(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl.
G10L 13/00 (2006.01)

(45) 공고일자 2006년12월01일
(11) 등록번호 10-0574031
(24) 등록일자 2006년04월19일

(21) 출원번호 10-1998-0044279
(22) 출원일자 1998년10월22일
심사청구일자 2003년09월23일

(65) 공개번호 10-1999-0037291
(43) 공개일자 1999년05월25일

(30) 우선권주장 97-291405 1997년10월23일 일본(JP)

(73) 특허권자 소니 가부시끼 가이샤
일본국 도쿄도 시나가와쿠 키타시나가와 6쵸메 7반 35고

(72) 발명자 오모리 시로
일본국 도쿄도 시나가와구 기다시나가와 6초메 7방 35고 소니가부시키가이샤내

니시구치 마사유키
일본국 도쿄도 시나가와구 기다시나가와 6초메 7방 35고 소니가부시키가이샤내

(74) 대리인 신관호

심사관 : 안준형

전체 청구항 수 : 총 38 항

(54) 음성합성방법및장치그리고음성대역확장방법및장치

(57) 요약

음성대역 확장장치는 광대역 유/무성음에서 각각 추출된 유/무성음 파라미터로부터 형성된 광대역 유/무성음 코드북(12, 14)과 예를 들면 광대역 음성의 대역을 제한함으로써 300~ 3400Hz의 주파수 대역을 가지는 협대역 음성신호에서 각각 추출된 유/무성음 파라미터로부터 형성된 협대역 유/무성음 코드북(8, 10)을 포함하여 구성된다.

대표도

도 1

특허청구의 범위

청구항 1.

복수 종류의 입력 부호화된 파라미터의 음성을 합성하기 위하여 소정된 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 특성 파라미터로부터 형성된 광대역 유성음 코드북과 광대역 무성음 코드북, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성 파라미터로부터 형성된 협대역 유성음 코드북과 협대역 무성음 코드북을 사용하는 음성합성방법에 있어서,

복수의 부호화된 파라미터를 복호화하는 단계와,

복수의 복호화된 첫 번째 파라미터로부터 여진원(innovation)을 형성하는 단계와,

두 번째 복호화된 파라미터를 음성합성 특성파라미터로 변환하는 단계와,

세 번째 복호화된 파라미터를 참조하여 유/무성음을 판별하는 단계와,

협대역 유/무성음 코드북을 사용함으로써 판별의 결과에 기초한 음성합성 특성파라미터를 양자화하는 단계와,

광대역 유/무성음 코드북을 사용함으로써, 협대역 유/무성음 코드북을 사용하여 양자화된 협대역 유/무성음 데이터를 역양자화하는 단계와,

역양자화된 데이터와 여진원을 기초로 한 음성을 합성하는 단계를 포함하여 구성하는 것을 특징으로 하는 음성합성방법.

청구항 2.

제 1항에 있어서,

복수 종류의 부호화된 파라미터는 협대역 음성을 부호화함으로써 얻어지고, 부호화된 첫 번째 파라미터는 여진원에 관련된 파라미터이고, 두 번째 파라미터는 선형예측 파라미터이고, 세 번째 파라미터는 유/무성음 판별 표지(flag)인 것을 특징으로 하는 음성합성방법.

청구항 3.

제 1항에 있어서,

광대역 유/무성음 코드북을 형성하기 위해 실행된 유/무성음 판별은 세 번째 부호화된 파라미터를 사용하는 것과는 다른 것을 특징으로 하는 음성합성방법.

청구항 4.

제 3항에 있어서,

유/무성음을 판별할 수 없는 파라미터를 제외하고, 입력 음성으로부터 광대역 유/무성음 코드북과 협대역 유/무성음 코드북을 형성하기 위해 파라미터를 추출하는 단계를 더 포함하여 구성되는 것을 특징으로 하는 음성합성방법.

청구항 5.

제 1항에 있어서,

자기상관(autocorrelation)이 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 6.

제 1항에 있어서,

켑스트럼(cepstrum)이 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 7.

제 1항에 있어서,

스펙트럼 엔벌로프(envelope)가 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 8.

제 1항에 있어서,

부호화된 첫 번째 파라미터의 피치 성분이 강한 것으로 판단될 때, 임펄스열이 여진원으로써 취해지는 것을 특징으로 하는 음성합성방법.

청구항 9.

복수 종류의 입력 부호화된 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 특성파라미터로부터 미리 형성된 광대역 유성음 코드북과 광대역 무성음 코드북과, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성파라미터로부터 미리 형성된 협대역 유성음 코드북과 협대역 무성음 코드북을 사용하는 음성합성장치에 있어서,

복수의 부호화 파라미터를 복호화하는 수단과,

복호화 수단에 의해 복호화된 복수의 파라미터들 중 첫 번째 파라미터로부터 여진원을 형성하는 수단과,

복호화 수단에 의해 복호화된 부호화 파라미터중 제 2파라미터로부터 음성합성 특성파라미터를 얻는 수단과,

복호화수단에 의해 복호화된 부호화 파라미터의 제 3파라미터를 참조하여 유/무성음을 판별하는 수단과,

협대역 유/무성음 코드북을 사용함으로써, 유/무성음 판별의 결과에 의거한 음성합성 특성파라미터를 양자화하는 수단과,

광대역 유/무성음 코드북을 사용함으로써 유/무성음 양자화 수단으로부터 양자화된 유/무성음 데이터를 역양자화하는 수단과,

광대역 유/무성음 역양자화 수단에서의 역양자화된 데이터와 여진원 형성 수단으로부터의 여진원에 의거하는 음성을 합성하는 수단을 포함하여 구성되어 있는 것을 특징으로 하는 음성합성장치.

청구항 10.

복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위해, 소정의 시간단위마다 광대역 유/무성음에서 추출된 특성파라미터로부터 미리 형성된 광대역 유/무성음 코드북이 사용되는 음성합성방법에 있어서,

복수의 부호화 파라미터를 복호화하는 단계와,

상기 복수의 복호화 파라미터들 중 첫 번째 파라미터로부터 여진원을 형성하는 단계와,

두 번째 복호화된 파라미터를 음성합성 특성파라미터로 변환하는 단계와,

상기 광대역 유/무성음 코드북에서 각 코드 벡터로부터 협대역 특성 파라미터를 연산하는 단계와,

상기 연산 수단에 의해 연산된 협대역 특성파라미터와 비교함으로써 음성합성 특성파라미터를 양자화하는 단계와,

상기 광대역 유/무성음 코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와,

상기 역양자화된 데이터와 여진원에 의거한 음성을 합성하는 단계를 포함하여 구성되어 있는 것을 특징으로 하는 음성합성방법.

청구항 11.

제 10항에 있어서,

복수 종류의 부호화 파라미터는 협대역 음성을 부호화함으로써 얻어지고, 부호화 파라미터의 첫 번째 파라미터는 여진원과 관련된 파라미터이고, 두 번째 파라미터는 선형예측계수이고, 세 번째 파라미터는 유/무성음 판별 표지인 것을 특징으로 하는 음성합성방법.

청구항 12.

제 10항에 있어서,

첫 번째 부호화된 파라미터의 피치 성분이 강한 것으로 판단될 때, 임펄스열은 여진원으로써 취해지는 것을 특징으로 하는 음성합성방법.

청구항 13.

제 10항에 있어서,

자기상관은 특성 파라미터로써 사용되고, 자기상관은 두 번째 부호화 파라미터로부터 생성되고, 자기상관은 상기 광대역 유/무성음 코드북에서의 광대역 자기상관과 대역스톱필터의 임펄스 응답의 자기상관을 컨벌루션에 의해 결정된 협대역 자기상관과 비교함으로써 양자화되고, 양자화된 데이터는 음성을 합성하기 위하여 상기 광대역 유/무성음 코드북을 사용하여 역양자화되는 것을 특징으로 하는 음성합성방법.

청구항 14.

제 10항에 있어서,

광대역 음성코드북은 소정의 시간단위 마다, 분리된 광대역 유/무성음에서 추출된 유/무성음 특성파라미터로부터 미리 형성된 광대역 유/무성음 코드북이고, 복수의 입력 부호화 파라미터의 세 번째 파라미터를 참조하여 판별가능한 유/무성음 판별의 결과에 의거하여, 음성합성 특성 파라미터는 광대역 유/무성음 코드북에서 각 코드북으로부터 연산에 의해 결정된 협대역 특성 파라미터와 비교함으로써 양자화되고, 양자화된 데이터는 광대역 유/무성음 코드북을 사용하여 역양자화되고, 음성은 역양자화된 데이터와 여진원에 의거하여 합성되는 것을 특징으로 하는 음성합성방법.

청구항 15.

제 14항에 있어서,

자기상관은 특성 파라미터로써 사용되고, 자기상관은 제 2부호화 파라미터로부터 생성되고, 자기상관은 광대역 음성코드북에서 광대역 자기상관과 대역스톱필터의 임펄스 응답의 자기상관을 컨벌루션에 의해 결정된 협대역 자기상관과 비교함으로써 양자화되고, 양자화된 데이터는 음성을 합성하기 위하여 광대역 음성코드북을 사용하여 역양자화되는 것을 특징으로 하는 음성합성방법.

청구항 16.

제 14항에 있어서,

광대역 유/무성음 코드북을 형성하기 위해 실행된 유/무성음 판별은 세 번째 부호화 파라미터를 사용하는 것과 다르게 구성되는 것을 특징으로 하는 음성합성방법.

청구항 17.

제 14항에 있어서,

광대역 유/무성음 코드북과 협대역 유/무성음 코드북을 형성하기 위하여, 유/무성음 판별이 불확실한 것을 제외하고, 입력 음성으로부터 파라미터를 추출하는 것을 특징으로 하는 음성합성방법.

청구항 18.

복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위 마다 광대역 음성에서 추출된 특성 파라미터로부터 미리 형성된 광대역 음성코드북을 사용하는 음성합성장치에 있어서,

복수 종류의 부호화 파라미터를 복호화하는 수단과,

복호화 수단에 의해 복호화된 복수 종류의 파라미터들 중 첫 번째 파라미터로부터 여진원을 형성하는 수단과,

복호화 수단에 의해 복호화된 복수 종류의 파라미터의 두 번째 복호화된 파라미터를 음성 합성 특성파라미터로 변환하는 수단과,

광대역 음성코드북에서 각 코드벡터로부터 협대역 특성 파라미터를 연산하는 수단과,

연산수단으로부터 협대역 특성 파라미터를 사용함으로써 파라미터 변환수단으로부터 음성합성 특성파라미터를 양자화하는 수단과,

광대역 음성코드북을 사용함으로써 양자화하는 수단으로부터 양자화된 데이터를 역양자화하는 수단과,

역양자화 수단에서의 역양자화된 데이터와 여진원 형성수단에서의 여진원에 의거하여 음성을 합성하는 수단을 포함하여 구성되어 있는 것을 특징으로 하는 음성합성장치.

청구항 19.

복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위마다, 광대역 음성에서 추출된 특성파라미터로부터 미리 형성된 광대역 음성코드북이 사용된 음성합성방법에 있어서,

복수 종류의 부호화 파라미터를 복호화하는 단계와,

복수 종류의 복호화된 파라미터중 첫 번째 파라미터로부터 여진원을 형성하는 단계와,

두 번째 복호화된 파라미터를 음성합성 특성파라미터로 변환하는 단계와,

광대역 음성코드북에서 각 코드 벡터로부터 부분추출에 의해 협대역 특성파라미터를 연산하는 단계와,

연산수단에 의해 연산된 협대역 특성 파라미터와 비교함으로써 음성합성 특성파라미터를 양자화하는 단계와,

광대역 음성코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와,

역양자화된 데이터와 여진원에 기초하여 음성을 합성하는 단계를 포함하여 구성되어 있는 것을 특징으로 하는 음성합성방법.

청구항 20.

제 19항에 있어서,

복수 종류의 부호화 파라미터는 협대역 음성을 부호화함으로써 얻어지고, 부호화 파라미터의 첫 번째 파라미터는 여진원에 관련된 파라미터이고, 두 번째 파라미터는 선형예측계수이고, 세 번째 파라미터는 유/무성음 판별 표지인 것을 특징으로 하는 음성합성방법.

청구항 21.

제 19항에 있어서,

자기상관은 특성 파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 22.

제 19항에 있어서,

켑스트럼은 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 23.

제 19항에 있어서,

스펙트럼 엔벌로프가 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 24.

제 19항에 있어서,

첫 번째 부호화 파라미터의 피치 성분이 강한 것으로 판단될 때, 임펄스열은 여진원으로써 취해지는 것을 특징으로 하는 음성합성방법.

청구항 25.

제 19항에 있어서,

광대역 음성코드북은 소정의 시간단위 마다. 분리된 광대역 유/무성음에서 추출된 유/무성음 특성파라미터로부터 미리 형성된 광대역 유/무성음 코드북이고, 복수의 입력 부호화 파라미터의 세 번째 파라미터를 참조하여 판별가능한 유/무성음 판별의 결과에 의거하여, 음성합성 특성 파라미터는 광대역 유/무성음 코드북에서 각 코드북으로부터 연산에 의해 결정된 협대역 특성 파라미터와 비교함으로써 양자화 되고, 양자화된 데이터는 광대역 유/무성음 코드북을 사용하여 역양자화되고, 음성은 역양자화된 데이터와 여진원에 의거하여 합성되는 것을 특징으로 하는 음성합성방법.

청구항 26.

제 25항에 있어서,

자기상관은 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 27.

제 25항에 있어서,

켑스트럼은 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 28.

제 25항에 있어서,

스펙트럼 엔벌로프가 특성파라미터로써 사용되는 것을 특징으로 하는 음성합성방법.

청구항 29.

제 25항에 있어서,

광대역 유/무성음 코드북을 형성하기 위해 실행된 유/무성음 판별은 세 번째 부호화된 파라미터를 사용하는 것과 다른 것을 특징으로 하는 음성합성방법.

청구항 30.

제 25항에 있어서,

유/무성음을 판별이 불확실한 파라미터를 제외하고, 광대역 유/무성음 코드북과 협대역 유/무성음 코드북을 형성하기 위하여 입력음성으로부터 파라미터를 추출하는 단계를 포함하여 구성되어 있는 것을 특징으로 하는 음성합성방법.

청구항 31.

제 25항에 있어서,

첫 번째 부호화 파라미터의 피치 성분이 강한 것으로 판단될 때, 임펄스열은 여진원으로써 취해지는 것을 특징으로 하는 음성합성방법.

청구항 32.

복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위마다, 광대역 유/무성음에서 추출된 특성파라미터로부터 미리 형성된 광대역 유/무성음 코드북을 사용하는 음성합성장치에 있어서,

복수 종류의 부호화 파라미터를 복호화하는 수단과,

상기 복호화 수단에 의해 복호화된 복수 종류의 파라미터중 첫 번째 파라미터로부터 여진원을 형성하는 수단과,

상기 복호화 수단에 의해 복호화된 복수 종류의 파라미터중 두 번째 복호화 파라미터를 음성합성 특성파라미터로 변환하는 수단과,

상기 광대역 유/무성음 코드북에서 각 코드벡터로부터 부분추출에 의해 협대역 특성파라미터를 연산하는 수단과,

상기 연산 수단으로부터 협대역 특성파라미터를 사용함으로써 파라미터 변환수단으로부터 음성합성 특성파라미터를 양자화하는 수단과,

상기 광대역 유/무성음 코드북을 사용함으로써 양자화 수단으로부터 양자화된 데이터를 역양자화하는 수단과,

상기 역양자화 수단에서의 역양자화된 데이터와 여진원을 형성하는 수단에서의 여진원에 의거하여 음성을 합성하는 수단을 포함하여 구성되어 있는 것을 특징으로 하는 음성합성장치.

청구항 33.

입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 파라미터로부터 미리 형성된 광대역 유성음 코드북과 광대역 무성음 코드북, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성 파라미터로부터 미리 형성된 협대역 유성음 코드북과 협대역 무성음 코드북이 사용되는 음성대역 확장방법에 있어서,

소정의 시간 단위마다 입력 협대역 음성에서 유성음과 무성음을 판별하는 단계와,

협대역 유/무성음으로부터 유성파라미터와 무성파라미터를 생성하는 단계와,

협대역 유/무성음 코드북을 사용함으로써 협대역 음성의 협대역 유/무성음 파라미터를 양자화하는 단계와,

협대역 유/무성음 코드북을 사용하여 양자화된 협대역 유/무성음 데이터를 협대역 유/무성음 코드북을 사용함으로써 역양자화하는 단계와,

역양자화된 데이터에 의거하여 협대역 음성의 대역을 확장하는 단계를 포함하여 구성되어 있는 것을 특징으로 하는 음성대역 확장방법.

