(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl.
G10L 15/00 (2006.01)

(45) 공고일자 2006년03월23일
(11) 등록번호 10-0538985
(24) 등록일자 2005년12월20일

(21) 출원번호 10-1997-0048769
(22) 출원일자 1997년09월25일

(65) 공개번호 10-1998-0024971
(43) 공개일자 1998년07월06일

(30) 우선권주장 96-257129 1996년09월27일 일본(JP)

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

(72) 발명자 이이지마 가즈유끼
일본국 도쿄도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

니시구찌 마사유끼
일본국 도쿄도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

마쯔모또 준
일본국 도쿄도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

(74) 대리인 신관호

심사관 : 안준형

(54) 음성부호화방법및장치와피치검출방법및장치

요약

본 발명은 반 피치 또는 두배 피치가 검출될 피치보다 더 강한 자기상관을 나타내는 음성신호에 대해서 조차 고정밀 피치 검출을 실행할 수 있는 피치검출 방법 및 장치에 관한 것이다. 입력 음성신호는 유성음 또는 무성음에 대해 판단되고 입력 음성신호의 유성음 부분 및 무성음 부분은 사인파 분석 부호화부(114)와 부호 여기 부호화부(120)에 의해 각각 부호화되어 상기 각 부호화 출력을 생성한다. 사인파 분석 부호화부(114)는 상기 부호화 출력에 피치탐색을 행하여 입력 음성신호로부터 피치정보를 찾고 이에 의해 검출된 피치정보에 따라 고신뢰성 피치정보를 설정한다. 피치검출의 결과는 소정의 고신뢰성 피치정보와 현재 부호화 단위과는 다른 부호화 단위의 유성음/무성음 판정의 결과를 사용하여 판단된다. 상기 피치검출 방법을 사용하는 음성 부호화 방법 및 음성 부호화 장치가 또한 나타나 있다.

대표도

도 1

명세서

도면의 간단한 설명

도 1은 본 발명에 의한 음성 부호화 방법을 행하는 음성 부호화 장치의 기본 구조를 나타내는 블록도이다.

도 2는 본 발명에 의한 음성 복호화 방법을 행하는 음성 복호화 장치의 기본 구조를 나타내는 블록도이다.

도 3은 본 발명을 구체화하는 음성 부호화 장치의 더 상세한 구조를 나타내는 블록도이다.

도 4는 고신뢰성 피치정보를 설정하기 위한 작동의 순서를 나타내는 순서도이다.

도 5는 고신뢰성 피치정보를 재설정하기 위한 작동의 순서를 나타내는 순서도이다.

도 6은 다양한 비트율의 데이터를 나타내는 표이다.

도 7은 도 3의 구조에서 피치검출을 위한 작동의 전형적인 순서를 나타내는 순서도이다.

도 8은 도 3의 구조에서 피치검출을 위한 작동의 전형적인 순서를 나타내는 순서도이다.

도 9는 도 3의 구조에서 피치검출을 위한 작동의 또다른 전형적인 순서를 나타내는 순서도이다.

도 10a, 10b, 10c는 도 3의 구조에서의 피치검출의 결과를 나타낸다.

도 11은 본 발명을 구체화하는 음성신호 부호화 장치를 사용하는 휴대용 단말기의 송신측의 구조를 나타내는 블록도이다.

도 12는 본 발명을 구체화하는 음성신호 복호화 장치를 사용하는 휴대용 단말기의 수신측의 구조를 나타내는 블록도이다.

* 도면의 주요부분에 대한 부호설명

110. 제 1부호화부 111. LPC 역필터

113. LPC 분석/양자화부 114. 사인파 분석 부호화부

115. V/UV판정부 120. 제 2 부호화부

121. 잡음 코드북 122. 가중 합성필터

123. 감산기 124. 거리계산회로

125. 청각 가중 필터

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 입력 음성신호가 부호화 단위로서 소정의 블록에 의해 시간축상에 구분되고 하나의 상기한 부호화 단위로부터 또다른 부호화 단위로 부호화되는 음성 부호화 방법 및 장치에 관한 것이다. 본 발명은 또한 상기 음성 부호화 방법 및 장치를 사용하는 피치검출 방법에 관한 것이다.

지금까지는, 음성 및 음향 신호를 포함한 오디오신호의 시간 영역 및 주파수 영역에서의 통계적 특성과 인간의 청감상 특성을 활용하므로써 신호 압축을 수행하기 위한 다양한 부호화 방법이 알려져 있다. 이 부호화 방법은 대략 시간 영역에서의 부호화와 주파수 영역에서의 부호화와 분석합성 부호화로 구분할 수 있다.

음성 신호의 고효율 부호화의 기술 중에서, 하모닉 부호화나 다중 대역 여기(MBE)부호화와 같은 사인파 분석부호화, 대역분할 부호화(SBC), 선형예측 부호화(LPC), 이산 코사인 변환(DCT), 변경 DCT(MDCT), 고속 푸리에 변환(FFT)이 알려져 있다.

한편, 파라미터로서 입력 음성신호의 피치를 사용하는 사인파 합성 부호화 생성 여기신호에서, 피치검출은 중요한 역할을 한다. 종래의 음성신호 부호화 회로에서 사용된 자기상관법을 사용하고, 피치검출 정확성을 향상시키기 위해 단지 1샘플의 샘플 이동으로 부분탐색하는데 알맞은 피치검출방법에 있어서, 만약 반 피치나 두배 피치가 음성신호에서 검출하기를 바라는 피치보다 더 강한 상관관계를 보인다면, 이 반 피치나 두배 피치는 에러에 의해 검출되기 쉽다. 또한, 음성신호의 무성음 부분에서는 중요한 피치가 없으므로, 무성음 부분의 피치검출의 결과는 종종 피치검출이 되지 않는다.

발명이 이루고자 하는 기술적 과제

따라서 본 발명의 목적은 반 피치나 두배 피치가 음성신호에서 검출하기를 원하는 피치보다 더 강한 상관관계를 나타내는 음성신호에 대한 피치를 올바르게 검출할 수 있는 피치검출 방법을 제공하는 것이다.

본 발명의 또다른 목적은 유성음/무성음의 검출의 결과가 피치검출에서 파라미터로서 사용되고 피치검출의 결과가 올바른 피치검출을 보증하기 위해 다른 음성신호에 대한 피치검출에 영향을 끼치는 것이 방지되는 피치검출 방법을 제공하는 것이다.

본 발명의 또다른 목적은 상기의 피치검출방법을 적용하므로써 이상 잡음이 없는 고정밀 자연재생음성을 생성할 수 있는 음성신호 부호화 방법 및 장치를 제공하는 것이다.

발명의 구성 및 작용

한 관점에서, 본 발명은 입력 음성신호가 소정의 부호화 단위에 의해 시간축상에서 구분되고 이 부호화 단위에 의거하여 음성신호가 유성음/무성음에 대해 판정되는 부호화 방법에서의 피치검출방법을 제공하고, 여기에서, 소정의 피치검출조건하에서 피치정보를 검출하는 피치 탐색단계에서, 음성신호의 현재 부호화 단위의 피치는, 또한 파라미터로서 시간축상의 현재 부호화 단위와 다른 음성신호의 부호화 단위에 의한 음성신호의 유성음/무성음의 판정의 결과를 사용하여 판정된다.

또다른 관점에서, 본 발명은 입력음성신호가 시간축상에서 부호화 단위에 의해 구분되고, 입력 음성신호의 피치를 검출하는 단계와, 입력 음성신호의 단기간 예측 잔차(殘差)를 구하는 예측 부호화 단계와, 상기와 같이 구해진 단기 예측 잔차에 사인파 분석 부호화를 행하는 사인파 분석 부호화 단계와, 입력 음성신호를 파형 부호화하는 파형 부호화 단계와, 상기 부호화 단위 기준으로 입력 음성신호의 유성음/무성음을 판정하는 판정단계를 포함하는 부호화 단위 기준으로 부호화되는 음성신호 부호화 방법을 제공한다. 현재 부호화 단위의 음성신호의 피치는 또한 시간축상의 현재 부호화 단위와 다른 부호화 단위의 음성신호의 판정 결과를 사용하여 판정된다.

본 발명의 음성신호 부호화 방법 및 장치에 있어서, 반 또는 두배 피치는 실수에 의해 검출되지 않아서 고정밀 피치검출을 가능하게 한다. 또한, p, k, t와 같은 파열음이나 마찰음은 완벽하게 재생될 수 있고, 유성음(V) 및 무성음(U) 부분사이의 천이 부분에서 재생된 이상음은 없어서, 윙윙거리는 소리가 없는 명확한 음을 얻을 수 있다.

