{n,j}$ 은 n 번째 잡음 코드북의 j 번째 코드백터의 선형 예측 계수이다.

[0013] 상기 슈퍼 코드북을 생성하는 단계는, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N 개를 상기 적응형 코드백터로 설정하여 상기 선택된 잡음 코드북에 저장시키는 단계, 그리고 상기 선택된 잡음 코드북에 상기 저장된 잡음 코드백터와 상기 N 개의 적응형 코드백터를 통합하여 상기 선택된 잡음 코드북을 상기 슈퍼 코드북으로 변경시키는 단계를 포함할 수 있다.

[0014] 상기 보정된 음성 신호를 추출하는 단계는, 상기 슈퍼 코드북에 포함되는 잡음 코드백터와 상기 클린 코드북에 포함되는 클린 코드백터를 조합하여 상기 통합 코드백터를 생성하는 단계, 상기 통합 코드백터에 대하여 클린 음성 신호의 게인 값과 상기 추정된 잡음 신호의 게인 값을 적용하여 각각 게인 조합 값을 생성하는 단계, 상기 입력 음성 신호의 스펙트럴 엔벨로프 값과 상기 게인 조합 값들 사이의 각각의 유사도를 구하는 단계, 그리고 상기 유사도와 게인 조합 값을 ML(Maximum Likelihood) 또는 베이시안 추정 알고리즘(MMSE)에 적용하여 상기 클린 음성 신호의 스펙트럴 엔벨로프 값과 상기 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계를 포함할 수 있다.

[0015] 상기 보정된 음성 신호를 생성하는 단계는, 위너 필터(Wiener filter)를 통해 상기 클린 음성 신호의 스펙트럴 엔벨로프 값과 상기 잡음 신호의 스펙트럴 엔벨로프 값으로부터 보정된 음성 파워 스펙트럼을 생성하는 단계, 그리고 상기 보정된 음성 파워 스펙트럼을 역푸리에 변환시켜 상기 보정된 음성 신호를 추출하는 단계를 더 포함할 수 있다.

[0016] 상기 클린 음성 신호의 게인 값과 상기 추정된 잡음 신호의 게인 값은 음성 부존재 확률(SAP) 알고리즘을 이용하여 획득할 수 있다.

[0017] 상기 클린 음성 신호의 게인 값(c_g)과 상기 추정된 잡음 신호의 게인 값(n_g)은 다음의 수학적식을 통해 연산될 수 있다.

$$\begin{bmatrix} \mathbf{C}_g \\ \mathbf{n}_g \end{bmatrix} = \text{SAP} \times \begin{bmatrix} \mathbf{C}_g' \\ \mathbf{n}_g' \end{bmatrix} + (1-\text{SAP}) \times \begin{bmatrix} \mathbf{C}_g'' \\ \mathbf{n}_g'' \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{C}_g' \\ \mathbf{n}_g' \end{bmatrix} = \begin{bmatrix} 0 \\ \| |Y(\omega)|^2 |A_d^n(\omega)|^2 \| \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{C}_g'' \\ \mathbf{n}_g'' \end{bmatrix} = \mathbf{E}^{-1} \cdot \mathbf{F}$$

[0018]

$$\mathbf{E} = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^4} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_d^n(\omega)|^4} \right\| \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^2 |A_x^m(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^2 |A_d^n(\omega)|^2} \right\| \end{bmatrix}$$

[0019]

[0020] 여기서, $\|f(\omega)\| = \int |f(\omega)| d\omega$ 이며, $Y(\omega)$ 는 입력 음성 신호의 주파수 응답 값을 나타내며, $A_d^n(\omega)$ 은 슈퍼 코드북에 저장된 n 번째 코드벡터를 나타내고, $A_x^m(\omega)$ 은 클린 코드북에 저장된 m 번째 코드벡터를 나타낸다.

[0021] 상기 입력 음성 신호에 이어서 입력되는 다음 프레임의 입력 음성 신호로부터 잡음 신호를 추정하고, 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 단계, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 새로운 적응형 코드벡터로 설정하는 단계, 그리고 상기 슈퍼 코드북에 저장된 상기 N 개의 적응형 코드벡터 중에서 적어도 하나의 적응형 코드벡터를 제거하고, 상기 새로운 적응형 코드벡터를 상기 슈퍼 코드북에 저장하는 단계를 더 포함할 수 있다.

[0022] 상기 슈퍼 코드북에 저장된 상기 N 개의 적응형 코드벡터 중에서 저장된 시기가 가장 빠른 적응형 코드벡터를 제거할 수 있다.

[0023] 본 발명의 다른 실시예에 따른 적응형 코드벡터를 이용한 코드북 기반 음성 향상 장치는, 클린 음성 신호와 잡음 신호가 혼합된 입력 음성 신호를 프레임 단위로 입력받는 음성 신호 입력부, 상기 입력 음성 신호를 푸리에 변환시켜 입력 음성 파워 스펙트럼을 생성하는 푸리에 변환부, 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 구하는 입력 음성 선형예측계수 추출부, 상기 입력 음성 파워 스펙트럼으로부터 상기 잡음 신호를 추정하고, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값을 구하는 잡음 신호 선형예측계수 추출부, 복수의 클린 코드벡터를 가지는 클린 음성 코드북과 복수의 잡음 코드벡터를 가지며 서로 다른 타입을 가지는 복수의 잡음 코드북을 포함하는 코드북 데이터베이스, 상기 복수의 잡음 코드북들 중에서, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값과 가장 유사한 잡음 코드벡터를 가지고 있는 잡음 코드북을 선택하는 코드북 선택부, 상기 추정된 잡음 신호의 스펙트럴 엔벨로프 값들 중에서 임의의 N 개를 적응형 코드벡터로 설정하고, 상기 N 개의 적응형 코드벡터를 상기 선택된 잡음 코드북에 저장시켜 슈퍼 코드북을 생성하는 제어부, 그리고 상기 슈퍼 코드북 내의 잡음 코드벡터들과 클린 코드북 내의 클린 코드벡터들을 조합한 통합 코드벡터와 상기 입력 음성 신호의 스펙트럴 엔벨로프 값을 이용하여 보정된 음성 신호를 추출하는 음성 합성부를 포함한다.