청구항 34.

입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 파라미터로부터 미리 형성된 광대역 유성음 코드북과 광대역 무성음 코드북과, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성파라미터로부터 미리 형성된 협대역 유성음 코드북과 협대역 무성음 코드북을 사용하는 음성대역 확장장치에 있어서,

소정의 시간 단위마다 입력 협대역 음성에서 유성음과 무성음을 판별하는 수단과,

유/무성음 판별수단에 의해 판별된 협대역 유/무성음으로부터 유성음 파라미터와 무성음 파라미터를 생성하는 수단과,

협대역 유/무성음 코드북을 사용함으로써 협대역 유/무성음 파라미터를 생성하는 수단으로부터 협대역 유/무성음 파라미터를 양자화하는 수단과,

광대역 유/무성음 코드북을 사용함으로써, 협대역 유/무성음 코드북의 사용에 의한 협대역 유/무성음 양자화 수단으로부터 협대역 유/무성음 데이터를 역양자화하는 수단과,

광대역 유/무성음을 역양자화하는 수단으로부터 역양자화된 데이터에 의거하여 확장된 협대역 음성의 대역을 포함하여 구성되어 있는 것을 특징으로 하는 음성대역 확장장치.

청구항 35.

입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 음성에서 추출된 파라미터로부터 미리 형성된 광대역 유/무성음 코드북이 사용된 음성대역 확장방법에 있어서,

입력 협대역 음성에서 협대역 파라미터를 생성하는 단계와,

상기 광대역 유/무성음 코드북에서 각 코드벡터로부터 협대역 파라미터를 연산하는 단계와,

상기 연산된 협대역 파라미터와 비교함으로써 입력 협대역 음성으로부터 생성된 협대역 파라미터를 양자화하는 단계와,

상기 광대역 유/무성음 코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와,

상기 역양자화된 데이터에 의거하여 협대역 음성의 대역을 확장하는 단계를 포함하여 구성되어 있는 것을 특징으로 하는 음성대역 확장방법.

청구항 36.

입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 유/무성음에서 추출된 파라미터로부터 미리 형성된 광대역 유/무성음 코드북을 사용하는 음성대역 확장장치에 있어서,

상기 입력 협대역 음성으로부터 협대역 파라미터를 생성하는 수단과,

상기 광대역 유/무성음 코드북에서 각 코드 벡터로부터 협대역 파라미터를 연산하는 수단과,

상기 협대역 파라미터 연산수단에서 협대역 파라미터와 비교함으로써 입력 협대역 파라미터 생성수단으로부터 협대역 파라미터를 양자화하는 수단과,

상기 광대역 유/무성음 코드북을 사용함으로써 협대역 음성 양자화수단으로부터 양자화된 협대역 데이터를 역양자화하는 수단과,

상기 광대역 유/무성음을 역양자화 수단으로부터 역양자화된 데이터에 의거하여 확장된 협대역 음성의 대역을 포함하여 구성되어 있는 것을 특징으로 하는 음성대역 확장장치.

청구항 37.

입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 유/무성음에서 추출된 파라미터로부터 미리 형성된 광대역 유/무성음 코드북이 사용된 음성대역 확장방법에 있어서,

입력 협대역 음성으로부터 협대역 파라미터를 생성하는 단계와,

상기 광대역 유/무성음 코드북에서 각 코드벡터로부터 부분추출에 의해 협대역 파라미터를 연산하는 단계와,

상기 연산된 협대역 파라미터와 비교함으로써 입력 협대역 음성으로부터 생성된 협대역 파라미터를 양자화하는 단계와,

상기 광대역 유/무성음 코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와,

상기 역양자화된 데이터에 의거하여 협대역 음성의 대역을 확장하는 단계를 포함하여 구성되어 있는 것을 특징으로 하는 음성대역 확장방법.

청구항 38.

입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 유/무성음에서 추출된 파라미터로부터 미리 형성된 광대역 유/무성음 코드북을 사용하는 음성대역 확장장치에 있어서,

입력 협대역 음성으로부터 협대역 파라미터를 생성하는 수단과,

상기 광대역 유/무성음 코드북에서 각 코드 벡터로부터 부분추출에 의해 협대역 파라미터를 연산하는 수단과,

상기 협대역 파라미터를 연산수단에서 협대역 파라미터를 사용함으로써 음성을 사용하는 협대역 파라미터를 생성수단으로부터 생성된 협대역 파라미터를 양자화하는 수단과,

상기 광대역 유/무성음 코드북을 사용함으로써 양자화수단으로부터 양자화된 협대역 데이터를 역양자화하는 수단과,

상기 역양자화 수단으로부터 역양자화된 데이터에 의거하여 확장된 협대역 음성의 대역을 포함하여 구성되어 있는 것을 특징으로 하는 음성대역 확장장치.

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 송신장치로부터 송신된 부호화된 파라미터로부터 음성을 합성하는 방법 및 장치에 관한 것이고, 또한 송신로를 통해 변화없이 대역폭을 유지하면서, 전화회선이나 방송망과 같은 통신망을 통한 송신장치로부터 수신장치로 송신된 좁은 주파수 대역 음성신호의 대역폭을 확장하는 방법 및 장치에 관한 것이다.

전화선은 예를 들어 300~3400Hz의 좁은 주파수 대역을 사용하도록 규정되어 있고, 따라서 전화망을 통해 전달되는 음성신호의 주파수 대역은 제한되어 있다. 그러므로 종래 아날로그 전화선은 좋은 음질을 보장하지 못했다. 이것은 디지털 휴대전화기의 음질에서도 마찬가지이다.

그러나, 전화전송선로에 대한 표준과 규정과 규약이 이미 엄격하게 정의되어 있기 때문에, 그런 특정한 통신에 대한 주파수 대역폭을 확장하는 것은 어렵다. 이런 상황하에서, 수신장치에서 대역외의 신호성분을 예측함으로써 광대역 신호를 생성하는 다양한 접근방법이 제시되어 왔다. 그런 기술적 제안들 중에서, 음성코드 북 매핑(mapping)을 사용함으로써 그런 결점을 극복하는 접근방식은 좋은 음질을 위해 최선이라고 여겨졌다. 이런 접근방식은 수신장치에 제공된 협대역 음성으로부터 광대역음성의 스팩트럼 엔벌로프(envelope)를 예측하기 위하여 음성 분석용과 합성용 2개의 음성코드북이 사용된 것을 특징으로 한다.

구체적으로, 상기 접근방식은 스펙트럼 엔벌로프를 나타내는 파라미터 즉 협대역용과 광대역용으로 2개의 음성코드북을 형성하는 LPC(Linear Predictive Code) 켑스트럼(cepstrum)을 이용한다. 이런 2개의 음성코드북에 있는 코드벡터들 사이에는 일대일 대응이 존재한다. 협대역 LPC 켑스트럼은 입력 협대역 음성으로부터 정해지고, 협대역 음성코드북에 있는 코드벡터와 비교함으로써 벡터로 양자화되고, 광대역 음성코드북내의 대응하는 코드벡터를 이용하여 역양자화되고, 광대역 LPC 켑스트럼을 구함으로써 구성된다.

코드벡터들 사이에 일대일 대응에 대한 2개의 음성코드북은 이하 서술되는 것과 같이 생성된다. 첫째로, 광대역 학습용 음성이 마련되고, 그것은 협대역 학습용 음성을 제공하기 위해 대역폭이 제한된다. 준비된 광대역 협대역 학습용 음성은 각각 프레임화되고, 협대역 음성으로부터 구해진 LPC 켑스트럼은 우선 학습하도록 사용되고, 협대역 음성코드북을 생성한다. 그 결과로 코드벡터로 양자화된 학습용 협대역 음성프레임에 대응하는 학습용 광대역 음성의 프레임이 집단화되고, 광대역 음성코드북을 형성하는 광대역 코드벡터를 제공하여 적재된다.

이런 접근방식의 다른 응용으로써, 광대역 음성코드북은 먼저 학습용 광대역 음성으로부터 생성되고, 대응하는 학습용 협대역 음성코드벡터들은 협대역 음성코드북을 생성하는 협대역 코드벡터을 제공하여 적재된다.

더욱이, 코드벡터화된 파라미터로써 자기상관이 사용되는 음성코드북생성 모드방식이 제안되어 왔다. 또한, 여진원(innovation)이 LPC분석과 합성에 필요요건이다. 그러한 여진원은 임펄스열과 노이즈와 업샘플된 협대역 여진원 등을 포함한다.

상기 언급된 접근방식의 응용은 만족할만한 음질에 도달하지 못한다. 특히, 현재 일본에서 널리 이용되고 있는 디지털 휴대전화시스템에서 채택되어 사용하고 있는, 소위 음성 부호화모드 CELP(Code Excited Linear Prediction : 부호화여기 선형예측)모드에 포함된 VSELP(Vector Sum Excited Linear Prediction:벡터합여기 선형예측)모드, PSI-CELP(Pitch Synchronous Innovation-Code Excited Linear Prediction:피치동기여진원-코드여기선형예측) 모드 등과 같은 저 비트레이트 음성부호화 모드에서 부호화된 음성에 이 접근방식이 적용될 때, 음질은 매우 불충분하다. 또한, 협대역과 광대역 음성코드북을 생성하는데 사용된 메모리의 크기는 불충분하다.

발명이 이루고자 하는 기술적 과제

따라서, 본 발명은 청감상 양질의 음을 갖는 광대역 음성을 제공할수 있는 음성합성방법 및 장치와 대역확장장치 및 방법을 제공함으로써 종래기술의 상기 언급된 문제점을 극복하는 것을 목적으로 하고 있다.

종래 기술의 문제점을 극복하기 위하여, 본 발명은 음성분석과 합성 모두 음성코드북을 사용함으로써 기억용량을 절약할 수 있는 음성합성방법 및 장치와 대역확장방법 및 장치를 제공하는 것을 또다른 목적으로 하고 있다.

발명의 구성

상기 목적은, 복수 종류의 입력 부호화된 파라미터의 음성을 합성하기 위하여 소정된 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 특성 파라미터로부터 형성된 광대역 유성음 코드북과 광대역 무성음 코드북, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성 파라미터로부터 형성된 협대역 유성음 코드북과 협대역 무성음 코드북을 사용하고, 본 발명에 따르면 복수의 부호화된 파라미터를 복호화하는 단계와, 복수의 복호화된 첫 번째 파라미터로부터 여진원(innovation)을 형성하는 단계와, 두 번째 복호화된 파라미터를 음성합성 특성파라미터로 변환하는 단계와, 세 번째 복호화된 파라미터를 참조하여 유/무성음을 판별하는 단계와, 협대역 유/무성음

코드북을 사용함으로써 판별의 결과에 기초한 음성합성 특성 파라미터를 양자화하는 단계와, 광대역 유/무성음 코드북을 사용함으로써, 협대역 유/무성음 코드북을 사용하여 양자화된 협대역 유/무성음 데이터를 역양자화하는 단계와, 역양자화된 데이터와 여진원을 기초로 한 음성을 합성하는 단계를 포함하는 음성합성방법을 제공함으로써 이루어질수 있다.

상기 목적은 복수 종류의 입력 부호화된 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 특성파라미터로부터 미리 형성된 광대역 유성음 코드북과 광대역 무성음 코드북과, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성파라미터로부터 미리 형성된 협대역 유성음 코드북과 협대역 무성음 코드북을 사용하고, 본 발명에 따르면 복수의 부호화 파라미터를 복호화하는 수단과, 복호화 수단에 의해 복호화된 복수의 파라미터들 중 첫 번째 파라미터로부터 여진원을 형성하는 수단과, 복호화 수단에 의해 복호화된 부호화 파라미터중 제 2파라미터로부터 음성합성 특성파라미터를 얻는 수단과, 복호화수단에 의해 복호화된 부호화 파라미터의 제 3파라미터를 참조하여 유/무성음을 판별하는 수단과, 협대역 유/무성음 코드북을 사용함으로써, 유/무성음 판별의 결과에 의거한 음성합성 특성파라미터를 양자화하는 수단과, 광대역 유/무성음 코드북을 사용함으로써 유/무성음 양자화 수단으로부터 양자화된 유/무성음 데이터를 역양자화하는 수단과, 광대역 유/무성음 역양자화 수단에서의 역양자화된 데이터와 여진원 형성 수단으로부터의 여진원에 의거하는 음성을 합성하는 수단을 포함하는 음성합성장치를 제공함으로써 이루어질수 있다.

상기 목적은 복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위해, 소정의 시간단위마다 광대역 음성에서 추출된 특성 파라미터로부터 미리 형성된 광대역 음성 코드북이 사용되고, 본 발명에 따르면, 복수의 부호화 파라미터를 복호화하는 단계와, 복수의 복호화 파라미터들 중 첫 번째 파라미터로부터 여진원을 형성하는 단계와, 두 번째 복호화된 파라미터를 음성합성 특성파라미터로 변환하는 단계와, 광대역 음성코드북에서 각 코드 벡터로부터 협대역 특성 파라미터를 연산하는 단계와, 연산 수단에 의해 연산된 협대역 특성파라미터와 비교함으로써 음성합성 특성파라미터를 양자화하는 단계와, 광대역 음성 코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와, 역양자화된 데이터와 여진원에 의거한 음성을 합성하는 단계를 포함하는 음성합성방법을 제공함으로써 이루어질수 있다.

상기 목적은 복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위 마다 광대역 음성에서 추출된 특성 파라미터로부터 미리 형성된 광대역 음성코드북을 사용하고, 본 발명에 따르면, 복수 종류의 부호화 파라미터를 복호화하는 수단과, 복호화 수단에 의해 복호화된 복수 종류의 파라미터들 중 첫 번째 파라미터로부터 여진원을 형성하는 수단과, 복호화 수단에 의해 복호화된 복수 종류의 파라미터의 두 번째 복호화된 파라미터를 음성 합성 특성파라미터로 변환하는 수단과, 광대역 음성코드북에서 각 코드북으로부터 협대역 특성 파라미터를 연산하는 수단과, 연산수단으로부터 협대역 특성 파라미터를 사용함으로써 파라미터 변환수단으로부터 음성합성 특성파라미터를 양자화하는 수단과, 광대역 음성코드북을 사용함으로써 양자화하는 수단으로부터 양자화된 데이터를 역양자화하는 수단과, 역양자화 수단에서의 역양자화된 데이터와 여진원 형성수단에서의 여진원에 의거하여 음성을 합성하는 수단을 포함하는 음성합성장치를 제공함으로써 이루어질수 있다.

상기 목적은 복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위마다, 광대역 음성에서 추출된 특성파라미터로부터 미리 형성된 광대역 음성코드북이 사용되고, 본 발명에 따르면, 복수 종류의 부호화 파라미터를 복호화하는 단계와, 복수 종류의 복호화된 파라미터중 첫 번째 파라미터로부터 여진원을 형성하는 단계와, 두 번째 복호화된 파라미터를 음성합성 특성파라미터로 변환하는 단계와, 광대역 음성코드북에서 각 코드 벡터로부터 부분추출에 의해 협대역 특성파라미터를 연산하는 단계와, 연산수단에 의해 연산된 협대역 특성 파라미터와 비교함으로써 음성합성 특성파라미터를 양자화하는 단계와, 광대역 음성코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와, 역양자화된 데이터와 여진원에 기초하여 음성을 합성하는 단계를 포함하는 음성합성방법을 제공함으로써 이루어질 수 있다.

상기 목적은 복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위마다, 광대역 음성에서 추출된 특성파라미터로부터 미리 형성된 광대역 음성코드북이 사용되고, 본 발명에 따르면, 복수 종류의 부호화 파라미터를 복호화하는 단계와, 복수 종류의 복호화 파라미터의 첫 번째 파라미터로부터 여진원을 형성하는 단계와, 두 번째 복호화된 파라미터를 음성합성 특성 파라미터로 변환하는 단계와, 광대역 음성코드북에서 각 코드 벡터로부터 부분추출에 의해 협대역 특성 파라미터를 연산하는 단계와, 연산수단에 의해 추출된 협대역 특성파라미터와 비교함으로써 음성합성 특성파라미터를 양자화하는 단계와, 광대역 음성코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와, 역양자화된 데이터와 여진원에 의거하여 음성을 합성하는 단계를 포함하는 음성합성방법을 제공함으로써 이루어질수 있다.