본 발명에 의한 피치검출방법에 있어서, 시간축 상의 현재 부호화 단위와 다른 부호화 단위의 음성신호에 대한 유성음/무성음 판정의 결과가 또한 현재 부호화 단위의 음성신호의 피치를 판정하기 위한 소정의 피치검출 조건하에서 피치정보를 검출하는 피치탐색단계에서 파라미터로서 사용되고, 고정밀 피치검출이 입력 음성신호에서 반 피치나 두배 피치의 오검출 없이 수행될 수 있다.

또한, 본 발명의 음성 부호화 방법 및 장치에서, 입력 음성신호의 유성음/무성음의 판정의 결과가 사용되어 사인파 분석 부호화 및 파형 부호화가 각각 입력 음성신호의 유성음 및 무성음 부분에 사용될 수 있다. 또한, 상기 부호화 방법 및 장치는

본 발명의 피치검출 방법을 이용하므로, 반 피치 또는 두배 피치의 잘못된 검출없이 고정밀 부호화가 수행될 수 있어서, 무성음 부분에서 윙윙거리는 소리가 없는 명확한 재생음을 얻을 수 있는 한편, 자연 합성음은 유성음 부분에서 얻어질 수 있다. 또한, 무성음 부분과 유성음 부분 사이의 천이부분에는 어떤 이상음도 생성되지 않는다.

도면을 참고하여, 본 발명의 바람직한 실시예가 상세하게 설명될 것이다.

도 1은 본 발명을 구체화하는 피치검출 방법 및 음성신호 부호화 방법을 이행하는 부호화 장치의 기본 구조를 나타낸다.

도 1의 음성 신호 부호기의 기초가 되는 기본 개념은, 이 부호기가, 일예로 하모닉 코딩(Coding)과 같은 사인파 분석 부호화를 행하기 위해, 입력 음성신호의, 예로 선형 예측 부호화(LPC) 잔차(Residuals)와 같은 단기 예측 잔차를 찾기 위한 제 1부호화부(110)와, 입력 음성신호를 위상재현성을 갖는 파형 부호화에 의해 부호화하기 위한 제 2부호화부(120)를 갖는다는 것이고, 상기 제 1부호화부(110)와 제 2부호화부(120)는 각각 입력 신호의 유성음(V)을 부호화하고 입력 신호의 무성음(UV) 부분을 부호화하는데 사용된다는 것이다.

제 1부호화부(110)는 일예로 하모닉 부호화나 다중대역여기(MBE) 부호화와 같은 사인파분석 부호화로 예를들어 LPC잔차를 부호화하는 구성을 사용한다. 제 2부호화부(120)는 최적 벡터의 폐쇄루프탐색에 의한 벡터양자화를 폐쇄루프에 의해 사용하고 또한 예를들어 합성에 의한 분석방법을 사용하여 부호 여기 선형예측(CELP)을 행하는 구성을 사용한다.

도 1에 도시된 실시예에서, 입력단자(101)로 공급된 음성신호는 제 1부호화부(110)의 LPC 역필터(111)와 LPC 분석 및 양자화부(113)로 보내진다. LPC 분석 양자화부(113)에 의해 얻어진 LPC계수 또는 일명 α파라미터는 제 1부호화부(110)의 LPC 역필터(111)로 보내진다. LPC 역필터(111)로부터 입력 음성신호의 선형 예측 잔차(LPC 잔차)를 출력한다. LPC 분석 양자화부(113)로부터 선형 스펙트럼쌍(LSPs)의 양자화 출력이 출력되어 후술될 출력 단자(102)로 보내진다. 사인파 분석 부호화부(114)는 V/UV 판정부(115)에 의한 V/UV판정과 함께 피치검출과, 스펙트럼 엔벌로프의 진폭의 계산을 행한다. 사인파 분석 부호화부(114)로부터의 스펙트럼 엔벌로프 진폭 데이터는 벡터양자화부(116)로 보내진다. 스펙트럼 엔벌로프의 벡터양자화 출력으로서, 벡터양자화부(116)로부터의 코드북 인덱스는 스위치(117)를 경유하여 출력단자(103)로 보내지는 반면, 사인파분석 부호화부(114)의 출력은 스위치(118)를 통해 출력단자(104)로 보내진다. V/UV판정부(115)의 V/UV판정출력은 출력단자(105)로 보내지고, 제어신호로서 스위치(117, 118)에 보내진다. 만약 입력 음성신호가 유성음(V)이면, 인덱스와 피치가 선택되어 출력단자(103, 104)에 출력된다.

본 실시예에서, 도 1의 제 2부호화부(120)는 부호 여기 선형 예측 부호화(CELP부호화)구성을 갖고, 합성에 의한 분석방법을 이용하는 폐쇄루프탐색을 사용하여 시간영역 파형을 벡터양자화하는데, 여기에서 잡음 코드북(121)의 출력은 가중 합성필터(122)에 의해 합성되고, 이 결과의 가중 음성은 감산기(123)로 보내지고, 이 가중 음성과 입력단자(101)에 그리고 이로부터 청각 가중 필터(125)를 통해 공급된 오차가 출력되고, 이와같이 구해진 오차는 거리계산회로(124)로 보내져서 거리계산을 이행하고 상기 오차를 최소화하는 벡터가 잡음 코드북(121)에 의해 탐색된다. 이 CELP부호화는 상기했듯이, 무성음부분을 부호화 하는데 사용된다. 잡음 코드북(121)으로부터의 UV데이터로서 코드북 인덱스는 V/UV판정부(115)로부터의 V/UV 판정의 결과가 무성음(UV)일 때 켜지는 스위치(127)를 통해 출력단자(107)에서 출력된다.

도 2는, 도 1의 음성 신호 부호기에 대응하는 장치로서, 본 발명에 의한 음성 복호화 방법을 수행하기 위한 음성신호 복호화 장치의 기본 구조를 나타내는 블록도이다.

도 2를 참고하면, 도 1의 출력단자(102)로부터의 선형 스펙트럼쌍(LSPs)의 양자화 출력으로서 코드북 인덱스는 입력단자(202)로 공급된다. 입력단자(203, 204, 205)에 도 1의 출력단자(103, 104, 105)의 출력, 즉 엔벌로프 양자화와 피치 및 V/UV판정결과의 출력으로서의 인덱스가 각각 입력된다. 입력단자(207)에 출력단자(107)로부터의 무성음(UV)에 대한 데이터로서의 인덱스가 입력된다.

입력단자(203)의 엔벌로프 양자화 출력으로서 인덱스는 역벡터양자화를 위해 역벡터양자화부(212)로 보내지고 LPC잔차의 스펙트럼 엔벌로프를 찾아서 이를 유성음 합성기(211)로 보낸다. 유성음 합성기(211)는 사인파 합성에 의해 유성음 부분의 선형 예측 부호화(LPC)잔차를 합성한다. 유성음 합성기(211)에는 또한 입력단자(204, 205)로부터의 피치 및 V/UV판정 출력이 공급된다. 유성음 합성부(211)로부터의 유성음의 LPC잔차는 LPC 합성필터(214)로 보내진다. 입력단자(207)로부터의 UV데이터의 인덱스 데이터는 무성음 합성부(220)로 보내지고 여기에서 무성음 부분의 LPC잔차를 꺼내기 위해 잡음 코드북이 참고가 되어야 한다. 이 LPC잔차는 또한 LPC 합성필터(214)로 보내진다. LPC 합성필터(214)에서, 유성음 부분의 LPC 잔차와 무성음 부분의 LPC 잔차가 LPC 합성에 의해 독립적으로 처리된다. 대신에, 서로 합해진 유성

음 부분의 LPC 잔차와 무성음 부분의 LPC 잔차가 LPC 합성처리될 수도 있다. 입력단자(202)로부터의 LSP 인덱스 데이터는 LPC파라미터 재생부(213)로 보내지고 여기에서 LPC의 *a*파라미터는 꺼내져서 LPC 합성필터(214)로 보내진다. LPC 합성필터(214)에 의해 합성된 음성신호는 출력 단자(201)에 꺼내진다.

도 3을 참고로 도 1에 도시된 음성 신호 부호기의 더 상세한 구조를 이제 설명한다. 도 3에서, 도 1과 비슷한 부분 또는 구성요소는 동일한 참고번호로 표시한다.

도 3에 도시된 음성 신호 부호기에서, 입력단자(101)에 공급된 음성 신호는 고역 통과 필터(HPF)(109)에 의해 필터되어 불필요한 범위의 신호가 제거되고 이로부터 LPC 분석/양자화부(113)의 LPC(선형 예측 부호화) 분석회로(132)와 LPC 역필터(111)에 공급된다.