발명의 효과

- [0024] 이와 같이, 본 발명에 따르면 입력신호의 잡음을 MS나 IMCRA(Improved Minima Controlled Recursive Averaging)와 같은 Long-term 잡음 추정 기법을 이용하여 추정된 잡음으로부터 선형예측계수를 추출하여 적응형 코드벡터를 생성하고, 생성된 적응형 코드벡터를 기존의 잡음 코드북에 결합하여 슈퍼 코드북을 생성함으로써, 학습한 잡음과 학습하지 않은 잡음, 모든 경우에 대해서 성능을 보장할 수 있다. 특히, 음성 부존재 확률(SAP)기법을 통해 획득한 클린 음성 신호의 게인 값과 잡음 신호의 게인 값을 이용하여, 더욱 성능이 향상된 음성 신호를 출력할 수 있다.

도면의 간단한 설명

- [0025] 도 1은 본 발명의 실시예에 따른 음성 인식 향상 장치의 구성을 나타내는 도면이다.
 도 2는 도 1에 따른 코드북 데이터베이스의 구성도이다.
 도 3은 본 발명의 실시예에 따른 음성 인식 향상 방법을 설명하기 위한 순서도이다.
 도 4는 본 발명의 실시예에 따른 슈퍼 코드북을 생성하는 과정을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0026] 기타 실시예들의 구체적인 사항들은 상세한 설명 및 도면들에 포함되어 있다.
- [0027] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술 되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0028] 도 1은 본 발명의 실시예에 따른 음성 인식 향상 장치의 구성을 나타내는 도면이고, 도 2는 도 1에 따른 코드북 데이터베이스의 구성도이다.
- [0029] 도 1을 참조하면, 음성 인식 향상 장치(100)는 음성 신호 입력부(110), 푸리에 변환부(120), 입력 음성 선형예측계수 추출부(130), 잡음 신호 선형예측계수 추출부(140), 코드북 데이터베이스(150), 코드북 선택부(160), 제어부(170) 및 음성 합성부(180)를 포함한다.
- [0030] 먼저 음성 신호 입력부(110)는 프레임 단위로 입력 음성 신호($y(n)$)를 수신한다. 여기서 입력 음성 신호는 오염된 음성 신호로서, 클린 음성 신호와 잡음 신호가 혼합된 음성 신호를 나타낸다.
- [0031] 푸리에 변환부(120)는 입력된 음성 신호를 시간 영역 신호에서 주파수 영역 신호로 푸리에 변환(FFT, Fast Fourier Transform)시켜 입력 음성 파워 스펙트럼을 생성한다.
- [0032] 입력 음성 선형예측계수 추출부(130)는 입력 음성 신호의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)을 구한다. 여기서, 입력 음성 선형예측계수 추출부(130)는 입력 음성 파워 스펙트럼($|Y(k)|^2$)을 역푸리에 변환(IFFT, Inverse Fast Fourier Transform)시켜 입력 음성 신호($y(n)$)의 자기상관 값($R_y(n)$)을 생성한 다음, 입력 음성 신호($y(n)$)의 자기상관 값($R_y(n)$)을 레빈슨 알고리즘(Levinson algorithm)에 적용하여 입력 음성 신호($y(n)$)에 대한 선형예측계수(a_y)를 추출한다. 그 다음, 입력 음성 선형예측계수 추출부(130)는 입력 음성 신호($y(n)$)에 대한 선형예측계수(a_y)의 주파수 응답 값을 이용하여 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)을 구한다. 여기서, 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)은 입력 음성 신호($y(n)$)에 대한 선형예측계수의 주파수 응답 값($Y_{LPC}(k)$)의 크기의 제곱을 나타낸다.
- [0033] 잡음 신호 선형예측계수 추출부(140)는 입력 음성 신호($y(n)$)로부터 잡음 신호를 추정하고, 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)을 구한다. 즉, 잡음 신호 선형예측계수 추출부(140)는 IMCRA(Improved Minima Controlled Recursive Averaging) 또는 MS(Minimum Statistic)와 같은 long term 잡음

추정 알고리즘을 이용하여 입력 음성 신호($y(n)$)로부터 잡음 신호를 추정한다. 그리고, 추정된 잡음 신호의 잡음 타입을 이용하여, 복수의 잡음 코드북들 중에서 1개의 잡음 코드북을 선택하며, 상기 추정된 잡음 신호에 대한 잡음 파워 스펙트럼($|D(k)|^2$)을 역푸리에 변환(IFFT)시켜 추정된 잡음 신호($d(n)$)의 자기 상관 값($R_d(n)$)을 생성한다. 그리고, 추정된 잡음 신호($d(n)$)의 자기 상관 값($R_d(n)$)을 레빈슨 알고리즘(Levinson algorithm)에 적용하여, 추정된 잡음 신호($d(n)$)의 선형예측계수(a_d)를 추출한다. 그리고 잡음 신호 선형예측계수 추출부(140)는 추정된 잡음 신호($d(n)$)에 대한 선형예측계수(a_d)의 주파수 응답 값을 이용하여 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)을 구한다.

[0034] 코드북 데이터베이스(150)는 도 2와 같이 클린 음성 코드북(151)과 복수의 잡음 코드북(152a, 152b, 152c, 152d, 152e)을 포함한다. 클린 음성 코드북(151)은 복수의 클린 코드벡터를 포함하며, 복수의 잡음 코드북(152)은 각각 서로 다른 잡음의 타입을 가지는데, 각각의 잡음 코드북(152)은 복수의 잡음 코드벡터를 포함한다. 본 발명의 실시예에서는 설명의 편의상 코드북 데이터베이스(150)는 1개의 클린 음성 코드북(151)과 5개의 잡음 코드북(152)으로 이루어지며, 5개의 잡음 코드북(152a, 152b, 152c, 152d, 152e)은 각각 잡음의 타입에 따라 White, Babble, F16, Factory2, Pink를 포함하는 것으로 가정한다. 또한 클린 음성 코드북(151)은 1024개의 클린 코드벡터로 이루어져 있으며, 잡음 코드북(152a, 152b, 152c, 152d, 152e)은 학습된 잡음(noise)에 대한 코드북으로서, 각각 8개의 잡음 코드벡터(fixed codevector)를 저장하고 있는 것으로 가정한다.