상기 목적은 복수 종류의 입력 부호화 파라미터로부터 음성을 합성하기 위하여, 소정의 시간단위마다, 광대역 음성에서 추출된 특성파라미터로부터 미리 형성된 광대역 음성코드북을 사용하고, 본 발명에 따르면, 복수 종류의 부호화 파라미터를 복호화하는 수단과, 복호화 수단에 의해 복호화된 복수 종류의 파라미터중 첫 번째 파라미터로부터 여진원을 형성하는 수단과, 복호화 수단에 의해 복호화된 복수 종류의 파라미터중 두 번째 복호화 파라미터를 음성합성 특성파라미터로 변환하

는 수단과, 광대역 음성코드북에서 각 코드벡터로부터 부분추출에 의해 협대역 특성파라미터를 연산하는 수단과, 연산 수단으로부터 협대역 특성파라미터를 사용함으로써 파라미터 변환수단으로부터 음성합성 특성파라미터를 양자화하는 수단과, 광대역 음성코드북을 사용함으로써 양자화 수단으로부터 양자화된 데이터를 역양자화하는 수단과, 역양자화 수단에서의 역양자화된 데이터와 여진원을 형성하는 수단에서의 여진원에 의거하여 음성을 합성하는 수단을 포함하는 음성합성장치를 제공함으로써 이루어질 수 있다.

상기 목적은 입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 파라미터로부터 미리 형성된 광대역 유성음 코드북과 광대역 무성음 코드북, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성 파라미터로부터 미리 형성된 협대역 유성음 코드북과 협대역 무성음 코드북이 사용되고, 본 발명에 따르면, 소정의 시간 단위마다 입력 협대역 음성에서 유성음과 무성음을 판별하는 단계와, 협대역 유/무성음으로부터 유성파라미터와 무성파라미터를 생성하는 단계와, 협대역 유/무성음 코드북을 사용함으로써 협대역 음성의 협대역 유/무성음 파라미터를 양자화하는 단계와, 협대역 유/무성음 코드북을 사용하여 양자화된 협대역 유/무성음 데이터를 협대역 유/무성음 코드북을 사용함으로써 역양자화하는 단계와, 역양자화된 데이터에 의거하여 협대역 음성의 대역을 확장하는 단계를 포함하는 음성대역 확장방법을 제공함으로써 이루어질수 있다.

상기 목적은 입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 분리된 광대역 유/무성음에서 각각 추출된 유/무성음 파라미터로부터 미리 형성된 광대역 유성음 코드북과 광대역 무성음 코드북과, 분리된 광대역 유/무성음의 주파수 대역을 제한함으로써 얻어진 협대역 음성에서 추출된 유/무성음 특성파라미터로부터 미리 형성된 협대역 유성음 코드북과 협대역 무성음 코드북을 사용하고, 본 발명에 따르면, 소정의 시간 단위마다 입력 협대역 음성에서 유성음과 무성음을 판별하는 수단과, 유/무성음 판별수단에 의해 판별된 협대역 유/무성음으로부터 유성음 파라미터와 무성음 파라미터를 생성하는 수단과, 협대역 유/무성음 코드북을 사용함으로써 협대역 유/무성음 파라미터를 생성하는 수단으로부터 협대역 유/무성음 파라미터를 양자화하는 수단과, 광대역 유/무성음 코드북을 사용함으로써, 협대역 유/무성음 코드북의 사용에 의한 협대역 유/무성음 양자화 수단으로부터 협대역 유/무성음 데이터를 역양자화하는 수단과, 광대역 유/무성음을 역양자화하는 수단으로부터 역양자화된 데이터에 의거하여 확장된 협대역 음성의 대역을 포함하는 음성대역 확장장치를 제공함으로써 이루어질수 있다.

상기 목적은 입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 음성에서 추출된 파라미터로부터 미리 형성된 광대역 음성코드북이 사용되고, 본 발명에 따르면, 입력 협대역 음성에서 협대역 파라미터를 생성하는 단계와, 광대역 음성코드북에서 각 코드벡터로부터 협대역 파라미터를 연산하는 단계와, 연산된 협대역 파라미터와 비교함으로써 입력 협대역 음성으로부터 생성된 협대역 파라미터를 양자화하는 단계와, 광대역 음성코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와, 역양자화된 데이터에 의거하여 협대역 음성의 대역을 확장하는 단계를 포함하는 음성대역 확장장치를 제공함으로써 이루어질수 있다.

상기 목적은 입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 음성에서 추출된 파라미터로부터 미리 형성된 광대역 음성코드북을 사용하고, 본 발명에 따르면, 입력 협대역 음성으로부터 협대역 파라미터를 생성하는 수단과, 광대역 음성코드북에서 각 코드 벡터로부터 협대역 파라미터를 연산하는 수단과, 협대역 파라미터 연산수단에서 협대역 파라미터와 비교함으로써 입력 협대역 파라미터 생성수단으로부터 협대역 파라미터를 양자화하는 수단과, 광대역 음성코드북을 사용함으로써 협대역 음성 양자화수단으로부터 양자화된 협대역 데이터를 역양자화하는 수단과, 광대역 음성을 역양자화 수단으로부터 역양자화된 데이터에 의거하여 확장된 협대역 음성의 대역을 포함하는 음성대역 확장장치를 제공함으로써 이루어질수 있다.

상기 목적은 입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 음성에서 추출된 파라미터로부터 미리 형성된 광대역 음성코드북이 사용되고, 본 발명에 따르면, 입력 협대역 음성으로부터 협대역 파라미터를 생성하는 단계와, 광대역 음성코드북에서 각 코드벡터로부터 부분추출에 의해 협대역 파라미터를 연산하는 단계와, 연산된 협대역 파라미터와 비교함으로써 입력 협대역 음성으로부터 생성된 협대역 파라미터를 양자화하는 단계와, 광대역 음성코드북을 사용함으로써 양자화된 데이터를 역양자화하는 단계와, 역양자화된 데이터에 의거하여 협대역 음성의 대역을 확장하는 단계를 포함하는 음성대역 확장방법을 제공함으로써 이루어질수 있다.

상기 목적은 입력 협대역 음성의 대역을 확장하기 위하여, 소정의 시간단위마다 광대역 음성에서 추출된 파라미터로부터 미리 형성된 광대역 음성코드북을 사용하고, 본 발명에 따르면, 입력 협대역 음성으로부터 협대역 파라미터를 생성하는 수단과, 광대역 음성코드북에서 각 코드 벡터로부터 부분추출에 의해 협대역 파라미터를 연산하는 수단과, 협대역 파라미터를 연산수단에서 협대역 파라미터를 사용함으로써 음성을 사용하는 협대역 파라미터를 생성수단으로부터 생성된 협대역

파라미터를 양자화하는 수단과, 광대역 음성코드북을 사용함으로써 양자화수단으로부터 양자화된 협대역 데이터를 역양자화하는 수단과, 역양자화 수단으로부터 역양자화된 데이터에 의거하여 확장된 협대역 음성의 대역을 포함하는 음성대역 확장장치를 제공함으로써 이루어질수 있다.

도 1을 참조하여, 본 발명의 음성대역 확장장치의 실시의 예와 협대역 음성의 대역을 설명한다. 여기에서, 음성대역 확장장치는 300~3400Hz의 주파수 대역과 8kHz의 샘플링 주파수를 가지는 협대역 음성신호로부터 입력에 제공되는 것으로 가정한다.

본 발명에 따르면 음성대역 확장장치는 광대역 유/무성음에서 추출된 유/무성음 파라미터를 이용하여 형성된 광대역 유성음 코드북(12)과 광대역 무성음 코드북(14)이 있고, 예를 들면 광대역 음성의 주파수 대역을 제한함으로써 생성된 300~3400Hz의 주파수대역을 가지는 협대역 음성신호에서 추출된 유/무성음 파라미터로부터 형성된 협대역 유성음 코드북(8)과 협대역 무성음 코드북(10)이 있다.

본 발명에 따르면, 음성대역 확장장치는 매 160샘플(1 프레임은 샘플링 주파수가 8kHz이기 때문에 20msec와 같다)로 입력단자(1)에 수신된 협대역 음성신호를 프레임화하기 위해 제공된 프레이밍 회로(2)와 프레임화된 협대역 음성신호를 기초로 한 여진원(innovation)을 형성하기 위한 제로필링(zerofilling)회로(16)와 매 20msec 프레임으로 협대역 음성신호상에서 유성음(V)과 무성음(UV)을 판별하기 위한 V/UV 판별기(5)와 V/UV 판별의 결과를 기초로 한 협대역 유/무성음에 대한 선형예측계수을 생성하기 위한 선형예측코드(LPC: linear prediction code) 분석기(31) 즉, LPC 분석장치(3)로부터 일종의 파라미터인 자기상관(γ)으로 선형예측계수(α)를 변환하기 위한 α/γ변환장치(4), 협대역 유성음 코드북(8)을 이용하는 α/γ변환장치(4)로부터 협대역 유성음 자기상관(γ)을 양자화하기 위한 협대역 유성음 양자화장치(7), 협대역 무성음 코드북(10)을 이용하는 α/γ변환장치(4)로부터 협대역 무성음 자기상관(γ)을 양자화하기 위한 협대역 무성음 양자화장치(9), 광대역 유성음 코드북(12)을 이용하는 협대역 유성음 양자화장치(7)로부터 협대역 유성음이 양자화된 데이터를 역양자화하기 위한 광대역 유성음 역양자화 장치(11), 광대역 무성음 코드북(14)을 이용하는 협대역 무성음 양자화 장치(9)로부터 협대역 무성음 양자화된 데이터를 역양자화하기 위한 광대역 무성음 역양자화 장치(13), 광대역 유성음 역양자화 장치(11)로부터 협대역 유성음 선형예측계수로 그리고 광대역 무성음 역양자화 장치(13)로부터 협대역 무성음 선형예측계수로 광대역 무성음 자기상관(역양자화된 데이터)을 변환하기 위한 γ/α 변환장치와 γ/α변환기(15)로부터의 협대역 유/무성음 선형예측계수와 제로필링회로(16)으로부터의 여진원을 기초로 한 광대역 음성을 합성하기 위한 LPC 합성장치(17)를 포함하여 구성한다.

음성대역 확장장치는 또한 8~16kHz의 프레이밍회로(2)로부터 프레임화된 협대역 음성의 샘플링 주파수를 바꾸기 위해 제공된 오버샘플링회로(19)와 LPC 합성장치(17)에서 합성된 출력으로부터 입력 협대역 유성음 신호의 주파수 대역에서 300~3400Hz의 신호를 삭제하거나 제거하기 위한 BSF(band stop filter:대역스톱필터)와 오버샘플링회로(19)로부터 본래 협대역 유성음 신호의 샘플링 주파수 16kHz의 주파수 대역 300~3400Hz 신호를 BSF(대역스톱필터)(18)로부터 출력에 더하기 위한 가산장치(20)를 포함하여 구성한다. 음성대역 확장장치는 300~7000Hz의 주파수 대역과 16kHz의 샘플링 주파수를 갖는 디지탈 음성신호를 출력단자(21)에 전달한다.

이제, 광대역 유/무성음 코드북(12, 14)과 협대역 유/무성음 코드북(8, 10)이 어떻게 구성되는지를 설명한다.

첫째, 예를 들어 프레이밍 회로(2)에서 매 20msec로 프레임화된 300~7000Hz의 주파수 대역을 갖는 광대역 음성신호는 유성음(V)과 무성음(UV)으로 분리된다. 유성음 파라미터와 무성음 파라미터는 각각 유/무성음으로부터 추출되고, 광대역 유/무성음 코드북(12, 14)을 생성한다.

또한, 협대역 유/무성음 코드북(8, 10)의 생성을 위해 광대역 음성은 예를 들면, 유성음 파라미터와 무성음 파라미터가 추출되는 300~3400Hz의 주파수 대역을 가지는 협대역 유성음 신호로 주파수 대역이 제한된다. 유/무성음 파라미터는 협대역 유/무성음 코드북(8, 10)을 생성하기 위해 사용된다.

도 2는 상기 언급된 4종류의 음성코드북의 생성을 위해 학습데이터의 준비를 나타내는 플로우 챠트이다. 도 2에서와 같이, 협대역 학습음성신호가 생성되고, 단계(S1)에서 매 20msec로 프레임화된다. 단계(S2)에서, 광대역 학습음성신호는 협대역 음성신호를 생성하기위해 대역제한된다. 단계(S3)에서, 협대역 음성신호는 단계(S1)에서와 같이 프레이밍 타임(20 msec/frame)에서 프레임화된다. 각 프레임의 협대역 음성신호는 프레임 에너지와 영교차(zero-cross)로 검파되고, 음성신호는 유성음(V) 또는 무성음(UV)으로 단계(S4)에서 판단된다.

고질의 음성코드북을 위해, 유성음(V)에서 무성음(UV)으로 (역도 같다) 변화에 있어 구성요소와 V와 UV를 판별하기 어려운 구성요소는 확실하게 V와 UV가 되는 음성만을 제공하기 위해 제거된다. 그리고, 학습용 협대역 V 프레임과 학습용 협대역 UV 프레임의 집합이 얻어진다.

다음에, 광대역 음성 프레임은 또한 V와 UV음성으로 분류된다. 그러나, 광대역 프레임은 협대역 프레임과 같은 타이밍에서 프레임되기 때문에 협대역 음성신호의 판별에서 V가 되도록 분류된 협대역 프레임으로써 동시에 처리된 광대역 프레임 V와 UV가 되도록 분류된 협대역 프레임으로써 동시에 처리된 광대역 프레임 UV로써 분류된다. 그리고 학습 데이터가 생성된다. 말할 필요도 없이, 협대역 프레임 판별에서 V 와 UV도 아닌 프레임은 분류되지 않는다.

또한, 학습 데이터는 설명되지 않은 반대 방법으로도 생성될 수 있다. 즉, V/UV 분류는 광대역 프레임상에서 이용된다. 그 분류의 결과는 협대역 프레임을 V나 UV로 분류한다.

다음에, 상기와 같이 생성된 학습 데이터는 도 3에서 보여지는 것과 같이 음성 코드북을 생성하기 위해 사용된다. 도 3은 음성 코드북의 생성을 나타내는 플로우 챠트이다. 도 3에서 보여지는 바와 같이, 광대역 V(UV) 프레임의 집합은 우선 광대역 V(UV) 음성 코드북을 생성하고 학습하는데 이용된다.

첫째, dn 차원(dimension)에까지 자기상관 파라미터는 단계(S6)에서와 같이, 각 광대역 프레임으로부터 추출된다. 자기상관 파라미터는 다음의 방정식(1)을 기초로 하여 연산된다.

[수학식 1]

$$\phi(xi) = \left(\sum_{j=0}^{N-1-i} xjxj + 1 \right) / \left(\sum_{j=0}^{N-1} x\,2j \right)$$

x: 입력신호, f(xi): n차 자기상관, 그리고 N: 프레임 길이.

단계(S7)에서, GLA(Generalized Lloyd Algorithm)는 각 광대역 프레임의 dw 차원의 자기상관 파라미터로부터 크기 sw의 dw-차원의 광대역 V(UV) 음성코드북을 생성하기 위해 사용된다.

그것은 각 광대역 V(UV)프레임의 자기상관 파라미터가 생성된 음성코드북의 코드벡터는 양자화된 부호화 결과로부터 검파된다. 각 코드 벡터에 대하여, 광대역 V(UV)프레임과 같이 동시에 처리된 각 협대역 U(UV)프레임으로부터 얻어진 벡터로 양자화된 광대역 V(UV)프레임에 대응하는 dn 차원의 자기상관 파라미터들은 단계(S8)에서 협대역 코드벡터로써 적재된다. 이 동작은 모든 코드벡터가 협대역 음성코드북을 생성하도록 한다.

도 4는 상기와 대칭적인 방법으로 음성코드북의 생성을 나타내는 플로우 챠트이다. 즉, 협대역 프레임 파라미터는 단계(S9, S10)에서 우선 학습용으로 협대역 음성코드북을 생성하기 위해 사용된다. 단계(S11)에서, 대응하는 광대역 프레임 파라미터가 적재된다.

상기 설명과 같이, 4개의 음성코드북은 협대역 V와 UV 음성 코드북과 광대역 V 와 UV 음성 코드북이다.

상기 언급된 음성대역 확장을 하는 음성대역 확장장치는 도 1에서 음성대역 확장장치의 동작을 나타내는 플로우 챠트인 도 5를 참조하여 설명되는 것과 같이, 상기 4개의 음성 코드북을 이용한 입력 협대역 음성을 협대역 음성으로 변환하기 위한 기능을한다.

먼저, 음성대역 확장장치의 입력단자(1)에 수신된 협대역 음성신호는 단계(S21)에서 프레이밍 회로(2)에 의해 매 160 샘플(20msec)로 프레임화된다. 프레이밍 회로(2)로부터 각 프레임은 LPC 분석장치(3)로 제공되고, 단계(S23)에서 LPC 분석을 필요로 한다. 프레임은 선형예측계수 파라미터(α)와 LPC 잔차(remainder)로 분리된다. 파라미터(α) 는 α/ɣ변환기(4)로 제공되고, 단계 S24에서 자기상관(ɣ)으로 변환된다.