LPC 분석/양자화부(113)의 LPC 분석 회로(132)는 입력 신호 파형의 256샘플 정도의 길이를 1블록으로서, 해밍창(Hamming Window)을 적용하고, 일명 *a*파라미터인 선형 예측 계수를 자기 상관법에 의해 구한다. 1데이터 출력부로서 프레임 간격은 대략 160샘플로 설정된다. 일예로, 만약 샘플링 주파수(fs)가 8kHz이면 1프레임 간격은 20msec 또는 160샘플이다.

LPC 분석 회로(132)로부터의 *a*파라미터는 *a*-LSP 변환회로(133)로 보내져서 선형 스펙트럼쌍(LSP) 파라미터로 변환된다. 이것은 직접형 필터계수로서 구해진 *a*파라미터를, 일예로 LSP 파라미터의 5쌍인 10으로 변환한다. 이 변환은, 일예로 뉴튼-랩슨법에 의해 수행된다. *a*파라미터가 LSP 파라미터로 변환되는 이유는 LSP 파라미터가 보간 특성에서 *a*파라미터보다 우수하기 때문이다.

a-LSP 변환회로(133)로부터의 LSP파라미터는 LSP 양자화기(134)에 의한 매트릭스-또는 벡터이다. 벡터양자화 이전의 프레임간 차이를 취하거나 여기에 매트릭스 양자화를 수행하기 위해 복수의 프레임을 모을 수 있다. 이 경우에, 20msec마다 계산된, 각각 20msec 길이의, LSP 파라미터의 2프레임이 함께 취급되어 매트릭스 양자화 및 벡터양자화 처리된다.

LSP 양자화의 인덱스 데이터인 양자화기(134)의 양자화 출력은 단자(102)에서 출력되는 반면, 양자화 LSP 벡터는 LSP 보간 회로(136)로 보내진다.

LSP 보간회로(136)는 20msec 또는 40msec 마다 양자화된 LSP 벡터를 보간한다. 즉, LSP 벡터는 2.5msec 마다 갱신된다. 이것은, 만약 잔차 파형이 하모닉 부호화/복호화 방법에 의해 분석/합성 처리되면, 합성 파형의 엔벌로프가 매우 매끄러운 파형을 나타내므로, LPC 계수가 20msec 마다 갑작스럽게 변화하면 이상음이 발생될 것 같다. 다시말해, 만약 LPC 계수가 2.5msec마다 서서히 변하면, 상기 이상잡음의 발생이 방지될 수 있다.

2.5msec 마다 생성된 보간 LSP 벡터를 사용하여 입력 음성을 역 필터링하기 위해, LSP 파라미터는 LSP→*a* 변환회로(137)에 의해 *a*파라미터로 변환된다(이것은 일예로 10차정도의 직접형 필터의 필터계수이다). LSP→*a*변환회로(137)의 출력은 LPC 변환 필터회로(111)로 보내져서, 역 필터링을 실행하여 2.5msec 마다 갱신된 *a*파라미터를 사용하여 매끄러운 출력을 생성한다. LPC 역필터(111)의 출력은, 일예로 하모닉 부호화 회로와 같은 사인파 분석 부호화부(114)의, 일예로 DCT 회로와 같은 직교 변환회로(145)로 보내진다.

LPC 분석/양자화부(113)의 LPC분석회로(132)로부터의 *a*파라미터는 청각가중 계산 회로(139)로 보내져서 청각 가중을 위한 데이터가 구해진다. 이 가중 데이터는 청각 가중 벡터양자화기(116)와 제 2부호화부(120)의 청각 가중 필터(125)와 청각 가중 합성필터(122)로 보내진다.

하모닉 부호화 회로의 사인파 분석 부호화부(114)는 하모닉 부호화 방법으로 LPC 역필터(111)의 출력을 분석한다. 다시말해, 피치검출, 각 하모닉스의 진폭(Am)산출 및 유성음(V)/무성음(UV) 판별이 행해지고, 피치로 변화된 각 하모닉스의 엔벌로프 또는 진폭(Am)의 개수는 차원변환에 의해 일정하게 된다.

도 3에 도시된 사인파 분석 부호화부(114)의 구체적인 예에서는, 일반적인 하모닉 부호화가 사용된다. 특히, 다중대역여기(MBE) 부호화에서, 모델화할 때 동시각(동일한 블록 또는 프레임)의 주파수영역 즉 대역마다 유성음 부분과 무성음 부분이 존재한다고 가정한다. 다른 하모닉 부호화 기술에서는, 1블록 또는 1프레임의 음성이 유성음인지 무성음인지가 택일적으로 판단된다. 다음 설명에서, MBE 부호화가 관련된 한에 있어서는, 대역 전체가 UV이면 소정의 프레임이 UV인 것으로 판정된다. 상기한 MBE에 대한 분석합성방법기술의 구체적인 예는 본 출원인의 이름으로 제출된 일본 특허출원번호 4-91442에 나와 있다.

도 3의 사인파 분석 부호화부(114)의 개방루프 피치 탐색부(141)와 영교차 카운터(142)에는 입력단자(101)로부터의 입력 음성신호와 고역 통과 필터(HPF)(109)로부터의 신호가 각각 공급된다. 사인파 분석 부호화부(114)의 직교 변환회로(145)에는 LPC 역필터(111)로부터의 LPC 잔차 즉 선형 예측 잔차가 공급된다.

개방루프 피치 탐색부(141)는 1.0단계 개방루프 피치 탐색을 행하기 위해 입력신호의 LPC잔차를 취한다. 추출된 개략 피치데이터는 다음에 설명될 폐쇄루프 탐색에 의해 고정밀 피치 탐색부(146)로 보내진다. 개방루프 피치 탐색부(141)는 다음에 설명될 폐쇄루프에 의해 0.25단계 고정밀 피치 탐색을 행한다.

개방루프 피치 탐색부(141)는 상기 추출된 개략 피치정보에 따라 고신뢰성 피치정보를 설정한다. 먼저 고신뢰성 피치정보의 후보값은 상기 개략 피치정보에 대한 조건보다 더 엄격한 조건하에서 설정되고 상기 개략 피치정보와 비교되어 무관한 후보값을 갱신하거나 버린다. 상기 고신뢰성 피치정보를 설정하거나 갱신하는 것은 다음에 설명된다.

개방루프 피치 탐색부(141)로부터, LPC잔차의 자기상관 피크값의 최대값을 정규화하여 얻어진 최대 정규화 자기상관값(r'(l))은 상기한 개략 피치정보와 고정밀 피치정보에 따라 출력되어서 V/UV판정부(114)로 보내진다.

다음에 설명될 V/UV판정부(115)의 판정출력은 또한 상기한 개방루프 탐색에 대한 파라미터로서 사용될 수도 있다. 유성음(V)이라고 판정된 음성신호의 부분으로부터 추출된 피치정보만이 상기한 개방루프 탐색에 대해 사용된다.

직교 변환회로(145)는, 일예로 이산(離散) 푸리에 변환(DFT)과 같은 직교변환을 행하여 타임축의 LPC잔차를 추파수축의 스펙트럼 진폭 데이터로 변환한다. 직교 변환회로(145)의 출력은 고정밀 피치 탐색부(146)와 스펙트럼 진폭 또는 엔벌로프를 평가하기 위해 구성된 스펙트럼 평가부(148)로 보내진다.

고정밀 피치 탐색부(146)에는 개방 루프 피치 탐색부(141)에 의해 추출된 비교적 대략적인 피치 데이터와 직교 변환부(145)에 의한 DFT에 의해 얻어진 주파수 영역 데이터가 공급된다. 고정밀 피치 탐색부(146)는 대략적인 피치 데이터값의 중심에 있는 0.2 내지 0.5의 비율로, ±몇몇 샘플로 피치데이터를 진동시켜서, 최적 소수점(플로우팅 포인트;Floating point)을 갖는 고정밀 피치 데이터의 값까지 도달하도록 한다. 합성에 의한 분석 방법은 피치를 선택하기 위한 고정밀 탐색 기술로서 사용되어서 파워 스펙트럼이 원음의 파워 스펙트럼에 가장 근접하게 될 것이다. 폐쇄루프 고정밀 피치 탐색부(146)로부터의 피치 데이터는 스위치(118)를 경유하여 출력단자(104)로 보내진다.

스펙트럼 평가부(148)에서, 각 하모닉스와 하모닉스의 집합으로서의 스펙트럼 엔벌로프의 진폭은 LPC잔차의 직교 변환 출력으로서의 스펙트럼 진폭과 피치를 기초로 평가되어, 고정밀 피치 탐색부(146)와 V/UV판정부(115)와 청각 가중 벡터 양자화부(116)로 보내진다.