[0035] 다음으로 코드북 선택부(160)는 복수의 잡음 코드북들 중에서, 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)과 가장 유사한 잡음 코드벡터를 가지고 있는 잡음 코드북을 선택한다.

[0036] 그리고, 제어부(170)는 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)들 중에서 임의의 N개를 적응형 코드벡터(adaptive codevector)로 설정하고, N개의 적응형 코드벡터를 상기 선택된 잡음 코드북에 저장시켜 슈퍼 코드북을 생성한다.

[0037] 또한 음성 합성부(180)는 슈퍼 코드북 내의 잡음 코드벡터들과 클린 코드북 내의 클린 코드벡터들을 조합한 통합 코드벡터와 상기 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)을 이용하여 음성인식이 향상된 보정 음성 신호를 추출한다.

[0038] 이하에서는 도 3 및 도 4를 통하여 본 발명의 실시예에 따른 음성 인식 향상 방법을 더욱 상세하게 설명한다. 도 3은 본 발명의 실시예에 따른 음성 인식 향상 방법을 설명하기 위한 순서도이고, 도 4는 본 발명의 실시예에 따른 슈퍼 코드북을 생성하는 과정을 설명하기 위한 도면이다.

[0039] 도 3에 나타난 것처럼, 먼저 음성 신호 입력부(110)는 외부로부터 프레임 단위로 입력 음성 신호($y(n)$)를 수신한다(S310). 여기서, 입력되는 음성 신호($y(n)$)는 클린 음성($x(n)$)과 잡음 신호($d(n)$)가 섞인 신호로서, 잡음으로 오염된 음성 신호($y(n)$)는 다음의 수학적 식 1과 같이 나타낼 수 있다.

수학적 식 1

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

[0041] 여기서 $x(n)$ 은 클린 음성 신호이며 $d(n)$ 은 잡음 신호이다.

[0042] 외부로부터 입력 음성 신호($y(n)$)가 수신되면, 푸리에 변환부(120)는 시간 도메인의 입력 음성 신호($y(n)$)를 푸리에 변환(FFT) 시켜서 주파수 도메인의 음성 신호($Y(k)$)를 생성하고, 음성 파워 스펙트럼(Noisy power spectrum, $|Y(k)|^2$)을 생성한다(S320).

[0043] 그리고, 입력 음성 선형예측계수 추출부(130)는 음성 파워 스펙트럼($|Y(k)|^2$)을 역푸리에 변환(Inverse Fast

Fourier Transform, IFFT)시켜 입력 음성 신호($y(n)$)의 자기 상관 값에 해당하는 $R_y(n)$ 을 생성한다(S330).

[0044] 입력 음성 선형예측계수 추출부(130)는 $R_y(n)$ 을 레빈슨 알고리즘(Levinson algorithm)에 적용하여, 입력 음성 신호($y(n)$)의 선형예측계수(a_y)를 추출한다(S340). 레빈슨 알고리즘을 이용하여 입력 음성 신호($y(n)$)로부터 선형예측계수(a_y)를 추출하는 방식은 당업자라면 용이하게 실시할 수 있는 것이므로 상세한 설명은 생략한다. 이와 같이 입력 음성 신호에 대한 선형예측계수(a_y)를 추출하면, 입력 음성 선형예측계수 추출부(130)는 입력 음성 신호($y(n)$)의 선형예측계수(a_y)를 이용하여 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)을 구한다(S350).

[0045] 한편, 잡음 신호 선형예측계수 추출부(140)는 IMCRA(Improved Minima Controlled Recursive Averaging) 또는 MS(Minimum Statistic)와 같은 Long-term 잡음 추정 알고리즘을 이용하여 음성 파워 스펙트럼($|Y(k)|^2$)로부터 입력 음성 신호($y(n)$)에 포함된 잡음 신호($d(n)$)를 추정한다(S360). 여기서, 추정된 잡음 신호($d(n)$)의 잡음 타입과 가장 유사한 잡음 코드북을 복수의 잡음 코드북들 중에서 선택한다.

[0046] 잡음 신호 선형예측계수 추출부(140)는 추정된 잡음 신호($d(n)$)로부터 잡음 파워 스펙트럼($|D(k)|^2$)을 생성하고, 추정된 잡음 파워 스펙트럼($|D(k)|^2$)을 역푸리에 변환(IFT)시켜 추정된 잡음 신호($d(n)$)의 자기 상관 값($R_d(n)$)을 생성한다(S370).

[0047] 잡음 신호 선형예측계수 추출부(140)는 추정된 잡음 신호($d(n)$)의 자기 상관 값($R_d(n)$)을 레빈슨 알고리즘(Levinson algorithm)에 적용하여, 추정된 잡음 신호($d(n)$)의 선형예측계수(a_d)를 추출한다(S380). 그리고 잡음 신호 선형예측계수 추출부(140)는 추정된 잡음 신호($d(n)$)의 선형예측계수(a_d)를 이용하여 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)을 구한다(S390).

[0048] 추정된 잡음 신호($d(n)$)의 P차 선형예측계수 $a_d = [1, a_d(1), \dots, a_d(P)]^T$ 로 나타낼 수 있으며, 주파수 도메인에서는 다음의 수학식 2와 같이 스펙트럼 형태로 나타낼 수 있다.

수학식 2

$$A_d(\omega) = \sum_{k=0}^P a_d[k] e^{-j\omega k}$$

[0049]

[0050] 이와 같이 추정된 잡음 신호($d(n)$)에 대한 선형예측계수(a_d)를 추출한 뒤, 잡음 신호 선형예측계수 추출부(140)는 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)을 구한다.

[0051] 다음으로 코드북 선택부(160)는 코드북 데이터베이스(150)에 저장된 복수의 잡음 코드북(152a, 152b, 152c, 152d, 152e)들 중에서, 상기 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)과 가장 유사한 잡음 코드북을 가지고 있는 잡음 코드북을 선택한다(S400).