또한, 프레임화된 신호는 단계(S22)의 V/UV 판별장치(5)에서 V(유성음)와 UV(무성음)를 판별된다. 도 1에 보여지는 것과 같이, 본 발명에 따른 음성확장장치는 또한 α/ɣ변환장치(4)의 출력을 협대역 V 음성 양자화장치(7)로 연결하기 위해 제

공된 스위치(6) 또는α/ɣ 변환기(4)의 하류부분으로 제공된 협대역 UV 음성 양자화장치(9)를 포함하여 구성한다. 프레임화된 신호가 V로 판단될 때, 스위치(6)는 신호경로를 협대역 유성음 양자화 장치(7)로 연결한다. 반대로, 신호가 UV로 판단될 때, 스위치(6)는 α/ɣ 변환기(4)의 출력을 협대역 UV 음성 양자화 장치(9)로 연결한다.

그러나, 이 단계(S22)에서 실행된 V/UV 판별은 음성코드북 생성을 위해 실행된 것과 다르다는 것에 주목하라. 즉, V와 UV에 속하지 않는 어떤 프레임이 생성된다. V/UV 판별장치(5)에서, 프레임신호는 틀림없이 V나 UV로 판단된다. 그러나, 사실상 고역상에서 음성신호는 큰 에너지를 나타낸다. UV음성은 V음성보다 큰 에너지를 갖고 있다. 큰 에너지를 갖는 음성신호는 UV신호로 판단되는 경향이 있다. 이런 경우, 비정상적인 음성이 생성된다. 이것을 피하기 위하여, V/UV 판별장치는 V 와 UV를 판별하기 어려운 음성신호를 V로써 받아들이도록 설정된다.

V/UV 판별장치(5)가 입력 음성신호를 V 음성으로 판단할 때, 스위치(6)로부터 유성음 자기상관(g)은 단계(S25)의 협대역 V 음성코드북(8)을 사용하여 양자화된 협대역 V 음성 양자화장치(7)로 제공된다. 반대로, V/UV판별장치(5)가 입력음성신호를 UV 음성으로 판단할 때, 스위치(6)에서 무성음 자기상관(ɣ)은 단계(S25)의 협대역 UV 음성코드북(10)을 사용하여 양자화된 협대역 UV 양자화 장치(9)로 제공된다.

단계(S26)에서, 광대역 V 역양자화장치(11) 또는 광대역 UV 역양자화 장치(13)는 광대역 V 음성코드북(12) 또는 광대역 UV 음성코드북(14)을 이용하여 양자화된 자기상관(ɣ)을 역양자화하고, 광대역 자기상관(ɣ)을 제공한다.

단계(S27)에서, 협대역 자기상관(ɣ)은 ɣ/α변환장치(15)에 의해 광대역 자기상관(α)으로 변환된다.

한편, LPC분석기(3)로부터 LPC 잔차(remainder)는 단계 S28에서 제로필링회로(16)에 의한 샘플들중 제로필링에 의해 광대역으로 에일리어스(alias)되고, 업샘플된다.

단계(S29)에서, 광대역 자기상관과 광대역 여진원은 광대역 음성신호를 제공하기 위해 LPC합성장치(17)에서 LPC합성을 필요로 한다.

그러나, 얻어진 광대역 음성신호는 단지 예측(prediction) 결과에서 생긴 것이고, 신호가 만약 처리되지 않는다면, 예측 에러를 포함한다. 특히, 입력 협대역 음성은 주파수 범위에서 처리되지 않고 그대로 남겨지는 것이 바람직하다.

그러므로, 단계(S30)에서 입력 협대역 음성은 대역스톱필터 (BSF: band stop filter)(18) 에 의해 필터링을 통하여 제거된 주파수 범위를 가지고, 단계(S32)의 오버샘플링회로(19)에서 오버샘플된 협대역 음성으로 단계 S31에서 가산된다. 거기에서 확장된 대역을 가지는 광대역 음성신호가 제공된다. 상기에서, 이득은 조절되고, 고역은 다소 억압되어 양질의 청감을 가지는 음성을 제공한다.

도 1에서 음성대역 확장장치는 자기상관을 이용하여 총 4개의 음성코드북을 생성한다. 그러나, 자기상관보다는 다른 어떤 파라미터가 사용될수도 있다. 예를 들면, LPC 켑스트럼(cepstrum)이 이 목적에 효과적으로 이용되고, 스펙트럼 엔벌로프(envelope)이 스펙트럼 엔벌로프예측으로부터 파라미터로써 직접 이용된다.

또한, 도 1에서 음성대역 확장장치는 협대역 V (UV) 음성코드북(8, 9)을 사용한다. 그러나, 그것은 음성코드북용 RAM 용량을 줄일 목적으로 생략될 수 있다.

도 6은 도 1에서 줄어든 수의 음성코드북이 사용된 음성대역 확장장치를 나타내는 블록 다이어그램이다. 도 6의 음성대역 확장장치는 협대역 V 와 UV 음성코드북(8, 10)을 사용하는 대신에 산술 회로(25, 26)를 사용한다. 산술회로(25, 26)는 광대역 음성코드북의 코드벡터로부터 산술에 의해 협대역 V와 UV 파라미터를 얻는다.

이 음성대역 확장장치의 나머지 부분은 도 1과 같이 배열된다.

자기상관이 음성 코드북상에서 파라미터로 사용될 때, 광대역과 협대역 음성 자기상관사이의 관계식은 아래와 같이 나타내어진다.

[수학식 2]

$$\phi(x_n) = \phi(x_w \otimes h) = \phi(x_w) \otimes \phi(h)$$

f: 자기상관 x_n: 협대역 음성신호 x_w: 광대역 음성신호 h: BSF의 임펄스 응답

협대역 자기상관 f(x_n)은 상기 관계식을 기초로 광대역 자기상관(x_w)으로부터 산술되며, 이론적으로 광대역과 협대역 벡터는 필요하지 않다.

즉, 협대역 자기상관은 광대역 자기상관과 BSF의 임펄스 응답의 자기상관과의 컨벌루션(convolution)에 의해 결정된다.

그러므로, 도 6에서 음성대역 확장장치는 도 5가 아니라 도 7에서와 같이 변형된 음성대역 확장장치의 동작을 나타내는 대역 확산을 실행한다. 특히, 입력단자(1)에 수신된 협대역 음성신호는 단계(S41)의 프레이밍회로(2)에서 매 160샘플(20msec)로 샘플링되고, 각 프레임이 단계(S43)의 LPC 분석을 필요로 하는 LPC 분석장치(3)에 제공되고, 선형예측계수 파라미터(α)와 LPC 잔차로 분리된다. 파라미터(α)는 α/ɣ 변환기(4)로 제공되어 단계(S44)에서 자기상관(ɣ)으로 변환된다.

또한, 프레임화된 신호는 단계(S42)의 V/UV 판별장치(5)상에서 V(유성음)와 UV(무성음)으로 판별된다. 프레임화된 신호가 V로 판단될 때, 스위치(6)은 α/ɣ변환기(4)로부터 협대역 유성음 양자화 장치(7)로 신호경로를 연결한다. 한편, 신호가 UV로 판단될 때, 스위치(6)은 α/ɣ변환기(4)의 출력을 협대역 UV 음성 양자화 장치(9)로 연결한다.

단계(S42)에서 실행된 V/UV 판별은 음성코드북 생성을 위해 실행된 것과는 다르다. 즉, V나 UV에 속하지 않는 어떤 프레임이 생성된다. V/UV 판별장치(5)에서, 프레임 신호는 틀림없이 V와 UV로 판별된다.

V/UV 판별장치(5)가 입력신호를 V 음성으로 판단할 때, 스위치(6)로부터 유성음 자기상관(ɣ)은 단계(S46)의 양자화된 협대역 V 음성 양자화 장치(7)로 제공된다. 그러나, 이와 같은 양자화 과정에서는, 협대역 음성코드북이 사용되는 것이 아니라, 이전에 설명된 것과 같이 단계(S45)의 연산회로(25)에 의해 정해진 협대역 V 파라미터가 사용된다.

반대로, V/UV 판별장치(5)가 입력 음성신호를 UV음성으로 판단할 때, 스위치(6)로부터 무성음 자기상관(ɣ)은 단계(S46)의 양자화된 협대역 UV 양자화 장치(9)로 제공된다. 그러나, 이때에도 또한 협대역 UV 음성코드북이 사용되는 것이 아니라 연산회로(26)에 의해 정해진 협대역 UV 파라미터가 사용된다.

단계(S47)에서, 광대역 V 역양자화 장치(11) 또는 광대역 UV 역양자화 장치(13)는 광대역 V 음성코드북(12) 또는 광대역 UV 음성코드북(14)을 사용하여 양자화된 자기상관(ɣ)을 역양자화하고, 그리고 광대역 자기상관(ɣ)을 제공한다.

단계(S48)에서, 협대역 자기상관(ɣ)은 ɣ/α변환장치(15)에 의해 광대역 자기상관(α)으로 변환된다.

한편, LPC 분석장치(3)로부터 LPC 잔차는 제로필링회로(16)의 샘플들 사이에서 제로필되고, 광대역을 갖도록 에일리어스되고, 업샘플(단계 (S49))된다. 그것은 LPC 합성장치(17)에 광대역 여진원으로써 제공된다.

단계(S50)에서, 광대역 자기상관과 협대역 여진원은 광대역 음성신호를 제공하기 위해 LPC 합성장치(17)에서 LPC 합성을 필요로 한다.

그러나, 얻어진 광대역 음성신호는 단지 예측으로부터 생긴 것이고, 그것은 처리되지 않는다면, 예측에러를 포함한다. 특히, 입력 협대역 음성은 되도력 이면 그것의 주파수 대역에서 만나는 것 없이 남겨져야한다.

그러므로, 단계(S51)에서, 입력 협대역 음성은 BSF(band stop filter:대역스톱필터)(18)에 의한 필터링을 통하여 제거된 주파수 범위를 가지고, 단계 (S53)에서, 단계(S52)의 오버샘플링 회로(19)에서 오버샘플된 협대역 음성에 가산된다.

도 6의 음성대역 확산장치에서, 양자화 과정은 협대역 음성코드북에서 코드북 벡터가 아니라 광대역 음성코드북으로부터 연산에 의해 결정된 코드벡터와 비교함으로써 실행된다. 그러므로, 광대역 음성코드북은 음성신호 분석과 합성을 위해 사용되고, 협대역 음성 코드북의 저장을 위한 메모리는 도 6의 음성대역 확장장치에는 필요가 없다.

그러나, 도 6의 음성대역 확장장치에서, 메모리 저장으로부터 생겨난 결과보다 음성대역 확장에 대한 동작으로써 더해진 연산이 더 문제가 된다. 이 문제를 피하기 위하여, 본 발명은 도 6에서 어떤 연산된 동작이 없는 음성대역 확장방법이 적용된 다양한 음성대역 확장장치를 제공한다. 도 8은 이러한 음성대역 확장장치의 변형을 나타낸다. 도 8에 나타내는 바와 같

이, 음성대역 확장장치는 도 6에서 나타내는 음성대역 확장장치에서 사용된 연산회로(25, 26)대신에 대역 음성코드북상에서 각 코드벡터에서 부분추출하기 위하여 부분추출회로(28, 29)를 사용한다. 이 음성대역 확장장치의 나머지 부분은 도 1이나 도 6에서 나타나는 것과 같이 구성된다.

상기 언급된 BSF(18)의 임펄스 응답의 자기상관은 다음의 관계식(3)으로 표현된 것과 같이 주파수 영역에서 BSF의 전력스펙트럼이다.

[수학식 3]

$$\phi(h) = F^{-1}(|H|^2)$$

H: BSF(18)의 특성주파수

여기에서 특성주파수을 가지는 또 다른 필터는 BSF(18)의 전력특성과 같고, 특성주파수는 H라고 가정한다. 관계식(4)은 다음과 같이 표현된다.

[수학식 4]

$$\phi(h) = F^{-1}(|H|^2) = F^{-1}(H') = h'$$

새로운 필터는 관계식(4)에 의해 표시되고 BSF(18)와 동등한 통과 및 저지구역을 가지고, 감쇠특성은 BSF(18)의 제곱이다. 그러므로, 새 필터는 대역스톱필터(BSF)라고 부른다.

상기 내용을 고려하면, 협대역 자기상관은 광대역 자기상관과 BSF의 임펄스 응답과의 컨벌루션 즉, 광대역 자기상관의 대역중지로부터 생겨난 다음의 관계식(5)으로 표현됨으로써 간략화된다.

[수학식 5]

$$\phi(x_n) = \phi(x_n) \otimes h'$$

음성코드북으로써 사용된 파라미터가 자기상관일 때, 유성음(V)상에서 자기상관 파라미터는 완만한 하향곡선 (즉 1차 자기상관 파라미터는 2차 파라미터보다 크고, 2차 파라미터는 3차 파라미터보다 크고,...)을 나타낸다.

한편, 협대역 음성신호와 광대역 음성신호의 관계는 광대역 음성신호가 협대역 음성신호를 제공하기 위하여 저역통과하는 것과 같다. 그러므로, 협대역 자기상관은 이론적으로 저역통과 광대역 자기상관에 의해 결정된다.

그러나, 광대역 자기상관이 완만하게 변하기 때문에 협대역 자기상관은 심지어 저역통과할지라도 거의 변화가 없는 것으로 나타난다. 그러므로, 저역통과는 생략해도 역효과가 없다. 즉, 광대역 자기상관은 협대역 자기상관으로써 사용될 수 있다. 그러나, 광대역 음성신호의 샘플링 주파수는 저대역 음성신호의 그것에 두배로 설정되기 때문에 저대역 자기상관은 실제로 매 다른 순서로 받아들여진다.

즉, 매 다른 순서로 받아들여진 광대역 자기상관 코드북은 협대역 자기상관 코드벡터와 동등하게 처리된다. 입력 협대역 음성의 자기상관은 광대역 음성코드북을 사용하여 양자화되고, 협대역 음성코드북은 필요없게 된다.

상기 언급된것과 같이, UV 음성은 V음성보다 더 큰 에너지를 가지고, 에러예측은 광범위한 영향을 미친다.

이것을 피하기 위하여, V/UV 판별장치는 V와 UV를 판별하기에 어려운 음성신호를 V로 취하도록 설정된다. 즉, 음성신호는 음성신호가 UV일 확률이 높을 때, UV로 판단된다. 이런 이유 때문에, UV음성코드북은 서로 다른 코드 벡터만을 기록하기 위하여 V 음성코드북보다 크기가 더 작다. 그러므로, UV의 자기상관은 매 다른 순서로 취해진 광역 자기상관 코드벡터와 비교하여 V의 자기상관만큼 완만한 곡선이 아닐지라도 입력 협대역 신호의 자기상관은 저역통과된 광대역 자기상관 코드벡터(즉 협대역 음성코드북을 입수할 수 있을 때 양자화)의 그것에 동등한 협대역 입력 음성신호의 양자화로 도달할 수 있게 한다. 즉, V와 UV 음성은 협대역 음성코드북 없이 양자화될 수 있다.

상기 설명된 바와 같이, 자기상관이 음성코드북에서 사용된 파라미터와 같이 취해질 때, 입력 협대역 음성의 자기상관은 매 다른 순서로 취해진 광대역 코드벡터와 비교하여 양자화될 수 있다. 이런 동작은 부분추출회로(28, 29)가 도 7의 매 다른 순서로 광역 음성코드북의 코드벡터를 취하도록 함으로써 실현될 수 있다. (단계(S45))

이제, 음성 코드북상에서 파라미터로써 스펙트럼 엔벌로프을 사용한 양자화는 이하에 설명한다. 이런 경우, 협대역 스펙트럼이 일부의 광대역 스펙트럼이기 때문에, 어떤 협대역 스펙트럼 음성코드북도 양자화하는데 요구되지 않는다. 말할 필요도 없이, 입력 협대역 음성의 스펙트럼 엔벌로프는 일부의 광대역 스펙트럼 엔벌로프 코드북과 비교를 통하여 양자화될 수 있다.

다음에, 본 발명에 따라서 음성합성방법과 음성합성장치는 본 발명의 음성 합성장치의 실시의 예에 따른 수신장치에 적용된 디지털 휴대전화기를 나타내는 블록 다이어그램인 도 9를 참조하여 설명한다. 이 실시의 예는 광대역 음성에서 매 소정의 시간단위로, 추출된 특성파라미터로부터 형성된 광대역 음성코드북을 포함하여 구성하고, 음성을 복수의 입력 부호화된 파라미터를 사용하여 합성하는데 적용된다.