V/UV판정부(115)는, 직교 변환회로(145)의 출력, 고정밀 피치 탐색부(146)로부터의 최적 피치, 스펙트럼 평가부(148)로부터의 스펙트럼 진폭 데이터, 개방루프 피치 탐색부(141)로부터의 정규화 자기 상관의 최대값(r'(l)) 및 영교차 카운터(142)로부터의 영교차 카운터값에 따라 프레임의 V/UV를 판정한다. 또한, MBE에 있어서 대역에 따른 V/UV판정의 경계위치는 V/UV 판정에 대한 조건으로 또한 사용될 수 있다. V/UV 판정부(115)의 V/UV 판정출력은 출력단자(105)에서 출력된다.

스펙트럼 평가부(148)의 출력부 또는 벡터양자화부(116)의 입력부에는 데이터수 변환부(일종의 샘플링율 변환을 수행하는 부)가 장치된다. 데이터수 변환부는 주파수축에 구분된 대역의 수와 데이터의 수가 피치와 다름을 고려하여 엔벌로프의 진폭 데이터(| Am |)를 일정한 값으로 설정하는데 사용된다. 다시말해, 만약 유효대역이 3400kHz까지이면, 유효대역은 피치에 따라 8 내지 63대역으로 구분될 수 있다. 각 대역 마다 얻어진 진폭 데이터 (| Am |)의 개수 $m_{MX}+1$은 8 내지 63의 범위에서 변화된다. 따라서 데이터수 변환부(119)는 가변개수 ($m_{MX}+1$)개의 진폭 데이터를, 일예로 44데이터와 같은 데이터의 일정개수 (M)개로 변환한다.

스펙트럼 평가부(148)의 출력부 즉 벡터양자화부(116)의 입력부에 제공된, 데이터 개수 변환부로부터의 일예로 44와 같은 일정개수 (M)개의 진폭데이터나 엔벌로프 데이터는 데이터 양자화부(116)에 의해, 한 단위로서 일예로 44데이터와 같은 일정개수의 데이터에 의하여 함께 모아져서 벡터양자화부(116)에 의해 가중 벡터양자화된다. 이 중량은 청각 가중 필터 연산부(139)의 출력에 의해 공급된다. 벡터양자화기(116)로부터의 엔벌로프의 인덱스는 스위치(117)에 의해 출력단자(103)에서 출력된다. 가중 벡터양자화에 앞서, 일정개수의 데이터로 이루어진 벡터에 대해 적당한 리키지(Leakage) 계수를 사용하여 부호화 단위간 차이를 취하는 것이 바람직하다.

제 2부호화부(120)를 설명한다. 제 2부호화부(120)는 부호여기선형예측(CELP) 부호화 구조이고 특히, 입력 음성신호의 무성음부를 부호화하기 위해 사용된다. 무성음 부분을 위한 CELP 부호화 구조에서, 일명 스토캐스틱 코드북(121)인 잡음 코드북의 대표출력으로서 무성음 부분의 LPC잔차에 해당하는 잡음출력은 이득회로(126)를 통해 청각 가중 합성필터(122)로 보내진다. 청각 가중 합성필터(122)는 입력 잡음을 LPC합성하고 이 결과의 가중 무성음 신호를 감산기(123)로 보낸다. 고역 통과 필터(HPF)(109)를 통해 입력단자(101)로부터 보내지고 청각 가중 필터(125)에 의해 청각 가중된 음성신호는 감산기(123)로 공급되고 여기서 청각 가중 음성신호와 합성필터(122)로부터의 신호와의 차이 또는 오차가 구해진다. 반면, 청각 가중 합성필터의 제로 입력응답은 미리 청각 가중 필터(125)의 출력으로부터 감산된다. 이 오차는 거리계산회로(124)로 공급되어 거리와 상기 오차를 최소화할 대표값 벡터가 잡음 코드북(121)에 의해 탐색된다. 상기한 것은 폐쇄루프 탐색과 그 다음으로 합성에 의한 분석법을 사용하는 시간영역 파형의 벡터양자화의 요약이다.

CELP 부호화 구조를 사용하는 제 2부호기(120)로부터 무성음(UV)부분을 위한 데이터로서 잡음 코드북(121)으로부터의 코드북의 형상(Shape) 인덱스와 이득 회로(126)로부터의 코드북의 이득 인덱스가 출력된다. 잡음 코드북(121)으로부터의 UV데이터인 형상(Shape) 인덱스는 스위치(127s)를 통해 출력 단자(107s)로 보내지는 반면, 이득 회로(126)의 UV데이터인 이득 인덱스는 스위치(127g)를 통해 출력 단자(107g)로 보내진다.

이 스위치(127s, 127g)와 스위치(117, 118)는 V/UV판정부(115)로부터의 V/UV판정결과에 따라 온/오프 된다. 구체적으로, 만약 현재 송신된 프레임의 음성신호의 V/UV판정의 결과가 유성음(V)을 나타내면, 스위치(117, 118)가 온되는 반면, 만약 현재 송신된 프레임의 음성신호가 무성음(UV)이면, 스위치(127s, 127g)가 온된다.

상기한 고신뢰성 피치정보를 설명한다.

고신뢰성 피치정보는 두배 또는 반 피치의 오검출을 방지하기 위한 종래의 피치정보에 부가하여 사용되는 평가 파라미터이다. 도 3에 도시된 음성신호 부호화 장치에 있어서, 고신뢰성 피치정보는, 고신뢰성 피치정보의 후보값으로서, 입력단자(101)로 입력된 입력 음성신호 피치정보와 음성레벨(부호화 단위 레벨)과 자기상관 피크값에 의거하여 사인파 분석 부호화부(114)의 개방루프 피치 탐색부(141)에 의해 설정된다. 상기와 같이 설정된 고신뢰성 피치정보의 후보값은 다음 프레임의 개방루프 탐색의 결과와 비교되고, 만약 두 피치값이 서로 충분히 가까우면 고신뢰성 피치정보로서 등록된다. 만약 아니라면, 후보값은 버려진다. 상기 등록된 고신뢰성 피치정보는 또한 만약 소정의 시간동안 갱신되지 않은 채로 있으면 버려진다.

상기한 고신뢰성 피치정보를 설정하고 재설정하기 위한 작동의 설명적인 순서의 알고리즘이 부호화 단위로서의 프레임과 함께 이제 설명된다.

다음은 이하의 설명에 사용되는 변수의 정의이다.

rblPch 고신뢰성 피치정보

rblPchCd 고신뢰성 피치정보의 후보값

rblPchHoldState 고신뢰성 피치정보 유지시간

lev 음성레벨(부호화 단위 레벨)(rms)

Ambiguous(p0, p1, range)는 만약 다음 네 조건 중 어느 것이라도 만족되면, 즉 만약 두 피치값(p0)와 (p1)이 서로에 대해 두배, 세배, 1/2배, 1/3배이면 참이되는 함수이다.

abs(p0-2.0×p1)/p0＜range

abs(p0-3.0×p1)/p0＜range

abs(p0-p1/2.0)/p0＜range

abs(p0-p1/3.0)/p0＜range

상기 부등식에서, range는 소정의 상수이다. 반면, 다음을 가정한다.

pitch[0] : 직전 프레임의 피치

pitch[1] : 현재 프레임의 피치

pitch[2] : 다음에 올(미래) 프레임의 피치

r'(n) : 자기상관 피크값

lag(n) : 피치래그(많은 샘플에 의해 표시되는 피치주기)

여기에서 r'(n)는 자기상관의 0번째 피크(R_0)(파워)에 의해 정규화되고 증가량의 순서로 배열된 자기상관의 연산값(R_k)을 나타내고, n은 순서를 나타낸다.

또한 자기상관 피크의 값(r'(n))과 피치래그(lag(n))는 또한 현재 프레임 동안 보존된다고 가정한다. 이것은 각각 crntR'(n)와 crntR'(n)으로 표시된다. 또한

rp[0] : 직전(과거) 부호화 단위(r'(l))의 자기상관 피크의 최대값

rp[1] : 현재 부호화 단위(r'[l])의 자기상관 피크의 최대값

rp[2] : 다음에 올(미래) 부호화 단위(r'[l])의 자기상관 피크의 최대값

또한 고신뢰성 피치정보의 후보값이 어떤 소정의 조건을 만족시키는 현재 프레임의피치 또는 자기상관 피크값 또는 피치값에 의해 설정되고 고신뢰성 피치정보는 이 후보값과 다음 프레임의 피치간의 차이가 소정의 값보다 더 작을 때만 등록된다고 가정한다.

다음에, 개략 검출 피치정보에 의거하여 고신뢰성 피치정보를 정하는 설명적인 알고리즘을 나타내었다.