[0052] 여기서, 코드북 선택부(140)는 IS-D(Itakura-Saito Distortion) 알고리즘을 이용하여 잡음신호의 코드북 타입을 선정하는데, IS-D를 이용하는 잡음신호의 코드북 타입의 선정은 다음의 수학식 3에 의해 이루어진다.

수학식 3

$$n' = \arg \max_n \left\{ \max_{1 \leq j \leq v_n} p(\hat{\mathbb{D}} | a_d^{n,j}) \right\}, 1 \leq n \leq N_d$$

[0053]

[0054] 여기서 n' 는 선택된 잡음 코드북, $\hat{\mathbb{D}} = [d[0], d[1], \dots, d[K-1]]^T$ 는 추정된 잡음신호, N_d 는 코드북 데이터베이스 내 잡음 코드북의 개수, v_n 은 n 번째 코드북 내 코드벡터의 개수이고, $a_d^{n,j}$ 은 n 번째 잡음 코드북의 j 번째 코드벡터의 선형 예측 계수이다. 따라서, 도 2에서 예를 든 것에 따르면, N_d 는 5이고 v_n 은 8에 해당한다.

[0055] 그리고, 제어부(170)는 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)들 중에서 임의의 N 개를 적응형 코드벡터(adaptive codevector)로 설정하고, N 개의 적응형 코드벡터를 선택된 잡음 코드북(n')에 저장시켜 슈퍼 코드북을 생성한다(S410).

[0056] 여기서, 추정된 잡음 신호($d(n)$)의 스펙트럴 엔벨로프 값($|D_{LPC}(k)|^2$)의 개수(N)는 임의로 설정이 가능하며, 본 발명의 실시예에서는 설명의 편의상 $N=4$ 개로 가정한다. 또한 설명의 편의상 선택된 잡음 코드북(n')을 White 타입의 잡음 코드북이라고 가정하면, 도 4와 같이 White 타입의 잡음 코드북(152a)에는 기존의 8개의 잡음 코드벡터(fixed code vector)에 4개의 적응형 코드벡터(adaptive code vector)가 추가된다. 따라서, White 타입의 잡음 코드북(152a)는 12개의 잡음 코드벡터를 포함하는 슈퍼 코드북으로 변경된다.

[0057] 다음으로 음성 합성부(170)는 슈퍼 코드북 내의 잡음 코드벡터들과 클린 코드북(151) 내의 클린 코드벡터들을 조합한 통합 코드벡터와 상기 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)을 이용하여 보정된 음성 신호($x''(n)$)를 출력한다(S420).

[0058] 더욱 상세하게 설명하면, 먼저 음성 합성부(170)는 S410 단계에서 생성된 슈퍼 코드북(즉, White 타입의 잡음 코드북(152a))에 포함되는 12개의 잡음 코드벡터와 클린 코드북에 포함되는 1024개의 클린 코드벡터를 조합하여 통합 코드벡터를 생성한다. 즉, 클린 코드북에 저장된 클린 코드벡터들의 개수가 1024개이고, 선택된 슈퍼 코드북에 저장된 잡음 코드벡터들의 개수가 12개이므로, 클린 코드벡터와 잡음 코드벡터가 조합된 통합 코드벡터의 경우의 수는 총 12288개가 가능하다.

[0059] 그리고, 음성 합성부(170)는 12288개의 통합 코드벡터들과 S320단계에서 추정된 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)의 유사도를 IS-D를 이용하여 측정하여, 입력 음성 신호에 포함된 클린 음성 신호와 잡음 신호를 추정한다.

[0060] 유사도를 구하는 과정을 더욱 상세하게 설명하면, 입력 음성 신호($y(n)$) 중 클린 음성 신호($x(n)$)의 계인을 c_g 라 하고 잡음 신호($d(n)$)의 계인을 n_g 라고 하면, 다음의 수학식 4와 같이 12288개의 통합 코드벡터에 대하여 계인 조합 값(A_1 내지 A_{12288})을 구할 수 있다.

수학식 4

$$\left\{ \begin{array}{l} C_g \times C_{\textcircled{1}} + n_g \times n_{\textcircled{1}} = A_1 \\ C_g \times C_{\textcircled{1}} + n_g \times n_{\textcircled{2}} = A_2 \\ \vdots \\ C_g \times C_{\textcircled{1024}} + n_g \times n_{\textcircled{11}} = A_{12287} \\ C_g \times C_{\textcircled{1024}} + n_g \times n_{\textcircled{12}} = A_{12288} \end{array} \right.$$

[0061]

[0062] 상기 수학식 4와 같이 클린 코드북의 1024개의 코드벡터($C_{\textcircled{1}}, C_{\textcircled{2}}, \dots$)와 잡음 코드북(즉, 슈퍼 코드북)의 12개의 코드벡터($n_{\textcircled{1}}, n_{\textcircled{2}}, \dots, n_{\textcircled{11}}, n_{\textcircled{12}}$) 사이에는 12288개의 계인 조합 값(A_1 내지 A_{12288})이 생성된다.

수학식 4에서 클린 음성 신호($x(n)$)의 계인 값(C_g)들과 잡음 신호($d(n)$)의 계인 값(n_g)들은 동일하게 표현하였으나, 각각 다른 값을 가질 수 있으며, 조합에 따라 계인 값(C_g, n_g)들이 다르게 산출될 수 있다.

[0063] 그러면, 음성 합성부(170)는 하나의 프레임에 해당하는 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)과 계인 조합 값(A_1 내지 A_{12288}) 사이의 각각의 유사도(L_1 내지 L_{12288})를 IS-D를 이용하여 구한다.

IS-D를 이용하여 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)과 계인 조합 값 사이의 유사도(L , Likelyhood)를 구하는 방식은 당업자라면 용이하게 실시할 수 있으므로 이에 대한 상세한 설명은 생략한다.

[0064] 그런 다음, 음성 합성부(170)는 ML(Maximum Likelihood) 기법 또는 베이시안 추정 기법(Bayesian Minimum Mean Squared Error Estimation (MMSE))을 통해 입력 음성 신호로부터 클린 음성 신호와 잡음 신호를 구분한다.