도 9에서 나타내는 휴대용 디지털 전화 시스템의 수신장치부측의 음성합성장치는 음성 디코더(38)과 음성 합성장치(39)를 포함하여 구성된다.

휴대용 디지털 전화기는 아래 설명된 것과 같이 구성된다. 물론, 송신장치와 수신장치는 실제로 휴대 전화기셋트로 함께 합체되지만, 설명의 편리를 위하여 분리하여 설명될 것이다.

디지털 휴대 전화 시스템의 송신부측에서, 마이크로폰(31)을 통하여 입력으로써 공급된 음성신호는 A/D변환장치(32)에 의해 디지털 신호로, 음성 엔코더 장치(33)에 의해 엔코드로 변환되고, 안테나(35)로부터 그것을 송신하는 송신장치(34)에 의해 출력비트로 처리된다. 음성 엔코더 장치(33)는 송신장치(34)에 부호화된 파라미터를 협대역 신호로 경로제한된 송신 변환을 고려하여 제공한다. 예를 들면, 부호화된 파라미터는 여진원관련 파라미터와 선형예측계수(α) 등을 포함한다.

수신장치부에서, 안테나(36)에 의해 포획된 파는 수신장치(37)에 의해 검파되고, 파에 의해 옮겨진 부호화된 파라미터는 음성 디코더 장치(38)에 의해 디코더되고, 음성은 음성합성장치(39)에 의해 부호화된 파라미터를 사용하여 합성되고, 합성된 음성은 D/A 변환 장치(40)에 의해 아날로그 음성으로 변환되고, 스피커(41)로 옮겨진다.

도 10은 본 발명이 디지털 휴대 전화셋에서 사용된 음성합성장치의 첫 번째 실시의 예를 나타내는 블록 다이어그램이다. 도 10의 음성합성장치는 음성을 음성 엔코더장치(33)으로부터 디지털 휴대용 전화 시스템의 송신장치부로 보내진 부호화된 파라미터를 사용하여 합성하고, 수신부측의 음성 디코더(38)는 엔코드된 음성신호를 음성이 송신장치부의 음성 엔코더(33)에 의해 엔코드되는 모드로 디코드하도록 되어있다.

즉, 음성신호가 PSI-CELP(Pitch Synchronous Innovation-Code Excited Linear Prediction) 모드에서 음성 엔코더(33)에 의해 부호화될 때, 음성 디코더(38)은 PSI-CELP모드를 송신장치측으로부터 엔코드된 음성 신호를 디코드하기위해 채택한다.

음성 디코더(38)는 부호화된 첫 번째 파라미터인 여진원관련 파라미터를 협대역 여진원으로 디코드하고, 그것을 제로필링 회로(16)에 제공한다. 그것은 또한 α/ɣ변환장치(4)(α: 선형예측계수, ɣ: 자기상관)에서 부호화된 두 번째 파라미터를 선형예측계수로 변환한다. 더욱이, 그것은 부호화된 세 번째 파라미터인 유/무성음 표지관련 신호를 가지고 있는 V/UV 판별장치(5)를 제공한다.

음성합성장치는 또한 음성 디코더(38)과 제로필링회로(16)과 α/ɣ변환장치(4)와 V/UV 판별장치(5)에 더하여 광대역과 무성음으로부터 추출된 유/무성음 파라미터를 사용하여 형성된 광대역 유성음 코드북(12)과 광대역 무성음 코드북(14)을 포함하여 구성한다.

도 10에서와 같이, 음성합성장치는 또한 광대역 유성음 코드북(12)과 광대역 무성음 코드북(14)에서 각 코드벡터의 부분추출을 통하여 협대역 파라미터를 결정하기 위한 부분추출회로(28, 29)와 부분추출회로(28)에서 협대역 파라미터를 이용한 α/ɣ변환 장치(4)로부터 협대역 유성음 자기상관을 양자화하기 위한 협대역 유성음 양자화 장치(7)와 부분추출회로(29)에서 협대역 파라미터를 사용한 α/ɣ변환 장치(4)로부터 협대역 무성음 자기상관을 양자화하기 위한 협대역 무성음 양자화 장치(9)와 협대역 유성음 양자화 장치(7)로부터 광대역 유성음 코드북(12)을 이용하여 협대역 유성음 양자화된 데이터를 역양자화하기 위한 광대역 유성음 역양자화 장치(11)와 협대역 무성음 양자화 장치(9)로부터 협대역 무성음 코드북(14)을

이용하여 협대역 무성의 양자화된 데이터를 역양자화하기 위한 광대역 무성음 역양자화 장치(13)와 협대역 유성음 역양자화 장치(11)로부터 협대역 유성음 선형예측계수로 광대역 유성음 자기상관 (역양자화된 데이터)과 광대역 무성음 역양자화 장치(13)로부터 협대역 무성음 선형예측계수로 광대역 무성음 자기상관(역양자화된 데이터)을 변환하기 위한 γ/α변환기(15)와 그리고 γ/α변환기(15)로부터 협대역 유/무성음 선형예측계수와 제로필링회로(16)으로부터 여진원을 기초로 한 광대역 음성을 합성하기 위한 LPC 합성장치(17)을 포함하여 구성한다.

음성 합성장치는 또한 8kH에서 16zkHz의 음성디코더(38)에 의해 디코더된 협대역 음성 데이터의 샘플링 주파수를 변화하기 위해 제공된 오버샘플링회로(19)와 LPC 합성장치(17)에서 합성된 출력으로부터 입력 협대역 유성음 신호의 주파수 대역에서 300~3400Hz의 신호를 제거하기 위한 대역스톱필터(band stop filter)와 그리고 오버샘플링회로(19)로부터 협대역 유성음 신호의 샘플링 주파수상에서 16kHz와 주파수 대역상에서 300~3400kHz의 신호를 BSF 필터(18)로부터 출력으로 더하기 위한 가산장치(20)를 포함하여 구성한다.

광대역 유/무성음 코드북(12, 14)은 도 2에서 도 4까지 나타나는 절차에 따라서 형성된다. 양질의 음성코드북을 위하여, 유성음(V)으로부터 무성음(UV)까지의 (또한 그역도 같음) 전송에 있어서의 성분과 V와 UV를 판별하기 위해 어려운 성분은 확실하게 V 와 UV 되는 음성만 공급하기 위하여 제거된다. 따라서, 학습 협대역 V 프레임의 집합과 학습 협대역 UV 프레임의 집합이 얻어진다.

광대역 유/무성음 코드북(12, 14)을 사용하는 음성 합성장치뿐만아니라 실제 전송장치부로부터 전송된 부호화된 파라미터는 도 10의 음성합성장치의 동작을 나타내는 플로우 챠트인 도 11을 참조하여 설명된다.

우선, 음성 디코더(38)에 의해 복호화된 선형예측계수(α)는 단계 S61에서 α/γ변환장치(4)에 의해 자기상관(γ)으로 변환된다.

유/무성(V/UV) 판별 표지관련 파라미터는 단계 S62의 V/UV 판별장치(5)에서 V와 UV 음성이 판별된 음성디코더(38)에 의해서 복호화된다.

프레임화된 신호가 V로 판단될 때, 스위치(6)은 협대역 유성음 양자화 장치(7)로 신호경로를 연결한다. 반대로, 신호가 UV로 판단될 때, 스위치(6)은 α/γ변환장치(4)의 출력을 협대역 UV 음성 양자화장치(9)로 연결한다.

그러나, 단계 S22에서 실행된 V/UV 판별은 음성코드북 생성을 위해 실행된 것과는 다르다는 것에 주목하라. 즉, V 도 UV 도 아닌 어떤 프레임이 생성될 수 있다. V/UV 판별장치(5)에서, 프레임 신호는 틀림없이 V나 UV중 하나로 판단될 것이다.

V/UV 판별장치(5)가 입력 음성신호를 V음성으로 판단할 때, 스위치(6)로부터 유성음 자기상관(γ)은 협대역 V 음성 양자화 장치(7)로 제공되고, 그것은 협대역 음성코드북을 사용하지 않는 단계 S63의 부분추출회로(28)에 의해 정해진 협대역 V음성 파라미터를 이용하여 단계 S64에서 양자화된다.

반대로, V/UV 판독장치(5)가 입력 음성신호를 UV로 판단할 때, 스위치(6)으로부터 무성음 자기상관(g)는 협대역 UV 양자화 장치(9)로 제공되고, 협대역 UV 장치는 협대역 UV 음성코드북을 사용하지 않고, 부분추출회로(29)에서 연산에 의해 결정된 협대역 UV 파라미터를 사용하는 것에 의해 단계(S63)에서 양자화된다.

단계(S65)에서, 광대역 V양자화 장치(11) 또는 광대역 UV 역양자화 장치(13)는 광대역 V 음성코드북(12) 또는 광대역 UV 음성 코드북(14)을 각각 사용하여 양자화된 자기상관(g)을 역양자화하고, 광대역 자기상관을 제공한다.

단계 S66에서, 광대역 자기상관(γ)은 γ/α변환장치(15)에 의해 광대역 자기상관(α)으로 변환된다.

한편, 음성 디코더(38)로부터 여진원관련 파라미터는 단계 S67의 제로필링 회로(16)에 의해 샘플들을 제로필링함으로써 광대역을 갖기 위해 엘리어스(alias)되고, 업샘플된다. 그것은광대역 여진원으로써 LPC 합성장치(17)로 제공된다.

단계 S68에서, 광대역 자기상관(α)과 광대역 여진원은 광대역 음성신호를 제공하기 위하여 LPC 합성장치(17)에서 LPC 합성이 필요하다.

그러나, 얻어진 광대역 음성신호는 예측으로부터 생성된 것이고, 그것은 그밖에 처리되는 것이 없다면, 예측에러를 포함한다. 특히, 입력 협대역 음성은 가능한한 주파수 대역에서 만나지 않고 남겨져야 한다.

그러므로, 단계 S69에서, 입력 협대역 음성은 BSF(18)에 의해 필터링을 통하여 제거된 주파수 범위를 가지고, 단계 S71의 오버샘플링회로(19)에 의해 샘플된 엔코드된 음성 데이터로 단계 S70에서 가산된다.

도 10의 음성 합성장치는 어떤 협대역 음성 코드북에서 코드북과 비교하는 것이 아니라 광대역 음성 코드북으로부터 부분 추출에 의해 결정된 코드 벡터를 비교함으로써 양자화되도록 적용된다.

즉, 파라미터(α)가 복호화의 과정에서 얻어지기 때문에, 그것은 협대역 자기상관(ɣ)으로 변환된다. 협대역 자기상관(ɣ)은 광대역 음성코드북상에서 매 다른 순서로 취해지며 각 벡터와 비교함으로써 양자화된다. 그리고, 양자화된 협대역 자기상관은 광대역 자기상관을 제공하기 위해 모든 벡터를 사용하여 역양자화된다. 이 광대역 자기상관은 광대역 선형 예측펙터(a)로 변환된다. 이 이득제어와 고 대역의 약간의 억제는 상기에서 청강상의 음질은 향상시키기 위하여 설명된 것과 같이 영향을 받는다. 그러므로, 광대역 음성코드북은 음성신호를 분석하고 합성하는데 사용되고, 협대역 음성코드북을 저장하기 위한 메모리는 불필요하다.

도 12는 도10의 다양한 음성합성장치를 나타내는 블록 다이어그램이고, 그것은 PSI-CELP 부호화 모드를 채택하는 음성 디코더(38)로부터 부호화된 파리미터가 적용된다. 도 12에서 나타내는 음성합성장치는 부분추출회로(18, 19)대신에 광대역 음성코드북상에서 각 코드벡터를 연산함으로써 협대역 V (UV) 파라미터를 제공하기 위한 연산회로(28, 29)를 사용한다. 이 음성 합성장치의 나머지부분은 도 10에서 나타내는 것과 같이 구성된다.

도 13을 디지털 휴대 전화셋에서 사용된 본 발명 음성합성 장치의 실시의 예를 나타내는 블록 다이어그램이다. 도 13에서 나타내는 음성합성장치는 음성 엔코더(33)으로부터 디지털 휴대전화시스템의 전송장치부에서 보내진 부호화된 파라미터를 사용하는 음성을 합성하도록 되어지고, 수신장치부측 음성합성장치에 있는 음성 디코더(46)는 음성이 송신부를 음성 엔코더(33)에 의해 엔코더된 모드에서 엔코드된 음성신호를 디코드한다.

즉, 음성신호가 VSELP(Vector Sum Excited Linear Prediction) 모드상에서 음성 엔코더에 의해 엔코드될 때, 음성 디코더(6)는 송신장치부측으로부터 엔코드된 신호를 디코드하기위해 VSELP 모드를 채택한다.

음성 디코더(46)는 부호화된 첫 번째 파라미터가 되는 여진원 관련 파라미터를 여진원 선택장치(47)로 제공한다. 또한 부호화된 두 번째 파리미터가 되는 선형여진원 펙터를 α/ɣ변환장치(4)로 제공한다. 또한, V/UV 판별장치(5)에 부호화된 세 번째 파라미터가 되는 유/무성음 표지 관련 신호를 공급한다.

음성 디코더에서 VSELP 모드를 사용하는 본 발명의 음성합성 장치를 나타내는 도 13의 음성합성장치는 도 10과 12에서 나타내는 그것들과 다르고, 여진원 선택장치(47)는 제로필링회로(16)의 상부로 제공되는 PSI-CELP를 사용한다.

PSI-CELP 모드상에서, CODEC(coder/decoder)은 청감상 부드러운 음성을 제공하기 위해 유성음 신호를 처리하는 반면, VSELP모드에서는 CODEC은 청감상 부드럽지 않고, 일부 노이즈를 포함하는 대역 확장된 음성을 제공한다. 도 13의 음성합성 장치의 동작을 나타내는 플로우챠트인 도 14에서와 같이 이노베이션 선택장치(47)에 의해 처리된다. 도 14에서의 처리절차는 단지 단계 S87에서 S89까지 첨가적으로 실행되지만, 도 11과는 다르다.

VSELP 모드에 대해, 여진원은 CODEC상에서 사용된 파라미터 베타(long-term prediction factor)와 *bL[i](long-term filtering)와 cl[i](excited code vector)로부터 베타 *bL[i]+ 감마(gamma) 1 *cl[i]로써 형성된다. 베타 *bL[i]는 감마 1*cl[i]가 노이즈 성분을 나타내는 동안 피치성분을 나타낸다. 그러므로, 여진원은 베타 *bL[i]와 감마 1 *cl[i] 로 분리된다. 전자가 단계 S87에서 소정의 시간에 대해 고에너지를 나타낼 때, 입력음성신호는 강한 피치를 가지는 유성으로 된다. 그러므로, 동작은 단계 S88 에서 YES이고, 여진원으로써 임펄스열을 취한다. 여진원이 피치성분을 가지지 않는 것으로 판단될 때, 동작은 여진원을 0으로 억제하기 위해 NO이다. 또한, 형성된 협대역 여진원이 단계 S89에서 PSL-CELP 모드와 같이 제로 필링회로(16)에 의해 제로필링에 의해 업샘플되고, 광대역 여진원을 생성한다. 그리고 나서, VSELP모드에서 생성된 유성음은 청감상의 질을 향상시킨다.

더욱이, VSELP 모드를 채택하는 음성 디코더(46)로부터 부호화된 파라미터를 사용하여 음성을 합성하기 위한 음성합성장치는 음성 디코더상에 VSELP 모드를 채택하여 음성합성장치를 나타내는 블록 다이어그램인 도 15에서와 같이 본 발명에 따라서 제공된다. 도 15에서 음성합성장치는 부분추출회로(28, 29)대신에 광대역 음성코드북상에서 각 코드벡터의 연산에 의해 협대역 V/UV파라미터를 제공하기 위한 연산회로(25, 26)를 포함하여 구성한다. 이 음성합성장치의 나머지 부분은 도 13에서 나타내는 것과 같이 구성된다.

도 15의 음성합성장치는 도 1에서 나타내는 바와 같이 광대역 유/무성음으로부터 추출된 유/무성음 파라미터를 사용하여 형성된 광대역 유/무성음 코드북(12, 14)과 또한 광대역 유성음의 주파수 대역을 제한함으로써 생성된주파수 대역상에서 300~3400Hz의 협대역 음성신호로부터 추출된 유/무성음 파라미터를 사용하여 형성된 협대역 유/무성음 코드북(8, 10)을 사용하여 음성을 합성할 수 있다.

이 음성합성장치는 저주파수 대역으로부터 고주파수 대역으로의 예측으로 제한되지 않는다. 또한, 광대역 스펙트럼을 예측하는 수단에서, 신호는 음성에 제한되지 않는다.