[조건 1]

if rblPch×0.6＜pitch[1]＜rblPch×1.8

and

rp[1]＞0.39

and

lev＞200.0

or

rp[1]＞0.65

or

rp[1]＞0.30 and abs(pitch[1]-rblPchCd)＜8.0 and lev＞400.0

then

[조건 2]

if rblPchCd≠0.0 and abs(pitch[1]-rblPchCd)＜8

and !Ambiguous(rblPch, pitch[1], 0.11)

then

[처리 1]

rblPch = pitch[1]

endif

[처리 2]

rblPchCd = pitch[1]

else

[처리 3]

rblPchCd = 0.0

endif

상기한 알고리즘에 의한 고신뢰성 피치정보를 설정하는 작동의 순서는 도 4의 순서도를 참고하여 설명한다.

만약 스텝(S1)에서, '조건 1'이 만족되면, 처리는 스텝(S2)으로 전달되어 '조건 2'가 만족되는지 여부가 판단된다. 만약 '조건 1이' 스텝(S1)에서 만족되지 않으면, 스텝(S5)에 나타난 '처리 3'이 실행되고 이 실행의 결과가 고신뢰성 피치정보로서 받아들여진다.

만약 '조건 2'가 스텝(S2)에서 만족되면, 스텝(S3)의 '처리 1'이 실행되고, 그리고나서 스텝(S2)의 '처리 2'로 간다. 반면, 만약 '조건 2'가 스텝(S2)에서 만족되지 않으면, 스텝(S3)의 '처리 1'은 실행되지 않고 스텝(S4)의 '처리 2'가 실행된다.

스텝(S4)의 '처리 2'의 실행의 결과는 고신뢰성 피치정보로서 출력된다.

만약 고신뢰성 피치정보의 등록후에, 고신뢰성 피치정보가 예를들어 계속해서 5프레임 동안 새롭게 등록되지 않으면, 상기 등록된 고신뢰성 피치정보가 재설정된다.

한번 설정된 고신뢰성 피치정보를 재설정하는 알고리즘의 일예가 이제 설명된다.

[조건 3]

if rblPchHoldState = 5

then

[처리 4]

rblPch = 0.0

rblPchHoldState = 0

else

[처리 5]

rblPchHoldState++

endif

상기 알고리즘에 의한 고신뢰성 피치정보를 재설정하는 작동의 순서가 도 5의 순서도를 참고하여 설명된다.

만약 스텝(S6)에서, '조건 3'이 만족되면, 스텝(S7)에 나타난 '처리 4'가 실행되어 고신뢰성 피치정보가 재설정된다. 역으로, 만약 '조건 3'이 스텝(S6)에서 만족되지 않으면, 스텝(S7)의 '처리 4'가 실행되지 않고 스텝(S8)에 나타난 '처리 5'가 실행되어 고신뢰성 피치정보가 재설정된다.

상기한 것은 고신뢰성 피치정보를 설정하고 재설정한다.

상기한 음성신호 부호기는 필요한 음질에 따라 다른 비트율의 데이터를 출력할 수 있다. 다시말해, 출력데이터는 다양한 비트율로 출력될 수 있다.

특히, 출력 데이터의 비트율이 저비트율과 고비트율간에 전환될 수 있다. 예를들어, 만약 저비트율이 2kbps이고 고비트율이 6kbps이면, 출력 데이터는 도 6에 나타난 다음 비트율을 갖는 데이터이다.

출력단자(104)로부터의 피치데이터는 항상 유성음에 대해 8bits/20msec의 비트율로 출력되고, 출력단자(105)로부터의 V/UV 판정출력은 항상 1bit/20msec이라고 나타나 있다. 출력단자(102)로부터 출력된 LSP 양자화에 대한 인덱스는 32bits/40msec와 48bits/40msec사이에서 전환된다. 반면, 출력단자(103)에 의해 출력된 15bits/20msec와 87bits/20msec사이에서 전환된다. 출력단자(107s, 107g)로부터 출력된 무성음(UV)에 대한 인덱스는 11bits/10msec와 23bits/5msec사이에서 전환된다. 유성음(UV)에 대한 출력 데이터는 2kbps에 대해서는 40bits/20msec이고 6kbps에 대해서는 120kbps/20msec이다. 한편, 유성음(UV)에 대한 출력 데이터는 2kbps에 대해서는 39bits/20msec이고 6kbps에 대해서는 117kbps/20msec이다. LSP양자화에 대한 인덱스와 유성음(V)에 대한 인덱스와 무성음(UV)에 대한 인덱스가 다양한 구성요소의 구조와 관련하여 다음에 설명될 것이다.

도 3의 음성 부호기에서, 유성음/무성음(V/UV) 판정부(115)의 상세한 예가 이제 설명될 것이다.

이 V/UV 판정부(115)는 입력 음성신호의 프레임 평균 에너지(lev), 정규화된 자기상관 피크값(rp), 스펙트럼 유사도(pos), 제로크로싱(nZero), 피치 래그(pch)에 따라 종류별로 부호화 단위의 V/UV 판정을 수행한다.

즉, V/UV 판정부(115)에는 직교 변환회로(145)의 출력에 의거하여, 입력 음성신호의 스펙트럼 엔벌로프의 프레임 평균 에너지(lev) 다시말해 프레임 평균 rms 또는 그에 준한 값(lev)과, 개방루프 피치 탐색부(141)로부터의 정규화된 자기상관 피크값(rp)과, 제로크로싱 카운터(142)로부터의 크로스 카운트값(nZero)과 제로크로싱 카운터(142)로부터의 최적 피치로서의 피치 래그(pch)가 공급된다. 제로크로싱의 수는 샘플의 수로서 표시되는 피치주기이다. MBE에서와 유사한 대역을 기초로 한 V/UV판정의 대역에 따른 결과의 경계 위치는 또한 프레임의 V/UV판정에 대한 조건으로서 사용된다. 이것은 스펙트럼 유사도(pos)로서 V/UV 판정부(115)로 보내진다.

대역에 의거한 V/UV 판정의 결과를 사용하는 MBE에 대한 V/UV 판정의 조건이 이제 설명된다.

MBE의 경우에 m번째 하모닉스의 크기를 나타내는 파라미터 또는 진폭 | Am | 은 다음에 의해 표현된다.

$$\therefore |A_m| = \sum_{j=a_m}^{b_m} |S(j)| |E(j)| \Big/ \sum_{j=a_m}^{b_m} |E(j)|^2$$

이 식에서 | S(j) | 는 LPC 잔차를 DFT하여 얻어진 스펙트럼이고 | E(j) | 는 기본 신호의 스펙트럼, 구체적으로 256포인트 해밍창이고, a_m, b_m은 인덱스 j에 의해 표시되는, 다음으로 m번째 하모닉스의 해당하는 m번째 대역에 대응하는 주파수의 하한 및 상한값이다. 대역기준 V/UV판정에 대해, 신호 대 잡음비(NSR)가 사용된다. m번째 대역의 NSR은 다음식에 의해 표시된다.

$$NSR = \frac{\sum_{j=a_m}^{b_m} \{ | S(j) | - | A_m | | E(j) | \}^2}{\sum_{j=a_m}^{b_m} | S(j) |^2}$$

만약 NSR값이 소정의 임계값, 일예로 0.3보다 크면, 즉 오차가 보다 더 크면, 적용된 대역에서의 | A_m | | E(j) | 에 의한 | S(j) | 의 근사가 좋지 않다고, 즉 여기신호 | E(j) | 가 베이스로서 적당하지 않다고 판단된다. 따라서 적용된 대역은 무성음이라고 판단되고, 만약 아니라면, 근사는 매우 잘 행해지고 따라서 대역은 유성음(V)이라고 판단된다.

한편, 기본 피치 주파수로 구분된 대역의 수가 음성신호의 피치에 따라 대략 8 내지 63의 범위에서 변한다면, 대역에 의거한 V/UV 플래그의 대역에 따른 개수도 또한 변한다. 따라서 V/UV 판정의 결과는 고정 주파수 범위를 구분하여 얻어진 소정의 수의 대역에 의해 분류(축퇴)된다. 구체적으로, 가청범위를 포함한 소정 주파수 범위는 예를들어 각각 V/UV 판정이 이루어진 12대역으로 구분된다. 구체적으로, 대역 기준 V/UV 판정 데이터의 모든 대역에서 유성음(V)과 무성음(UV)사이의 1분할 위치 또는 경계 위치 정도만을 나타내는 데이터가 스펙트럼 유사도(pos)로서 사용된다. 스펙트럼 유사도(pos)에 의해 가정될 수 있는 값은 1≤pos≤12이다.