[0065] 이를 더욱 상세하게 설명하면, 먼저 음성 합성부(170)는 다음의 수학식 5와 같이 입력 음성 신호($y(n)$)의 스펙트럴 엔벨로프 값($|Y_{LPC}(k)|^2$)을 구한다.

수학식 5

$$\begin{aligned}
 \text{입력 음성 신호의} &= \frac{L_1}{L_1 + L_2 + \dots + L_{12288}} A_1 \\
 \text{스펙트럴 엔벨로프 값} &+ \frac{L_2}{L_1 + L_2 + \dots + L_{12288}} A_2 \\
 &+ \vdots \\
 &+ \frac{L_{12287}}{L_1 + L_2 + \dots + L_{12288}} A_{12287} \\
 &+ \frac{L_{12288}}{L_1 + L_2 + \dots + L_{12288}} A_{12288}
 \end{aligned}$$

[0066]

[0067] 수학식 5에 나타난 A_1 내지 A_{12288} 는 수학식 4에서 구한 계인 조합 값이기 때문에, 이를 수학식 5에 대입하면 다음의 수학식 6과 같이 표현될 수 있다.

수학식 6

$$\begin{aligned}
 & \frac{L_1}{L_1 + L_2 + \dots + L_{12288}} (C_g \times C_{\textcircled{1}} + n_g \times n_{\textcircled{1}}) \\
 & + \\
 & \frac{L_2}{L_1 + L_2 + \dots + L_{12288}} (C_g \times C_{\textcircled{1}} + n_g \times n_{\textcircled{2}}) \\
 & + \\
 & \vdots \\
 & + \\
 & \frac{L_{12287}}{L_1 + L_2 + \dots + L_{12288}} (C_g \times C_{\textcircled{1024}} + n_g \times n_{\textcircled{11}}) \\
 & + \\
 & \frac{L_{12288}}{L_1 + L_2 + \dots + L_{12288}} (C_g \times C_{\textcircled{1024}} + n_g \times n_{\textcircled{12}}) \\
 \\
 = & \frac{L_1 (C_g \times C_{\textcircled{1}}) + L_2 (C_g \times C_{\textcircled{1}}) + \dots + L_{12288} (C_g \times C_{\textcircled{1024}})}{\sum_{i=1}^{12288} L_i} \\
 + & \frac{L_1 (n_g \times n_{\textcircled{1}}) + L_2 (n_g \times n_{\textcircled{2}}) + \dots + L_{12288} (n_g \times n_{\textcircled{12}})}{\sum_{i=1}^{12288} L_i}
 \end{aligned}$$

[0068]

[0069] 수학식 6에서 구한 입력 음성 신호(y(n))의 스펙트럴 엔벨로프 값을 클린 음성 신호의 스펙트럴 엔벨로프 값(X')과 추정된 잡음 신호의 스펙트럴 엔벨로프 값(d')으로 구분하면 수학식 7과 같이 나타낼 수 있다.

수학식 7

$$\begin{aligned}
 \text{클린 음성 신호의} & \\
 \text{스펙트럴 엔벨로프 값(X')} &= \frac{L_1 (C_g \times C_{\textcircled{1}}) + L_2 (C_g \times C_{\textcircled{1}}) + \dots + L_{12288} (C_g \times C_{\textcircled{1024}})}{\sum_{i=1}^{12288} L_i} \\
 \\
 \text{잡음 신호의} & \\
 \text{스펙트럴 엔벨로프 값(d')} &= \frac{L_1 (n_g \times n_{\textcircled{1}}) + L_2 (n_g \times n_{\textcircled{2}}) + \dots + L_{12288} (n_g \times n_{\textcircled{12}})}{\sum_{i=1}^{12288} L_i}
 \end{aligned}$$

[0070]

[0071] 그런 다음, 음성 합성부(170)는 다음의 수학식 8과 같이 추정된 클린 음성 신호의 스펙트럴 엔벨로프 값(X')과 잡음 신호의 스펙트럴 엔벨로프 값(d')으로 이루어진 위너 필터(Wiener filter, $X'/(X'+d')$)에 입력 음성 신호에 대한 입력 음성 파워 스펙트럼(y)을 곱함으로써, 추정된 잡음 신호 부분이 제거되어, 보정된 음성 파워 스펙트럼(X'')을 추정할 수 있다. 여기서, 보정된 음성 파워 스펙트럼(X'')이란 입력 음성 신호에서 추정된 잡음 신

호가 제거되어 클린 음성 신호의 비율이 증가된 상태로서, 음성의 질이 향상된 음성 신호에 대한 파워 스펙트럼을 나타낸다.

수학식 8

$$\frac{x'}{x' + d'} \times y = \frac{x'}{x' + d'} \times (x + d) \approx x''$$

[0072]

[0073] 여기서, y 는 입력 음성 신호에 대한 입력 음성 파워 스펙트럼($|Y(k)|^2$)을 나타내고, x 는 클린 음성 파워 스펙트럼을 나타내며, d 는 잡음 파워 스펙트럼을 나타낸다. 그리고, 음성 합성부(170)는 보정된 음성 파워 스펙트럼(X'')을 역푸리에 변환(IFFT)시켜 음성의 질이 향상된 음성 신호를 출력시킨다.

[0074] 이와 같이 S420 단계가 완료된 이후에, 음성 신호 입력부(110)에 다음 프레임의 입력 음성 신호가 입력되면, S310 단계부터 S420 단계를 반복하여 음성의 질이 향상된 음성 신호를 출력시키는데, 제어부(170)는 적응형 코드벡터를 갱신시키도록 한다.

[0075] 즉, 제어부(170)는 S310 단계 내지 S400 단계를 거쳐서 새로운 적응형 코드벡터(adaptive codevector)가 생성되면, 슈퍼 코드북(151)에 기 저장되어 있는 N 개의 적응형 코드벡터 중에서 적어도 하나의 적응형 코드벡터를 제거하고, 새로 생성된 적응형 코드벡터를 상기 슈퍼 코드북(151)에 저장시킨다. 여기서, 기 저장되어 있는 N 개의 적응형 코드벡터 중에서 가장 저장 시기가 빠른, 즉 가장 오래된 적응형 코드벡터를 제거하도록 한다.