발명의 효과

광대역 여진원으로 임펄스열을 받아들임으로써 음성피치가 강할 때, 특히 본발명에 따라서 청감상 유성음의 질은 향상될 수 있다.

도면의 간단한 설명

도 1은 본 발명의 음성 대역폭 확장장치의 실시의 예를 나타내는 블록 다이어그램이다.

도 2는 도 1의 음성 대역폭 확장장치에서 사용된 음성 코드북용 데이터 생성을 나타내는 플로우 챠트이다.

도 3은 도 1의 음성 대역폭 확장장치에서 사용된 음성 코드북의 생성을 나타내는 플로우 챠트이다.

도 4는 도 1의 음성 대역폭 확장 장치에서 사용된 다른 음성 코드북의 생성을 나타내는 플로우 챠트이다.

도 5는 도 1의 음성 대역폭 확장장치의 동작을 나타내는 플로우 챠트이다.

도 6은 도 1에서 감소된 수의 음성코드북이 사용된 또다른 음성 대역폭 확장장치를 나타내는 플로우 챠트이다.

도 7은 도 6의 또다른 음성 대역폭 음성장치의 동작을 나타내는 플로우 챠트이다.

도 8은 도 1에서 감소된 수의 음성코드북이 사용된 또다른 음성 대역폭 확장장치를 나타내는 블록 다이어그램이다.

도 9는 본 발명의 음성합성장치상에서 수신기에서 적용되고 있는 디지털 휴대 전화장치의 구성을 나타내는 블록 다이어그램이다.

도 10은 음성합성장치의 음성디코더에서 PSI-CELP 부호화 모드를 채택하는 음성합성장치를 나타내는 블록 다이어그램이다.

도 11은 도 10의 음성합성장치의 동작을 나타내는 플로우 챠트이다.

도 12는 도 10의 음성합성장치의 음성디코더에서 PSI-CELP 부호화 모드를 채택하는 또다른 음성합성장치를 나타내는 블록 다이어그램이다.

도 13은 음성합성장치의 음성 디코더에서 VSELP 모드를 채택하는 음성합성장치를 나타내는 블록 다이어그램이다.

도 14는 도 13의 음성합성장치의 동작을 나타내는 플로우 챠트이다.

도 15는 음성합성장치의 음성디코더에서 VSELP 모드를 채택하는 음성합성장치를 나타내는 블록 다이어그램이다.

* 도면의 주요부분에 대한 부호설명

2. 프레이밍회로 3. LPC 분석장치

4. α/γ변환장치 5. V/UV 판별장치

6. 스위치 7. 협대역 V 양자화장치

8. 협대역 V 코드북 9. 협대역 UV 양자화장치

10. 협대역 UV 코드북 11. 광대역 V 역양자화장치

12. 광대역 V 코드북 13. 광대역 UV 역양자화장치

14. 광대역 UV 코드북 15. γ/α변환장치

16. 제로필링회로 17. LPC 합성장치

18. 대역스톱필터 19. 오버 샘플링 회로

20. 가산장치 25. 26. 연산회로

28. 29. 부분추출회로 32. A/D 변환장치

33. 음성엔코더 34. 송신장치

37. 수신장치 38. 음성엔코더(PSI-CELP)

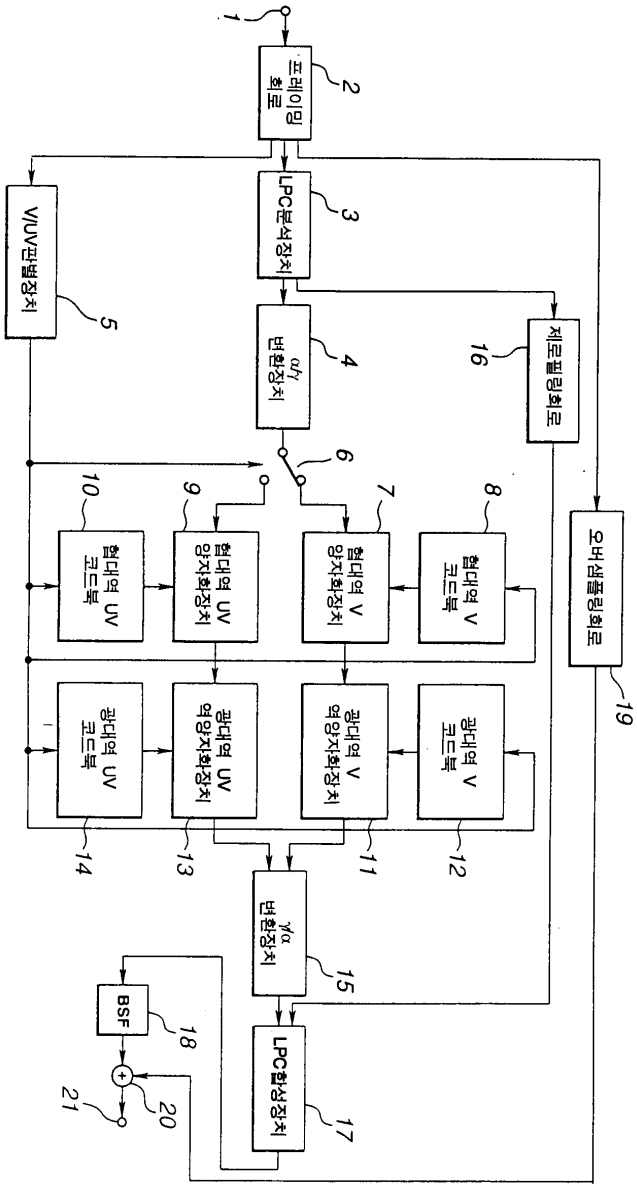
39. 음성합성장치 40. D/A 변환장치

41. 스피커 46. 음성디코더 (VSELP)

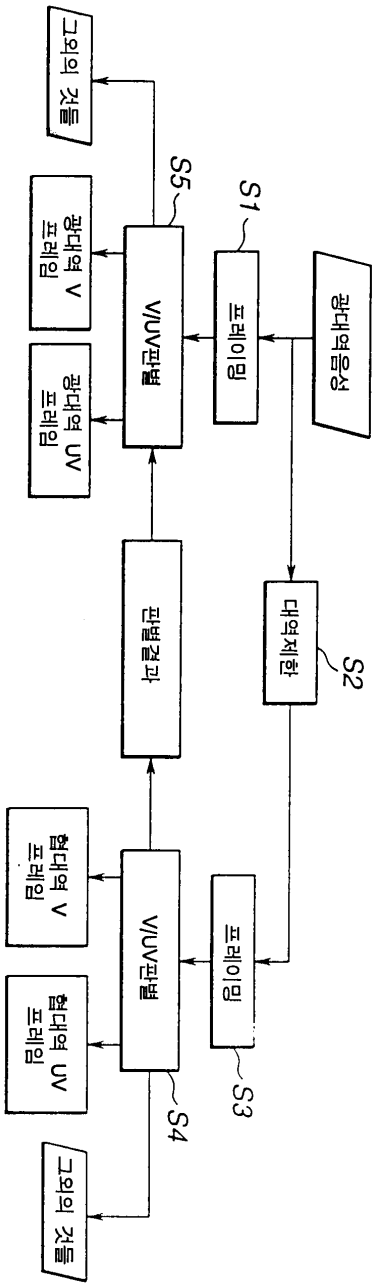
47. 여진원 선택장치

도면

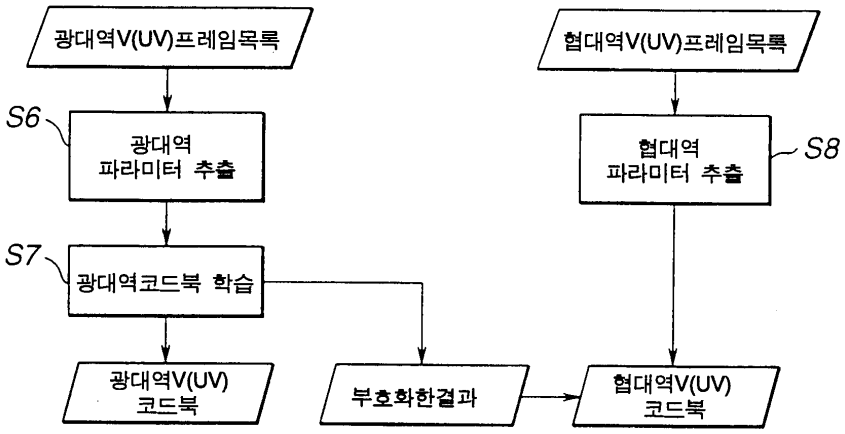
도면1


1
프레이밍 회로
2
LPC분석장치
3
V/UV판별장치
5
α/γ 변환장치
4
제로필링회로
16
6
협대역 UV 코드북
10
협대역 UV 양자화장치
9
협대역 V 양자화장치
7
협대역 V 코드북
8
오버샘플링회로
19
광대역 UV 코드북
14
광대역 UV 역양자화장치
13
광대역 V 역양자화장치
11
광대역 V 코드북
12
γ/α 변환장치
15
BSF
18
LPC합성장치
17
20
21

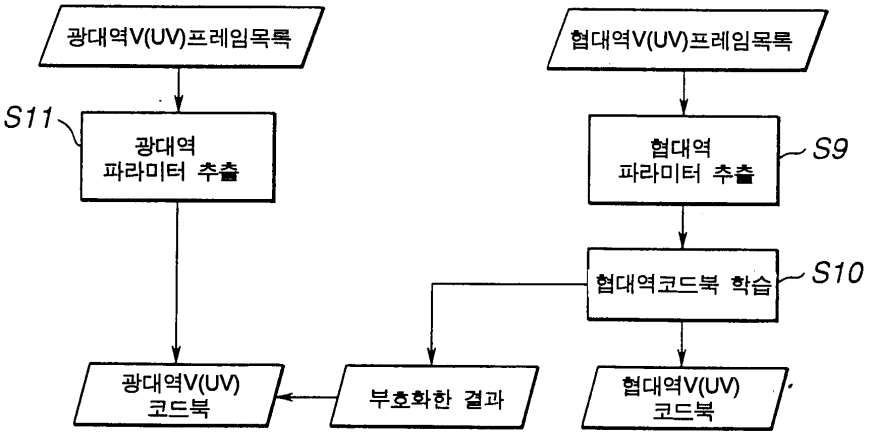
도면2


그외의 것들
광대역 V 프레임
광대역 UV 프레임
V/UV판별
S5
S1
프레이밍
광대역음성
판별결과
대역제한
S2
협대역 V 프레임
협대역 UV 프레임
V/UV판별
프레이밍
S4
S3
그외의 것들

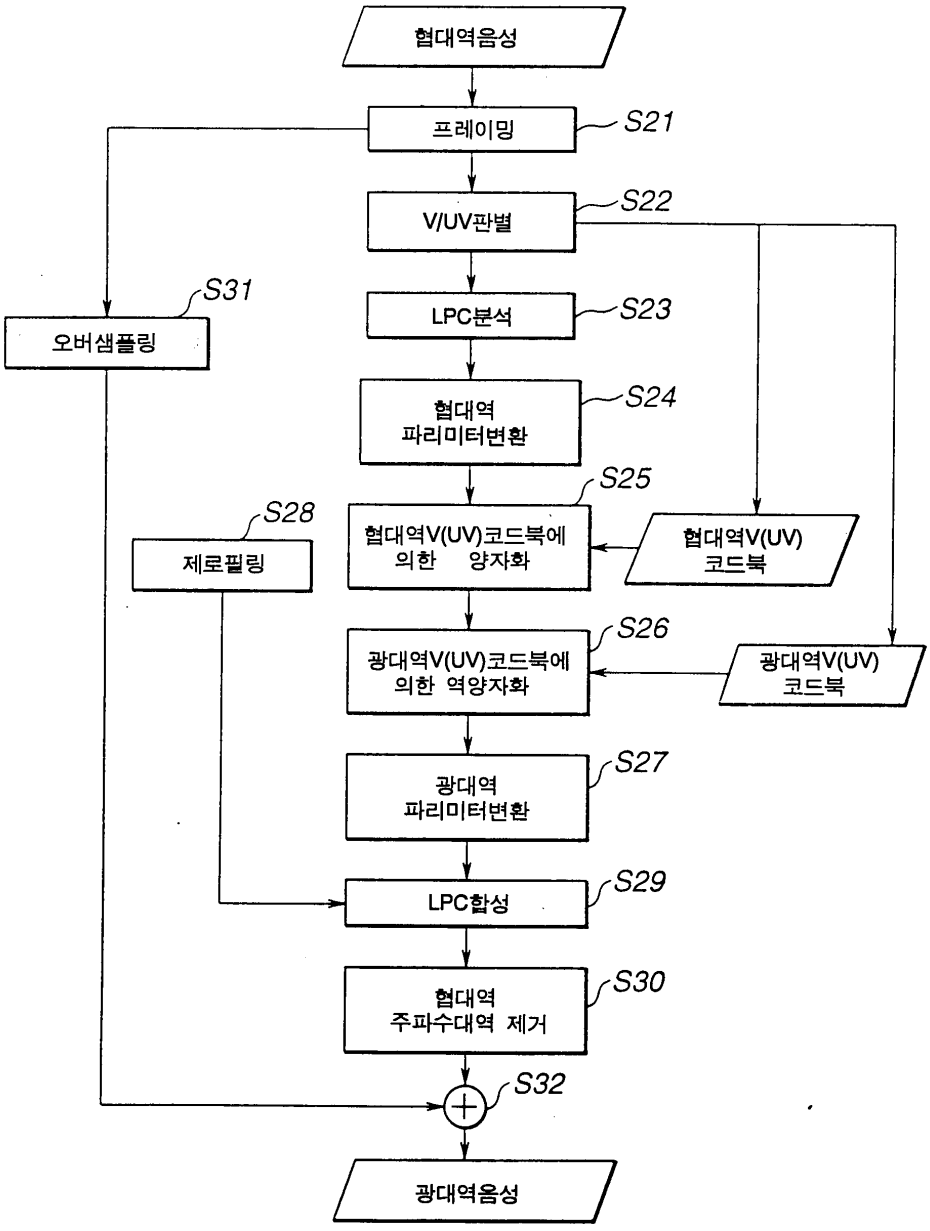
도면3


광대역V(UV)프레임목록
협대역V(UV)프레임목록
S6
광대역 파라미터 추출
협대역 파라미터 추출
S8
S7
광대역코드북 학습
광대역V(UV) 코드북
부호화한결과
협대역V(UV) 코드북

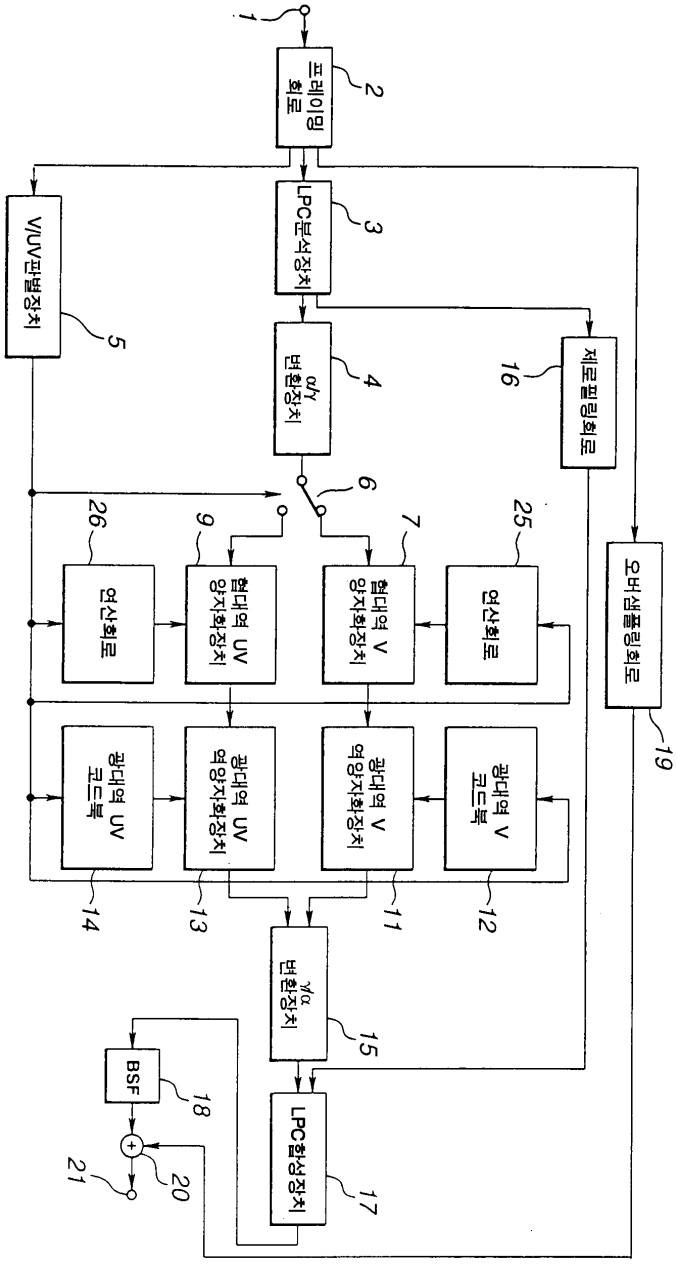
도면4


광대역V(UV)프레임목록
협대역V(UV)프레임목록
S11
광대역
파라미터 추출
협대역
파라미터 추출
S9
협대역코드북 학습
S10
광대역V(UV)
코드북
부호화한 결과
협대역V(UV)
코드북