V/UV 판정부(115)로 공급된 입력 파라미터는 유성음(V)과 유사도를 나타내는 함수값을 실행하기 위해 함수 연산된다. 이 함수의 구체적인 예가 이제 설명된다.

먼저, 함수(pLev(lev))의 값이 입력 음성신호의 프레임 평균 에너지의 값(lev)에 기초하여 연산된다. 이 함수(pLev(lev))로서,

pLev(lev) = 1.0/(1.0 + exp(-(lev - 400.0)/100.0))이 사용된다.

그리고나서, 함수(pR0r(rp))의 값이 정규화 자기상관 피크값(rp)(0≤rp≤1.0)의 값에 따라 연산된다. 함수(pR0r(rp))의 구체적인 예는

pR0r(rp) = 1.0/(1.0 + exp(-(rp - 0.3)/0.06))이다.

그리고나서 함수(pP0s(pos))의 값은 유사도(pos)(1≤pos≤12)에 따라 연산된다. 함수(pP0s(pos))의 구체적인 예는

pP0s(pos) = 1.0/(1.0 + exp(-(pos - 1.5)/0.8))

그리고나서 함수(pNZero(nZero))의 값은 영교차(nZero)(1≤nZero≤160)의 수의 값에 따라 구해진다. 함수(pNZero(nZero))의 구체적인 예는

pNZero(nZero) = 1.0/(1.0 + exp((nZero - 70.0)/12.0))이다.

그리고나서 함수(pPch(pch))의 값은 피치래그(pch)(20≤pch≤147)에 따라 구해진다. 함수(pPch(pch))의 구체적인 예는

pPch(pch) = 1.0/(1.0 + exp(-(pch - 12.0)/2.5))

×1.0/(1.0 + exp(-(pch - 105.0)/6.0))이다.

상기 함수 pLev(lev), pR0r(rp), pNZero(nZero), pPch(pch)에 의해 연산된 파라미터 lev, rp, pos, nZero, pch에 대해 V와의 유사도를 사용하여, V와의 최종적 유사도가 연산된다. 이 경우에, 다음 두 사항이 고려될 필요가 있다.

다시말해, 첫 번째 사항으로서, 만역 자기상관 피크값이 보다 더 작지만 프레임 평균 에너지가 매우 크다면, 음성은 유성음이라 생각되어야 한다. 즉, 가중된 합은 서로 강한 보완 관계를 갖는 두 파라미터에 대해 취해진다. 두 번째 사항으로서, V와의 유사도를 독립적으로 나타내는 파라미터가 곱에 의해 처리된다.

따라서, 서로 보완관계를 갖는 자기상관 피크값과 부호화 단위 평균 에너지가 가중 부가에 의해 합해지는 반면, 다른 파라미터는 곱에 의해 처리된다. V와의 최종적 유사도를 나타내는 함수(f(lev, rp, pos, nZero, pch))는 다음과 같이 연산된다.

f(lev, rp, pos, nZero, pch)

= ((1.2pR0r(rp) + 0.8pLev(lev))/2.0)

×pPos(pos)×pNZero(nZero)×pPch(pch).

가중 파라미터(α = 1.2, β = 0.8)는 경험적으로 구해진 값으로 알려져 있다.

V/UV판정은 소정의 임계값으로 함수(f)의 값을 판별하므로써 행해진다. 구체적으로 만약 f가 최종적으로 0.5정도이면, 프레임은 유성음(V)인 반면, f가 0.5보다 작으면, 부호화 단위은 무성음(UV)이다.

한편, 정규화된 자기상관 피크값(rp)에 대해 V와의 유사도를 구하기 위한 상기한 함수(pR0r(rp))는 함수(pR0r(rp))를 근사하는 함수들(pR0r'(rp)), 즉

pR0r'(rp) = 0.6x 0≤x≤7/34

pR0r'(rp) = 4.0(x - 0.175) 7/34≤x≤67/170

pR0r'(rp) = 0.6x + 0.64 67/170≤x≤0.6

pR0r'(rp) = 1 0.6≤x≤1.0

에 의해 대체될 수도 있다.

종합적으로, 상기한 V/UV판정의 기본 개념은 V/UV판정에 대한 파라미터(x), 일예로 상기한 입력 파라미터(lev, rp, pos, nZero, pch)가 다음에 의해 표현되는 S자형 함수(g(x))에 의해 변환된다.

g(x) = A/1(1 + exp(-(x-b)/a))

여기서 A, a, b는 상수이고, 이 S자형 함수(g(x))에 의해 변환된 파라미터는 V/UV 판정을 위해 사용된다.

만약 이 입력 파라미터(lev, rp, pos, nZero, pch)가 일반화되어서 n개의 입력 파라미터(여기서, n은 자연수)가 x_1, x_2, …, x_n에 의해 표현되면, 이 입력 파라미터 x_k(여기서, k=1, 2, …, n)에 의한 V와의 유사도는 함수 $g_k(x_k)$에 의해 표현되고 V와의 최종 유사도는

$f(x_1, x_2, \cdots, x_n) = F(g_1(x_1), g_2(x_2), \cdots, g_n(x_n))$

에 의해 평가된다.

함수($g_k(x_k)$)(k = 1, 2, …, n)로서, c_k 내지 d_k(c_k 및 d_k는 상수이고 $c_k < d_k$)의 어떤 값이라도 취할 수 있는 범위를 갖는 임의 함수가 사용될 수 있다. 함수($g_k(x_k)$)로서, c_k 내지 d_k의 어떤 값이라도 취할 수 있는 범위를 갖는, 다른 기울기를 갖는 복수의 직선으로 구성된 임의의 함수가 또한 사용될 수 있다.

함수($g_k(x_k)$)로서, c_k 내지 d_k의 어떤 값이라도 취할 수 있는 임의의 연속 함수가 마찬가지로 사용될 수 있다.

또한, 함수($g_k(x_k)$)로서,

$g_k(x_k) = A_k/(1 + \exp(-(x_k - b_k)/a_k))$

(여기서 k = 1, 2, …, n 및 A_k, a_k, b_k는 입력 파라미터 x_k와는 다른 상수이거나 또는 곱셈에 의한 이의 결합이다)

에 의해 표현되는 S형 함수가 사용될 수 있다.

S형 함수 또는 곱셈에 의한 이의 결합이 다른 기울기를 갖는 복수의 직선에 의해 근사될 수 있다.

입력 파라미터는 입력 음성신호의 상기한 부호화 단위 평균에너지(lev), 정규화 자기상관(rp), 스펙트럼 유사도(pos), 영교차의 수(nZero), 피치 래그(pch)로 열거될 수 있다.

만약 상기 입력 파라미터(lev, rp, pos, nZero, pch)에 대해 V와의 유사도를 나타내는 함수가 각각 pLev(lev), pR0(rp), pPos(pos), pNZero(nZero), pPch(pch)에 의해 표현되면, 이 함수들에 의해 V와의 최종 유사도를 표현하는 함수(f(lev, rp, pos, nZero, pch))는

f(lev, rp, pos, nZero, pch)

= ((αpR0(rp) + βpLev(lev))/(α+ β))

×pPos(pos)×pNZero(nZero)×pPch(pch)

(여기서, α, β는 각각 pR0r 및 pLev를 적당하게 가중하기 위한 상수이다.)

에 의해 계산될 수 있다.

상기한 바와 같이 얻어진 함수(f)의 값은 V/UV 판정을 하기 위한 소정의 임계값을 사용하여 판별된다.

피치검출이 고신뢰성 피치정보를 사용하여 행해지는 방식이 이제 설명된다.

피치검출은 기준값으로서 상기한 연산에 의해 구해진 고신뢰성 피치정보(rblPch)와 함께, 이전 프레임 (prevVUV)의 V/UV판정의 결과를 사용하여 행해진다고 가정한다.

이 경우에, 고신뢰성 피치정보(rblPch)와 이전 프레임(prevVUV)의 V/UV판정의 결과의 조합에 따라 다음의 네 경우((i) 내지 (iv))가 있다.

(i)prevVUV ≠ 0 이고 rblPch ≠ 0

피치검출은 고신뢰성 피치정보를 참고하여 행해진다. 직전 프레임이 이미 유성음(V)으로 판단되었으므로, 직전 부호화 단위의 정보를 우선적으로 피치검출에 관계시킨다.

(ii)prevVUV = 0 이고 rblPch ≠ 0

직전 프레임이 무성음(UV)이므로, 그 피치는 사용될 수 없고, 따라서 피치검출이 rblPch에만 관계하여 행해진다.

(iii)prevVUV = 1 이고 rblPch = 0

적어도 직전 프레임이 유성음(V)으로 판단되므로, 피치검출은 이것의 피치만을 사용하여 행해진다.