[0076] 이와 같이 본 발명의 실시예에 따르면 적응형 코드벡터를 프레임마다 갱신을 시키므로, 학습되지 않은 잡음이 섞인 입력 음성에 대해서도 효과적으로 음성의 질이 향상된 보정 음성 신호를 추출할 수 있다.

[0077] 한편, 본 발명의 실시예에 따르면, 수학식 4에서 음성 부존재 확률(SAP, Speech Absence Probability) 기법을 이용하면, 클린 음성 신호의 게인 값(c_g)과 잡음 신호의 게인 값(n_g)을 더욱 정확하게 추정할 수 있다. 음성 부존재 확률(SAP)을 이용하는 잡음 추정 모델은 음성이 존재하는 구간과 음성이 존재하지 않은 구간으로 나누어 각 구간에 대하여 별개로 각 게인 값을 연산하여 합산하는 방식을 가진다.

[0078] 먼저, 잡음 신호 선형예측계수 추출부(140)는 S360 단계에서 Long-term 잡음 추정 기법을 이용하여 잡음 신호($d(n)$)를 추정하면서 음성 부존재 확률 값(SAP)과 음성 존재 확률 값(1-SAP)을 산출한다. 그러면, 음성 합성부(170)는 산출된 음성 부존재 확률 값(SAP)과 음성 존재 확률 값(1-SAP)을 다음의 수학식 9에 적용하여 클린 음성 신호의 게인 값(c_g)과 잡음 신호의 게인 값(n_g)을 추정한다.

수학식 9

$$\begin{bmatrix} C_g \\ n_g \end{bmatrix} = \text{SAP} \times \begin{bmatrix} C_g' \\ n_g' \end{bmatrix} + (1-\text{SAP}) \times \begin{bmatrix} C_g'' \\ n_g'' \end{bmatrix}$$

$$\begin{bmatrix} C_g' \\ n_g' \end{bmatrix} = \begin{bmatrix} 0 \\ \| |Y(\omega)|^2 |A_d^n(\omega)|^2 \| \end{bmatrix}$$

$$\begin{bmatrix} C_g'' \\ n_g'' \end{bmatrix} = E^{-1} \cdot F$$

[0079]

[0080] 여기서, $Y(\omega)$ 는 입력 음성 신호의 주파수 응답 값을 나타내며, $A_d^n(\omega)$ 은 슈퍼 코드북에 저장된 n 번째 코드벡터를 나타낸다.

[0081] 여기서, 음성 존재 구간에서의 계인 값은 E 의 역행렬과 F 의 행렬 곱으로 나타낼 수 있으며, E 와 F 는 다음과 같다.

수학식 10

$$E = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^4} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^4 |A_x^m(\omega)|^2 |A_d^n(\omega)|^2} \right\| & \left\| \frac{1}{|Y(\omega)|^4 |A_d^n(\omega)|^4} \right\| \end{bmatrix}$$

$$F = \begin{bmatrix} \left\| \frac{1}{|Y(\omega)|^2 |A_x^m(\omega)|^2} \right\| \\ \left\| \frac{1}{|Y(\omega)|^2 |A_d^n(\omega)|^2} \right\| \end{bmatrix}$$

[0082]

[0083] 여기서, $\|f(\omega)\| = \int |f(\omega)| d\omega$ 이며, $A_x^m(\omega)$ 은 클린 코드북에 저장된 m 번째 코드벡터를 나타낸다.

[0084] 따라서, 본 발명의 실시예에 따르면 음성 부존재 확률(SAP, Speech Absence Probability) 기법을 이용하여, 클린 음성 신호의 계인 값(c_g)과 잡음 신호의 계인 값(n_g)을 더욱 정확하게 추정할 수 있으며, 나아가 음성 합성부(170)는 S420 단계에서 클린 음성 신호를 출력할 수 있으므로 더욱 음성의 질이 향상된 음성 신호를 추출시킬 수 있다.

[0085] 다음의 표 1은 본 발명의 실시예에 따른 잡음 추정 방법의 성능을 평가한 실험의 결과를 나타낸 것이다.

[0086] 본 실험예에서는 학습된 잡음인 내부 잡음(Inside noise)와 학습되지 않은 외부 잡음(Outside noise)에 대하여 기존의 IMCRA 방법, 기존의 코드북 기반 음성향상기법(CBSE), 본 발명의 실시예에 따른 슈퍼 코드북을 이용한

음성향상기법(CBSE(S)), 기존의 음성 부존재 확률(SAP) 기법을 이용한 코드북 기반 음성향상기법(SAP-CBSE), 본 발명의 실시예에 따른 음성 부존재 확률(SAP)기법과 슈퍼 코드북을 이용한 음성향상기법(SAP-CBSE(S))의 5가지 방법을 통하여 스펙트럼 왜곡도(LSD, Log-spectral Distortion)에 대한 실험을 진행하였다.

[0087] 특히, 학습된 잡음인 내부 잡음(Inside noise)은 white, babble, F16, factory2, pink에 대한 잡음으로서, 상기 5개의 잡음에 대해서는 잡음 코드북을 저장하고 있으며, 학습되지 않은 외부 잡음(Outside noise)은 Buccaneer1, car, restaurant, factory1, street에 대한 잡음을 포함한다. 본 실험예에서는 음성신호 코드북의 학습은 8000Hz로 샘플링 된 TIMIT 데이터베이스 4620개의 문장으로, 테스트는 1680개의 문장으로 수행하였다. 잡음의 경우 10가지 잡음에 대해서 실험을 하였으며 white, babble, F16, factory2, pink를 기존 잡음 코드북(fixed codebooks)으로 가지고 있는 환경하에서 실험을 수행하였다.