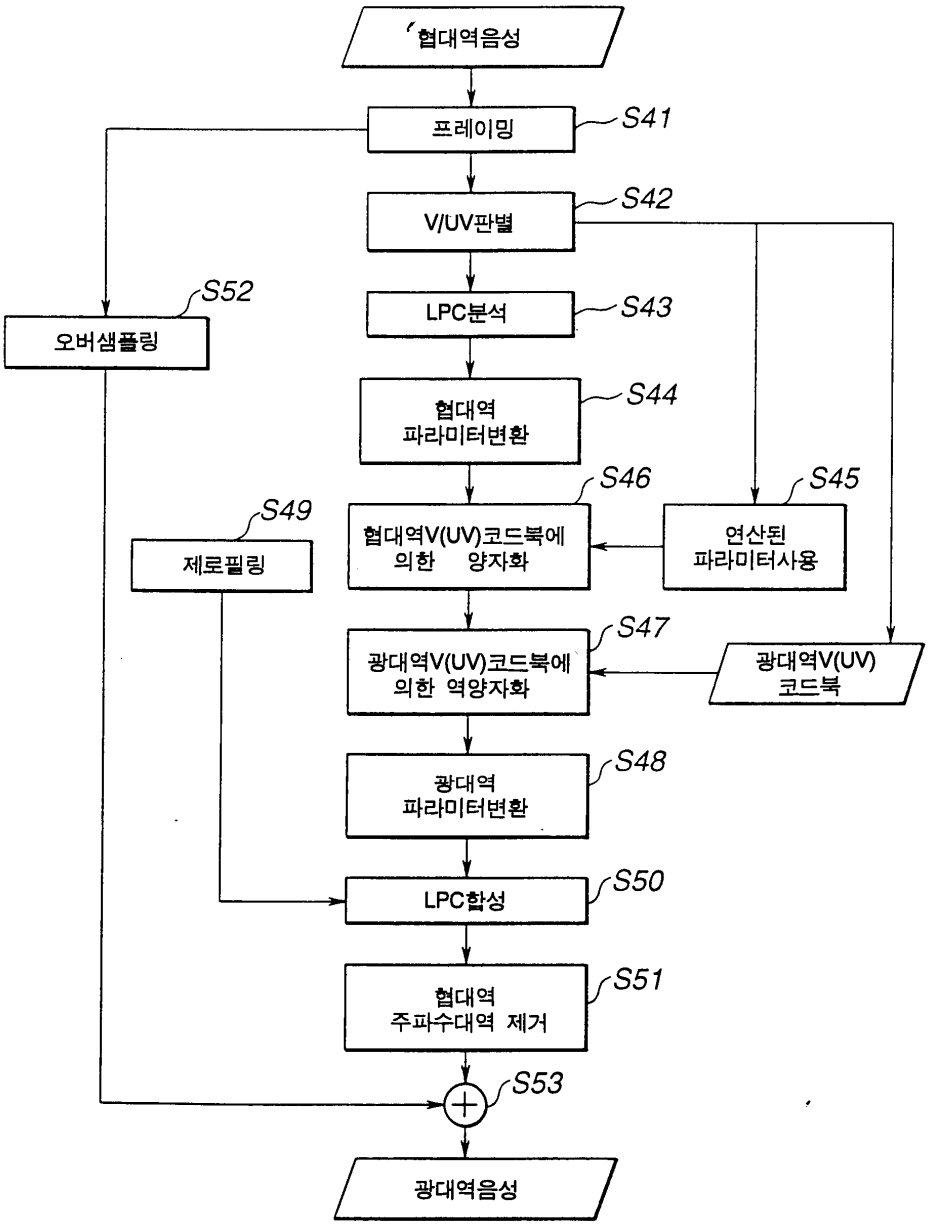
도면5


협대역음성
프레이밍
S21
V/UV판별
S22
S31
오버샘플링
LPC분석
S23
협대역
파리미터변환
S24
S25
S28
제로필링
협대역V(UV)코드북에
의한 양자화
협대역V(UV)
코드북
S26
광대역V(UV)코드북에
의한 역양자화
광대역V(UV)
코드북
광대역
파리미터변환
S27
LPC합성
S29
협대역
주파수대역 제거
S30
S32
광대역음성

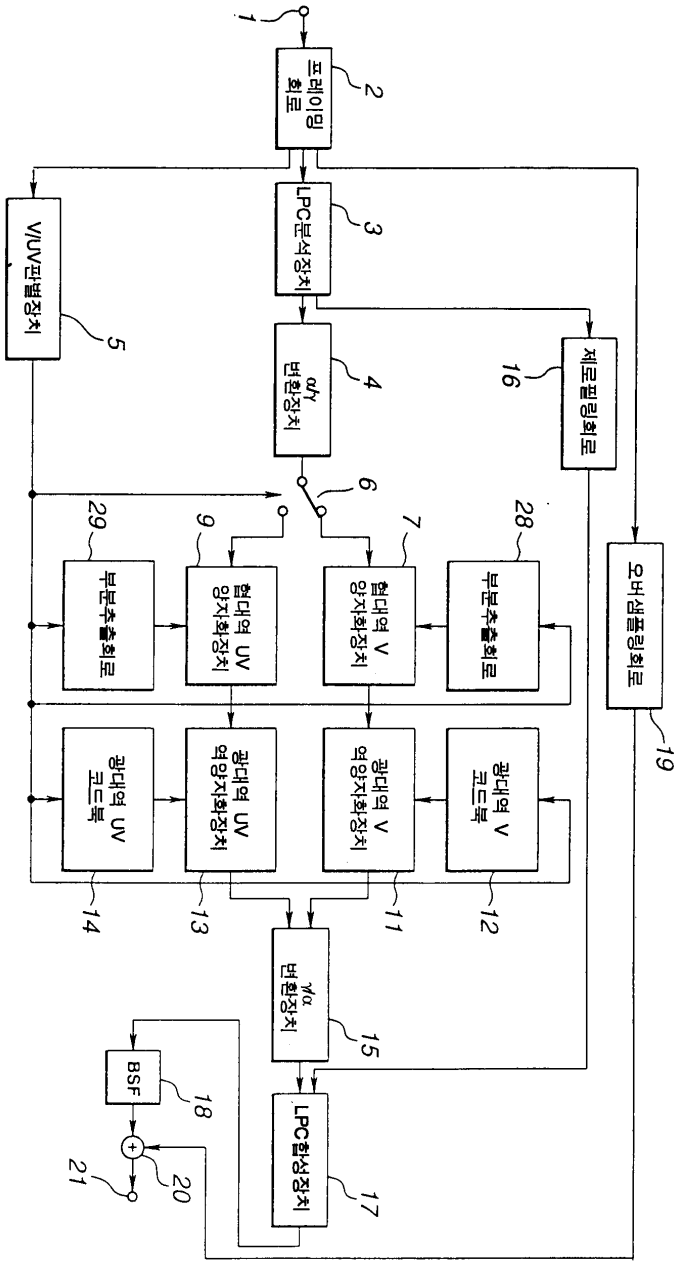
도면6


1
프레이밍 회로
2
LPC분석장치
3
V/UV판별장치
5
α/γ 변환장치
4
제로필링회로
16
6
26
9
7
25
연산회로
협대역 UV 양자화장치
협대역 V 양자화장치
연산회로
오버샘플링회로
19
광대역 UV 코드북
광대역 UV 역양자화장치
광대역 V 역양자화장치
광대역 V 코드북
14
13
11
12
γ/α 변환장치
15
BSF
18
LPC합성장치
17
20
21

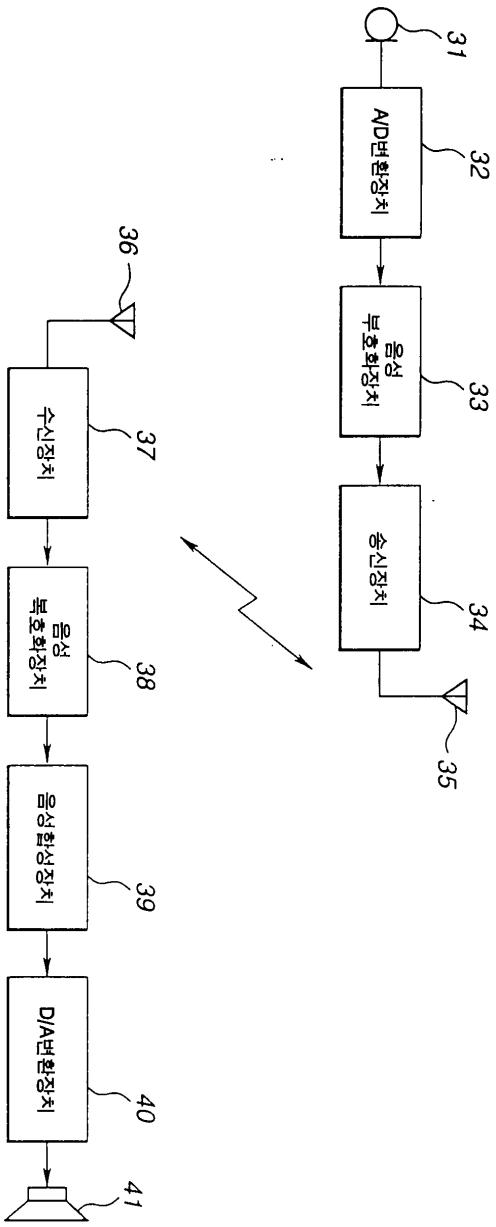
도면7


협대역음성
프레이밍
S41
V/UV판별
S42
LPC분석
S43
협대역
파라미터변환
S44
S52
오버샘플링
S49
제로필링
협대역V(UV)코드북에
의한 양자화
S46
연산된
파라미터사용
S45
광대역V(UV)코드북에
의한 역양자화
S47
광대역V(UV)
코드북
광대역
파라미터변환
S48
LPC합성
S50
협대역
주파수대역 제거
S51
S53
광대역음성

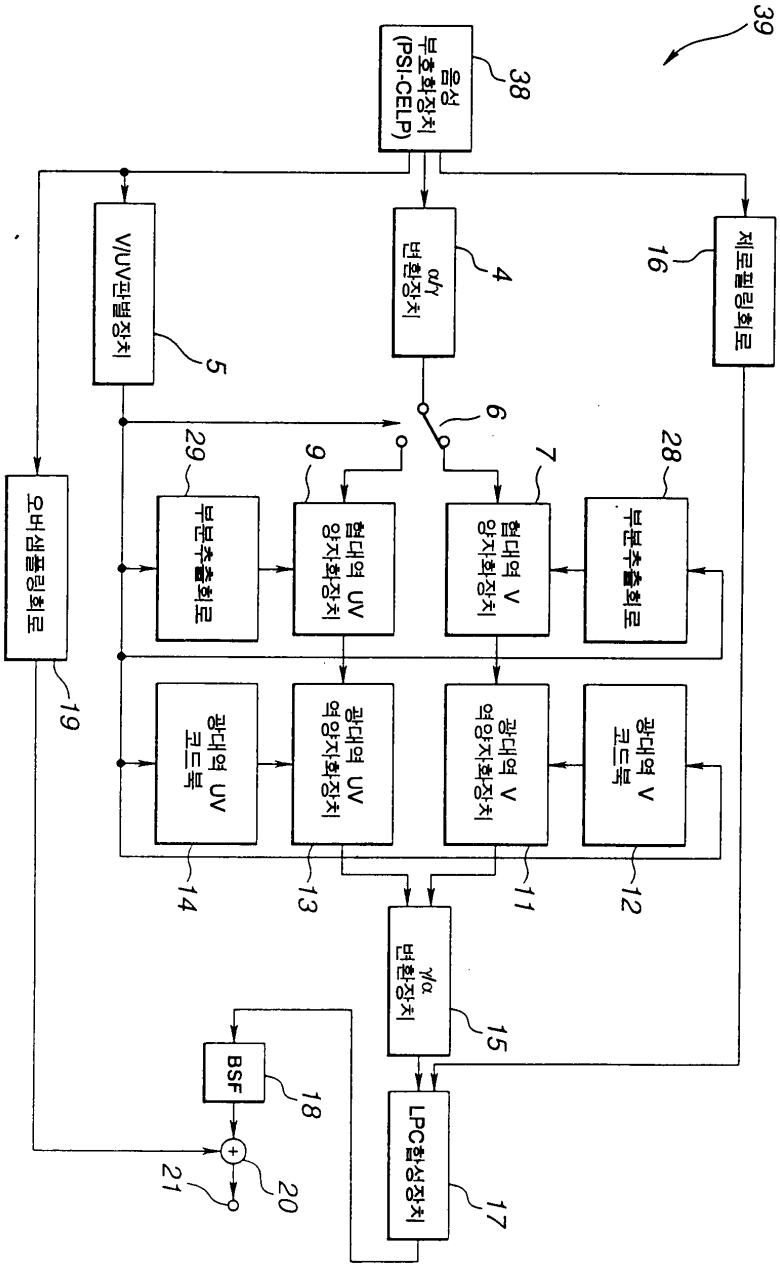
도면8


1
프레이밍 회로
2
LPC분석장치
3
V/UV판별장치
5
αγ 변환장치
4
제로필링회로
16
6
오버샘플링회로
19
29
부분추출회로
9
협대역 UV 양자화장치
7
협대역 V 양자화장치
28
부분추출회로
광대역 UV 코드북
14
광대역 UV 역양자화장치
13
광대역 V 역양자화장치
11
광대역 V 코드북
12
γα 변환장치
15
BSF
18
LPC합성장치
17
21
20