(iv)prevVUV = 0 이고 rblPch = 0

직전 프레임이 무성음(UV)으로 판단되므로, 피치검출은 다음에 올 미래 프레임 피치에 관하여 행해진다.

상기한 네 경우가 도 7 및 도 8의 순서도를 참고로 구체적으로 설명될 것이다.

도 7 및 도 8에서, ！는 부정을나타내고, &&는 'and'를 나타내고 trkPch는 최종적으로 검출된 피치인 피치를 나타낸다.

SearchPeaks(frm)(frm = {0, 2})는 만약 rp[1]≥rp[frm]이거나 만약 rp[1]＞0.7이면 pitch[1]인 그리고 그렇지 않다면 crntLag(n)를 n =0, 1, … 에 대해 순서대로 탐색할 때 0.81×pitch[frm]＜crntLag(n)＜1.2×pitch[frm]를 최초로 만족시키는 crntLag(n)를 그 값으로서 갖는 함수이다.

마찬가지로, SearchPeaks3Frms는 만약 비교시에 rp[0], rp[1] 및 rp[2], rp[1]가 rp[0] 또는 rp[2]보다 크거나 0.7보다 크면 pitch[1]와 같고, 그렇지 않으면 참조 프레임으로서 자기상관 피크 rp[0] 또는 rp[2]의 보다 큰 값을 갖는 프레임을 사용하여 상기한 SearchPeaks(frm)과 동일한 작동을 행하는 함수이다.

먼저, 스텝(S10)에서, '직전 프레임(prevVUV)의 V/UV판정의 결과가 0이 아닌 반면 고신뢰성 피치정보(rblPch)는 0.0이 아니다'는 조건이 만족되는지 여부가 판단된다. 만약 이 조건이 만족되지 않으면, 처리는 다음에 설명될 스텝(S29)으로 간다. 만약 조건이 만족되면, 처리는 스텝(S11)으로 간다.

스텝(S11)에서,

status0 = Ambiguous(pitch[0], rblPch, 0.11)

status1 = Ambiguous(pitch[1], rblPch, 0.11)

status2 = Ambiguous(pitch[2], rblPch, 0.11)

이 정의된다.

스텝(S12)에서, 조건'status0과 status1과 status2 중 아무것도 유효하지 않음'을 만족하는지 여부가 판단된다. 만약 이 조건을 만족하면 처리는 스텝(S13)으로 가고 만약 만족하지 않으면 스텝(S18)으로 간다.

스텝(S18)에서, 조건'status0 및 status2가 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면, 처리는 스텝(S19)으로 가서 피치로서 SearchPeaks(0)가 채택되고, 만약 만족되지 않으면, 처리는 스텝(S20)으로 간다.

스텝(S20)에서, 조건'status1 및 status2가 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면, 처리는 스텝(S21)으로 가서 피치로서 SearchPeaks(2)가 채택되고, 만약 만족되지 않으면 처리는 스텝(S22)로 간다.

스텝(S22)에서, 조건'status0가 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면, trkPch = pitch[0]이 피치로서 정해지고, 만약 만족되지 않으면, 처리는 스텝(S24)로 간다.

스텝(S24)에서, 조건'status1이 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면 trkPch = pitch[1]가 피치로서 정해지고 만약 만족되지 않으면 처리는 스텝(S26)으로 간다.

스텝(S26)에서, 'status2가 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면, trkPch = pitch[2]가 피치로서 정해지고, 만약 만족되지 않으면, 처리는 스텝(S28)로 가서 피치로서 trkPch = pitch[0]가 채택된다.

상기 스텝(S13)에서, 함수 Ambiguous(pitch[2], pitch[1], 0.11)가 참인지 거짓인지가 판단된다. 만약 이 함수가 참이면 처리는 스텝(S14)으로 가서 SearchPeaks(0)가 피치로서 채택된다. 만약 함수가 거짓이면, 처리는 스텝(S15)로 가서 SearchPeaks3frms()가 피치로서 채택된다.

스텝(S15)에서, 함수 Ambiguous(pitch[0], pitch[1], 0.11)가 참인지 거짓인지가 판단된다. 만약 이 함수가 참이면, 처리는 스텝(S16)으로 가서 SearchPeaks(2)이 피치로서 채택된다. 만약 이 함수가 거짓이면, 처리는 스텝(S17)로 가서 SearchPeaks3frms()가 피치로서 채택된다.

그리고나서, 상기 스텝(S29)에서, 조건'직전 프레임이 UV이고 고신뢰성 피치정보가 0.0임'이 만족되는지가 판단된다. 만약 이 조건이 만족되지 않으면, 처리는 스텝(S38)으로 가고 만약 만족되면 스텝(S30)으로 간다.

스텝(S30)에서,

status0 = Ambiguous(pitch[0], rblPch, 0.11)

status1 = Ambiguous(pitch[2], rblPch, 0.11)

가 정의된다.

스텝(S31)에서, 조건'status0 및 status1이 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면 처리는 스텝(S32)으로 가서 SearchPeaks(2)가 피치로서 채택되고, 만족되지 않으면 처리는 스텝(S33)으로 간다.

스텝(S33)에서, 조건'status0이 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면, trkPch = pitch[1]가 피치로서 정해지고, 만족되지 않으면 처리는 스텝(S35)으로 간다.

스텝(S35)에서 조건'status1이 유효하지 않음'이 만족되는지가 판단된다. 만약 이 조건이 만족되면 trkPch = pitch[2]가 피치로서 정해지고 만약 만족되지 않으면 처리는 스텝(S37)으로 가서 trkPch = rblPch가 피치로서 채택된다.

상기한 스텝(S38)에서, 조건 '직전 프레임이 UV가 아니고 고신뢰성 피치정보가 0.0임'이 만족되는지 여부가 판단된다. 만약 이 조건이 만족되지 않으면, 처리는 스텝(S40)으로 간다.

스텝(S40)에서, 함수 Ambiguous(pitch[0], pitch[2], 0.11)가 참인지 거짓인지가 판단된다. 만약 이 함수가 거짓이면, 처리는 스텝(S41)으로 가서 SearchPeaks3Frms()가 피치로서 채택된다. 만약 함수가 참이면, 처리는 스텝(S42)으로 가서 SearchPeaks(0)가 피치로서 채택된다.

상기한 작동의 순서로 고신뢰성 피치정보를 사용하는 피치검출이 행해진다.

상기의 구체적인 예에서, 피치검출은 고신뢰성 피치정보와 함께 V/UV검출의 결과를 사용하여 행해진다. 통상의 피치검출을 위해 V/UV검출의 결과만을 사용한 또다른 구체적인 예가 이하 설명된다.

피치검출에 대해 현재 부호화 단위와 다른 부호화 단위의 V/UV검출의 결과를 사용하기 위해, V/UV판정은

정규화 자기상관 피크값 r'(n) ($0 \leq r'(n) \leq 1.0$)

영교차의 수 nZero($0 \leq nZero \leq 160$)

부호화 단위 평균레벨(lev)

의 세 파라미터로부터만 주어진다.

이 세 파라미터에 대해, V와의 유사도가 다음식에 의해 연산된다.

[수학식 1]

pRp(rp) = 1.0/(1.0 + exp(-(rp - 0.3/0.06)))

[수학식 2]

pNZero(nZero) = 1.0/ {exp((nZero - 70.0)/12.0)}

[수학식 3]

pLev(lev) = 1.0/ {1.0 + exp(-(lev - 400.0/100.0))}

수학식 1 내지 수학식 3을 사용하여, V와의 최종 유사도는 다음 식에 의해 정의된다.

[수학식 4]

f(nZero,rp,lev)

= pNZero(nZero)× {1.2×pRp(rp)+ 0.8×pLev(lev)} /2.0

만약 f가 0.5정도이면, 프레임은 유성음(V)이라고 판단되고, 만약 f가 0.5보다 작으면, 프레임은 무성음(UV)이라고 판단된다.

V/UV판정의 결과만을 사용하는 피치검출의 작동의 구체적인 순서가 도 9의 순서도를 참고로 설명된다.

prevVUV는 직전 프레임의 V/UV 판정의 결과라고 나타나 있다. 1 및 0의 prevVUV의 값은 각각 V 및 UV를 나타낸다.

먼저, 스텝(S50)에서, V/UV판정은 현재 프레임에 대해 이루어져 '판정의 결과 (prevVUV)가 1의 값을 가짐'의 여부, 즉 프레임이 유성음인지 여부가 판단된다. 만약 프레임이 스텝(S50)에서 UV임이 판단되면, 처리를 스텝(S51)로 가서 trkPch = 0.0가 피치로서 채택된다. 반면, 만약 스테(S50)의 결과가 V이면, 처리를 스텝(S52)으로 간다.