표 1

LSD	Noise Type	IMCRA	CBSE	CBSE (S)	SAP-CBSE	SAP-CBSE (S)
Inside noise	white	8.35	6.86	6.85	5.70	5.71
	babble	5.80	5.63	5.29	5.09	5.03
	F16	5.64	5.18	5.11	4.58	4.52
	factory2	4.17	3.95	3.86	3.71	3.64
	pink	6.20	5.72	5.68	4.90	4.88
	avg	6.03	5.47	5.35	4.80	4.76
Outside noise	Buccaneer1	7.25	7.61	6.64	6.56	5.70
	car	4.56	5.69	5.09	5.21	4.66
	restaurant	6.17	6.83	6.12	6.57	5.63
	factory1	6.23	6.61	6.14	6.05	5.87
	street	5.48	5.37	4.84	5.06	4.64
	avg	5.94	6.42	5.77	5.89	5.30

[0088]

[0089] 실험 결과를 보면 본 발명의 실시예에 따른 슈퍼 코드북을 이용한 음성향상기법(CBSE(S))의 경우가 기존의 IMCRA 방법, 기존의 코드북 기반 음성향상기법(CBSE)에 비하여 우수한 잡음 추정 성능을 보였으며, 본 발명의 실시예에 따른 음성 부존재 확률(SAP)기법과 슈퍼 코드북을 이용한 음성향상기법(SAP-CBSE(S))의 경우가 기존의 SAP를 이용한 코드북 기반 음성향상기법(SAP-CBSE)에 비하여 잡음 추정 성능이 우수한 것으로 나타났다.

[0090] 특히, 본 발명의 실시예에 따른 슈퍼 코드북을 이용한 음성향상기법들(CBSE(S), SAP-CBSE(S))의 경우에는 학습된 잡음인 내부 잡음에서는 기존의 IMCRA 방법, 코드북 기반 음성향상기법들(CBSE, SAP-CBSE)과 유사한 잡음 추정 성능을 가지지만, 학습되지 않은 잡음(Outside noise)에 대해서는 월등히 우수한 잡음 추정 성능을 가지는 것으로 나타났다.

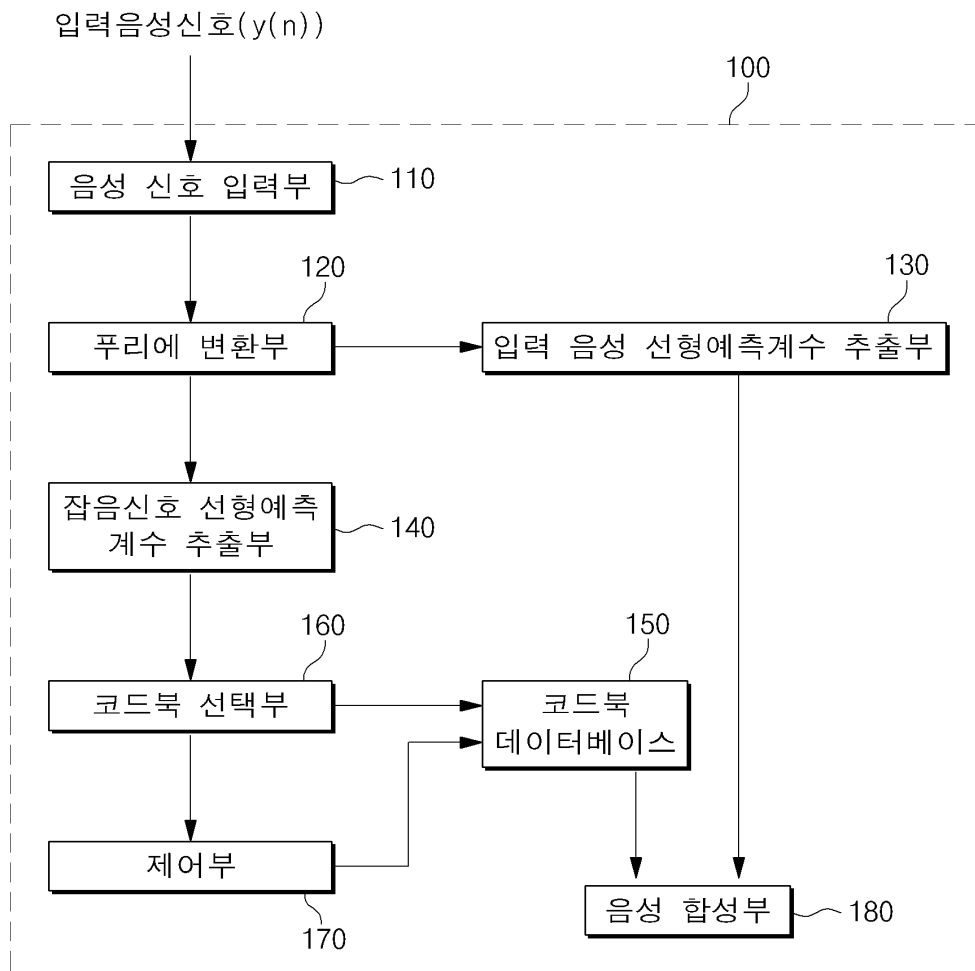
[0091] 이와 같이, 본 발명의 실시예에 따르면 입력신호의 잡음을 MS나 IMCRA와 같은 Long-term 잡음 추정 기법을 이용하여 추정된 잡음으로부터 선형예측계수를 추출하여 적응형 코드벡터를 생성하고, 생성된 적응형 코드벡터를 기존의 잡음 코드북에 결합하여 슈퍼 코드북을 생성함으로써, 학습한 잡음과 학습하지 않은 잡음, 모든 경우에 대해서 성능을 보장할 수 있다. 특히, 음성 부존재 확률(SAP)기법을 통해 획득한 클린 음성 신호의 게인 값(c_g)과 잡음 신호의 게인 값(n_g)을 이용하여, 더욱 성능이 향상된 음성 신호를 출력할 수 있다.

[0092] 본 발명은 첨부된 도면에 도시된 실시예를 참고하여 설명되었으나, 이는 예시적인 것에 불과하며, 본 발명과 관련된 기술 분야에서의 통상적인 지식을 가진 자라면, 이로부터 다양한 변형 또는 균등한 타 실시예가 존재 가능하다는 점을 이해할 것이다.