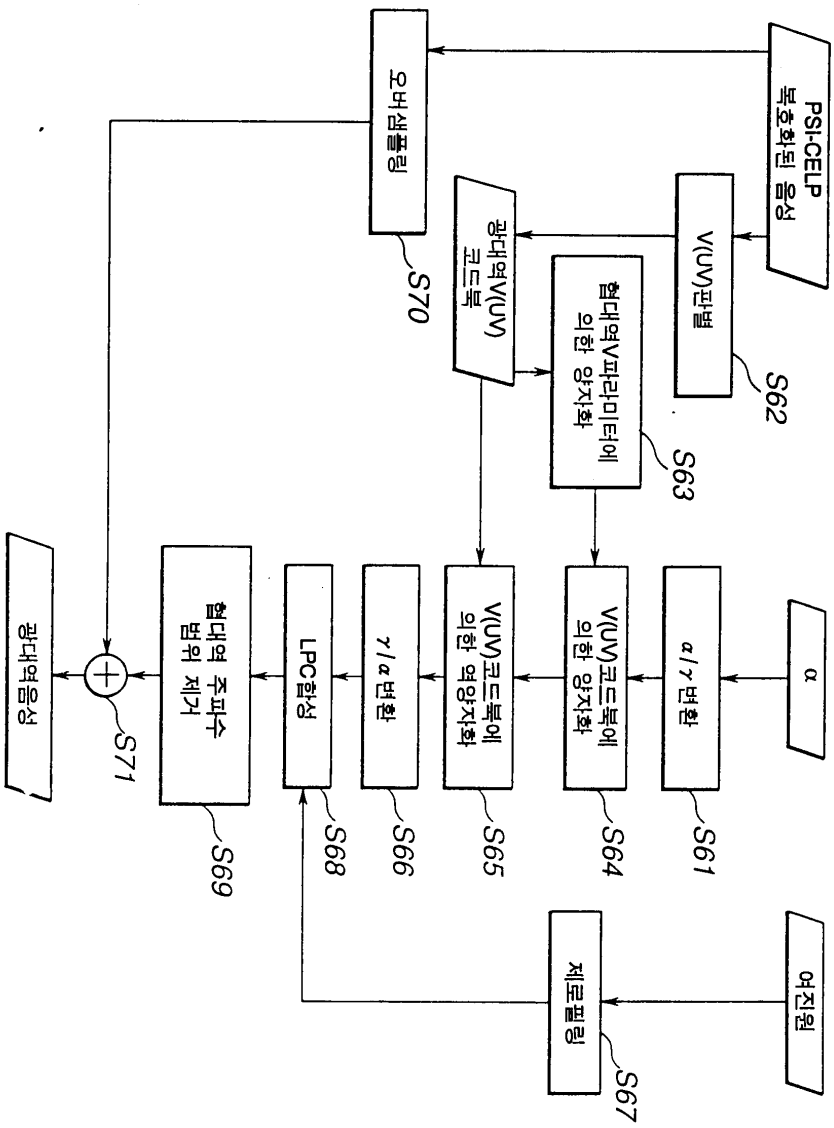
도면9


31
A/D변환장치
32
음성
부호화장치
33
송신장치
34
35
36
수신장치
37
음성
복호화장치
38
음성합성장치
39
D/A변환장치
40
41

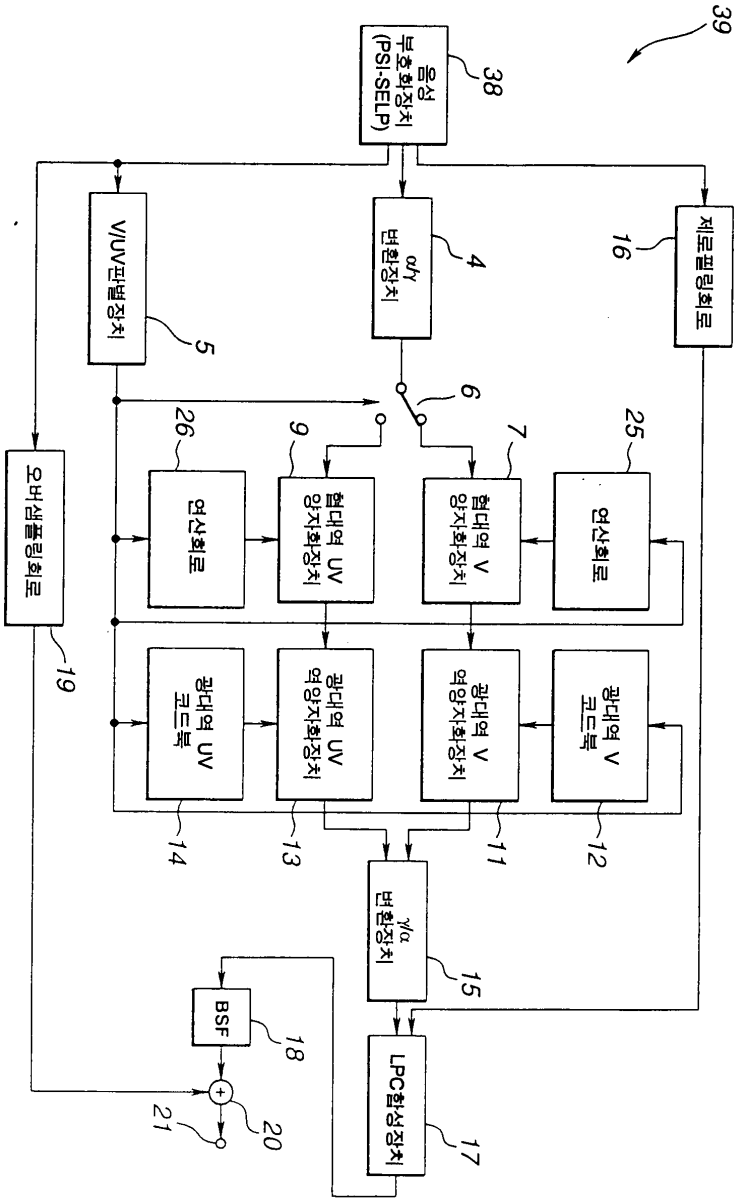
도면10


39
음성 부호화장치 (PSI-CELP)
38
V/UV판별장치
5
ω/γ 변환장치
4
제로필링회로
16
6
오버샘플링회로
19
부분추출회로
29
협대역 UV 양자화장치
9
협대역 V 양자화장치
7
부분추출회로
28
광대역 UV 코드북
14
광대역 UV 역양자화장치
13
광대역 V 역양자화장치
11
광대역 V 코드북
12
γ/α 변환장치
15
BSF
18
LPC합성장치
17
20
21

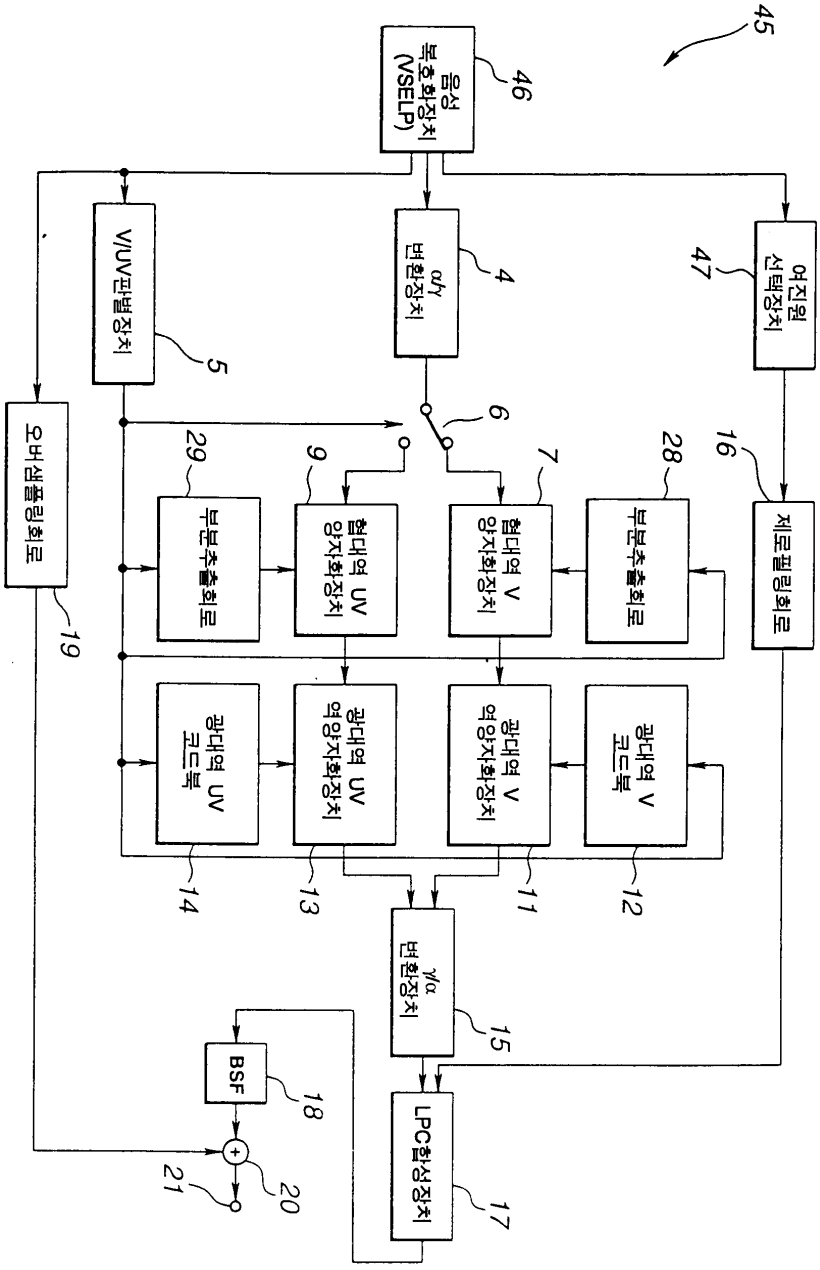
도면11


PSI-CELP 복호화된 음성
V(UV)판별
S62
오버샘플링
S70
광대역V(UV) 코드북
협대역V파라미터에 의한 양자화
S63
α
α/γ 변환
V(UV)코드북에 의한 양자화
S64
S61
V(UV)코드북에 의한 역양자화
S65
γ/α 변환
S66
LPC합성
S68
협대역 주파수 범위 제거
S69
광대역음성
S71
여진원
제로필링
S67

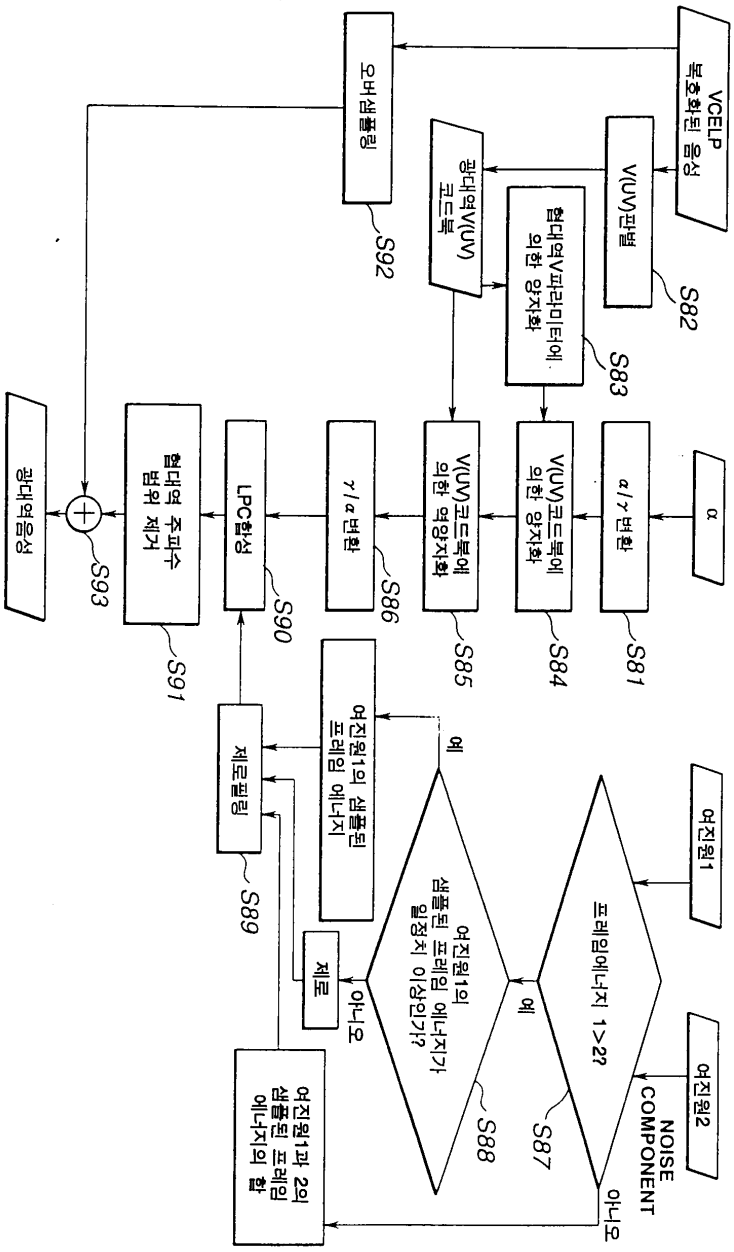
도면12


39
음성 부호화장치 (PSI-SELP)
38
V/UV판별장치
5
α/γ 변환장치
4
제로필링회로
16
6
26
9
7
25
연산회로
협대역 UV 양자화장치
협대역 V 양자화장치
연산회로
오버샘플링회로
19
광대역 UV 코드북
광대역 UV 역양자화장치
광대역 V 역양자화장치
광대역 V 코드북
14
13
11
12
γ/α 변환장치
15
BSF
18
LPC합성장치
17
20
21

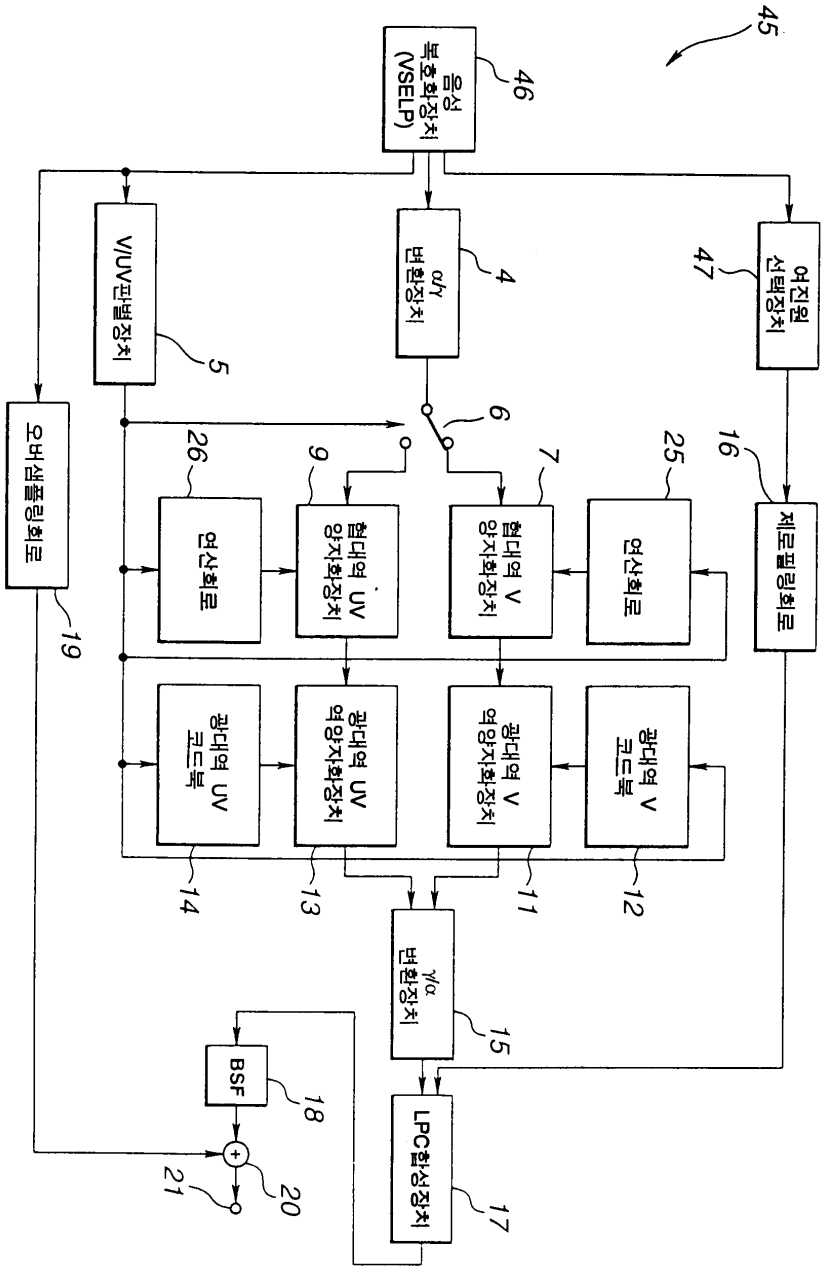
도면13


45
음성 복호화장치 (VSELP)
46
V/UV판별장치
5
αlγ 변환장치
4
여진원 선택장치
47
오버샘플링회로
19
6
29
9
7
28
16
부분추출회로
협대역 UV 양자화장치
협대역 V 양자화장치
부분추출회로
제로필링회로
광대역 UV 코드북
광대역 UV 역양자화장치
광대역 V 역양자화장치
광대역 V 코드북
14
13
11
12
γ/α 변환장치
15
BSF
18
LPC합성장치
17
20
21

도면14


VCELP 복호화된 음성
V(UV)판별
S82
협대역V파라미터에 의한 양자화
S83
α/γ 변환
S81
α
V(UV)코드북에 의한 역양자화
S84
V(UV)코드북에 의한 양자화
S85
γ/α 변환
S86
LPC합성
S90
협대역 주파수 범위 제거
S91
S93
광대역음성
광대역V(UV) 코드북
오버샘플링
S92
여진원1
여진원2
NOISE COMPONENT
프레임에너지 1>2?
S87
예
아니오
여진원1의 샘플된 프레임 에너지가 일정치 이상인가?
S88
예
아니오
여진원1의 샘플된 프레임 에너지
제로
제로필링
S89
여진원1과 2의 샘플된 프레임 에너지의 합

도면15


45
음성 복호화장치 (VSELP)
46
V/UV판별장치
5
α/γ 변환장치
4
여진원 선택장치
47
오버샘플링회로
19
6
26
9
7
25
16
연산회로
협대역 UV 양자화장치
협대역 V 양자화장치
연산회로
제로필링회로
광대역 UV 코드북
광대역 UV 역양자화장치
광대역 V 역양자화장치
광대역 V 코드북
14
13
11
12
γ/α 변환장치
15
BSF
18
21
20
LPC합성장치
17