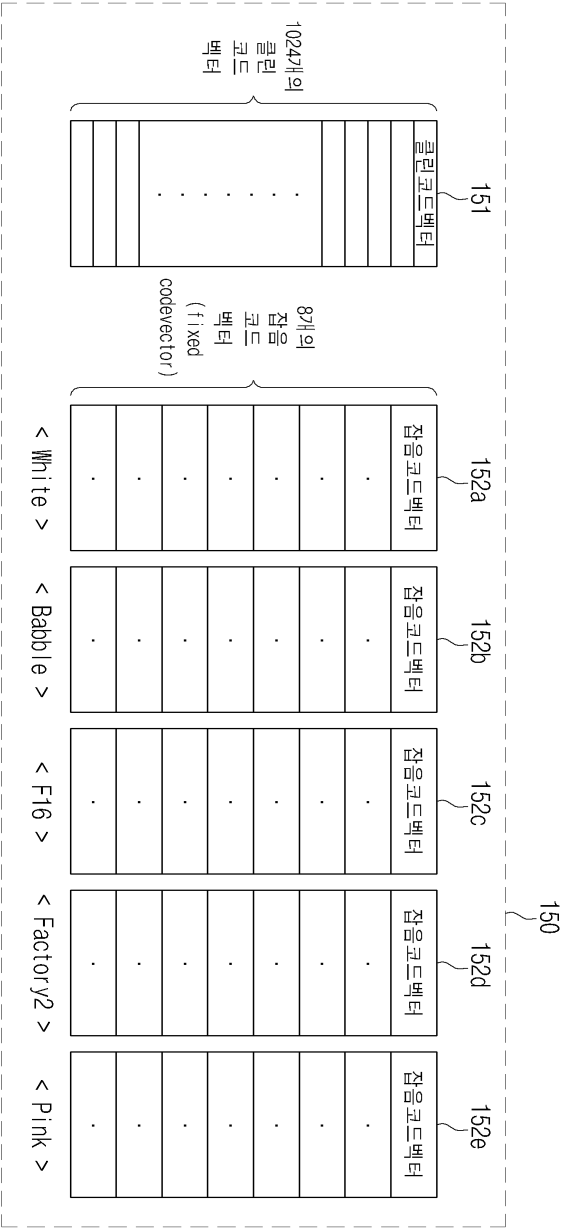
[0093] 따라서 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의해 정해져야 할 것이다.

도면

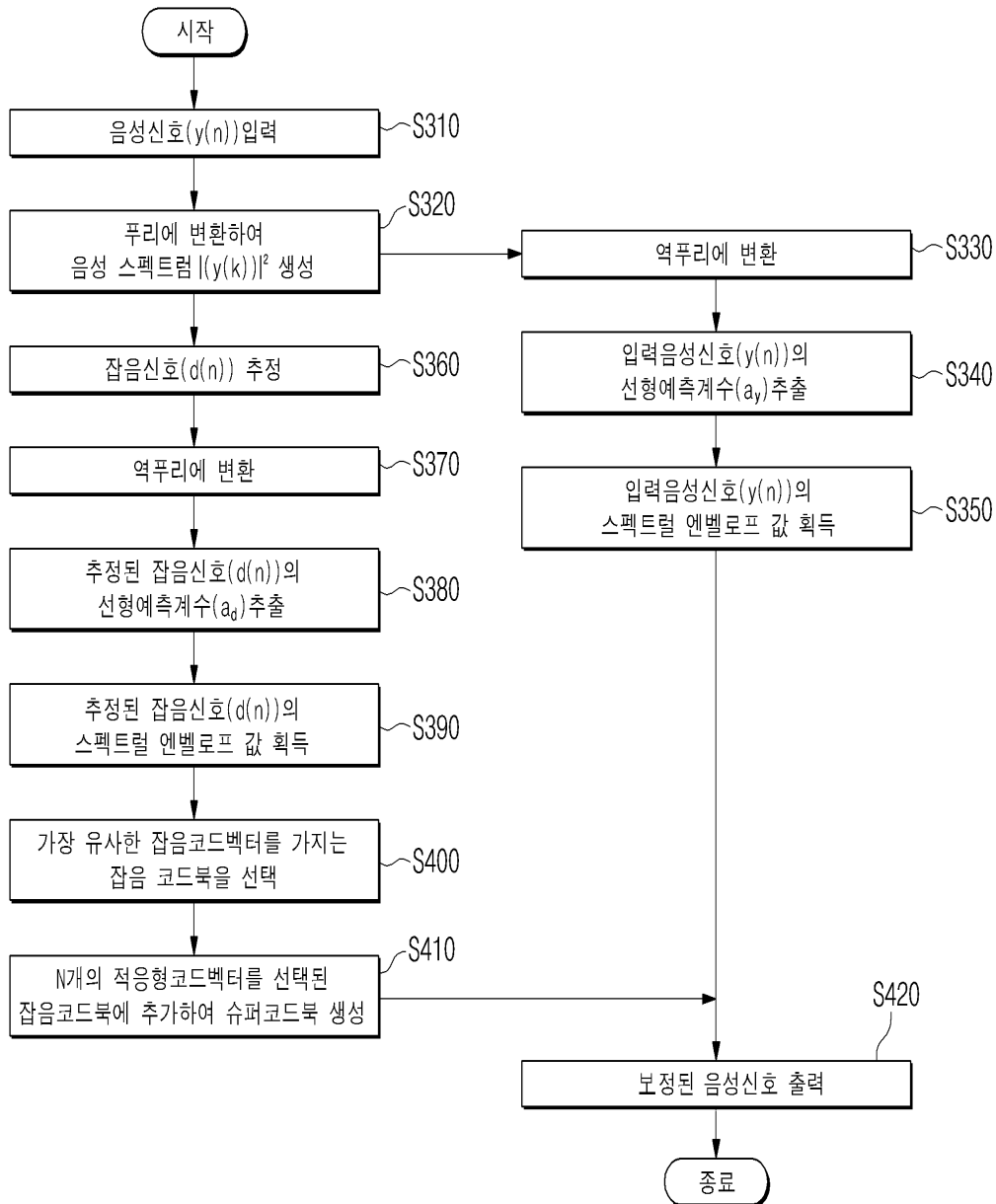
도면1



도면2



도면3



도면4