스텝(S52)에서, '과거 및 미래 프레임의 V/UV판정의 결과가 1임의 여부, 즉 두 부호화 단위 모두가 V인지 여부가 판단된다. 만약 결과가 부정이면, 처리는 이하 설명될 스텝(S53)으로 간다. 만약 두 프레임이 모두 V이면, 처리는 스텝(S54)으로 간다.

스텝(S54)에서, 두 피치 pitch[2], pitch[1]와 상수 0.11간의 관계를 설명하는 함수 Ambiguous(pitch[2], pitch[1], 0.11)가 참인지 거짓인지 여부가 판단된다. 만약 상기 함수가 참이면, 처리는 스텝(S55)으로 가서 trkPch = SearchPeaks(0)를 정한다. 즉 만약 rp[1]≥rp[0] 또는 rp[01]＞0.7이면 pitch[1]는 유효하다. 만약 그렇지 않으면, crntLag(n)가 n = 0, 1, 2, …의 순서로 탐색되어 0.81×pitch[0]＜crntLag(n)＜1.2×pitch[0]을 만족시키는 crntLag(n)가 설정된다. 만약 함수 Ambiguous(pitch[0], pitch[1], 0.11)가 거짓이면, 처리는 스텝(S56)으로 간다.

스텝(S56)에서, 두 피치 pitch[0], pitch[1]와 상수 0.11사이의 관계를 설명하는 함수 Ambiguous(pitch[0], pitch[1], 0.11)가 참인지 거짓인지가 판단된다. 만약 함수가 참이면, 처리는 스텝(S57)으로 가서 trkPch = SearchPeaks(2)를 정한다. 만약 함수 Ambiguous(pitch[0], pitch[1], 0.11)가 거짓이면, 처리는 스텝(S58(trkPch = SearchPeaks3Frm()))으로 가서 rp(0), rp(1), rp(2)를 비교한다. 만약 rp[1]가 rp[0] 또는 rp[2]정도이거나 0.7보다 크면, pitch[1]가 사용된다. 만약 그렇지 않으면 상기한 SearchPeaks(frm)과 동일한 작동이 참조 프레임으로서 자기상관 피크값(rp[0] 및 rp[2]) 중 보다 큰 값을 갖는 프레임을 사용하여 행해진다.

상기한 스텝(S53)에서, '과거 프레임의 V/UV판정의 결과가 1임'의 여부, 즉 프레임이 V인지 여부가 판단된다. 만약 과거 프레임이 V이면, 처리는 스텝(S59)으로 가서 피치로서 trkPch = SearchPeaks(0)를 정한다. 만약 과거 프레임이 UV이면 처리는 스텝(S60)으로 간다.

스텝(S60)에서, '미래 프레임에 대한 V/UV판정의 결과가 1임'의 여부, 즉 미래 프레임이 V인지 여부가 판단된다. 만약 결과가 긍정이면, 처리는 스텝(S61)으로 가서 trkPch = SearchPeaks(0)가 피치로서 받아들여진다. 만약 미래 프레임이 UV이면, 처리는 스텝(S62)으로 가고 여기서 프레임의 피치 pitch[1]가 trkPch에 대한 피치로서 받아들여진다.

도 10a 내지 도 10c는 V/UV판정의 상기한 결과를 음성샘플의 피치검출에 적용한 결과를 나타낸다. 도 10a 내지 도 10c에서, 가로축과 세로축은 각각 프레임의 수와 피치를 나타낸다.

도 10a는 종래의 피치검출방법에 의해 검출된 피치 궤적을 나타내고, 도 10b는 도 10c에 도시된 고신뢰성 피치정보와 V/UV판정의 결과 모두가 사용된 본발명의 피치검출 방법에 의해 검출된 피치 궤적을 나타낸다.

이 결과로부터 본 발명의 피치검출 방법은 유성음(V)이라고 판단된 음성신호의 부분에 대해 고신뢰성 피치정보가 정해지고 소정의 시간(여기에서는 5프레임)동안 상기 값이 유효하다는 것을 알 수 있다. 이 결과 도 10a의 150번째 샘플에 나타난 피치부분을 갑작스럽게 변화할 때 잘못된 피치검출이 발생되지 않는다.

상기 신호 부호화 및 신호 복호화 장치는 예를들어 도 11 및 도 12에 도시된 휴대용 통신단말기 또는 휴대용 전화세트에 있는 음성 코덱(codec)로서 사용될 수 있다.

도 11은 도 1 및 도 3에 도시된 바와같이 구성된 음성 부호화부(160)를 사용하는 휴대용 단말기의 송신측을 나타낸다. 도 1의 마이크(161)에 의해 모아진 음성신호는 증폭기(162)에 의해 증폭되고 아날로그/디지털(A/D)변환기(163)에 의해 디지털신호로 변환되어 도 1 및 도 3에 도시된 바와같이 구성된 음성 부호화부(160)로 보내진다. A/D변환기(163)로부터의 디지털 신호는 입력 단자(101)로 보내진다. 음성 부호화부(160)는 도 1 및 도 3과 관련하여 설명된 부호화를 행한다. 도 1 및 도 3의 출력단자의 출력신호는 음성 부호화부(160)의 출력신호로서 송신 채널 부호화부(164)로 보내져서 이렇게 공급된 신호에 채널 부호화를 행한다. 송신 채널 부호화부(164)의 출력신호는 변조를 위해 변조회로(165)로 보내지고 그리고 나서 디지털/아날로그 변환기(166)와 RF 증폭기(167)를 통해 안테나(168)로 공급된다.

도 12는 도 2에 도시된 바와같이 구성된 음성 복호화부(260)를 사용하는 휴대용 단말기의 수신측을 나타낸다. 도 12의 안테나(261)에 의해 수신된 음성신호는 RF 증폭기(262)에서 증폭되고 아날로그/디지털(A/D) 변환기(263)를 통해 복조 회로(264)로 보내지고, 이로부터의 복조 신호는 송신 채널 복호부(265)로 보내진다. 복호부(265)의 출력신호는 도 2에 도시된 바와같이 구성된 음성 복호부(260)로 공급된다. 음성 복호부(260)는 도 2와 관련하여 설명된 방식으로 신호를 복호한다. 도 2의 출력단자(201)에서의 출력신호는 디지털/아날로그(D/A) 변환기(266)로 음성 복호부(260)의 신호로서 보내진다. D/A 변환기(266)로부터의 아날로그 음성신호는 스피커(268)로 보내진다.

본 발명은 상기한 실시예에 제한되지 않는다. 비록 도 1 및 도 3의 음성 분석측(부호화측)의 구조와 도 2의 음성 합성측(복호화측)의 구조가 하드웨어로서 설명되었지만, 디지털 신호 처리기(DSP)를 사용하는 소프트웨어에 의해 이행될 수도 있다. 또한 본 발명의 범위는 송신 또는 기록 및/또는 재생뿐만아니라 피치 또는 음성변환이나 규칙에 의한 음성합성 또는 잡음 압축 등의 다양한 다른 분야에도 적용될 수 있음을 알 수 있다.

발명의 효과

이상 설명한 바와같이 본 발명의 피치검출방법에 의하면, 소정의 피치검출조건하에서 피치정보의 검출을 행하는 피치탐색처리로, 시간축상의 현재 이외의 부호화단위의 음성신호에 대한 유성음/무성음 판정결과를 파라미터로서 이용하여 현재의 부호화단위의 음성신호의 피치를 결정하기 때문에 입력음성신호 중의 반 피치나 두배 피치를 오검출하지 않고 고정밀도로 피치검출을 행할 수 있다.

또, 본 발명의 음성신호 부호화방법 및 장치에 의하면, 입력음성신호에 대한 유성음/무성음 판정결과에 의거하여 입력음성신호의 유성음부분에 대해서는 사인파분석 부호화를 행하고, 무성음부분에 대해서는 파형부호화에 의한 부호화를 행하는 구성을 가지며 또한 상기의 본 발명의 피치검출방법을 적용하였기 때문에 효율이 좋고, 더구나 반 피치나 두배 피치를 오검출하지 않고 고정밀도의 부호화를 행할 수 있고, 무성음부분에서도 코가 막힌 듯한 감이 없는 명료도가 높은 재생음이 얻어지고, 유성음부분에 있어서도 자연스러운 합성음을 얻을 수 있다. 또, 무성음부분과 유성음부분과의 천이부분에서 이상음 등이 발생하지도 않는다.

(57) 청구의 범위

청구항 1.

입력 음성신호가 소정의 프레임에 의해 시간축상에서 나누어지고 이 프레임에 기초한 음성신호가 유성음/무성음에 관하여 판단되는 부호화 방법에서의 피치검출 방법에 있어서,

소정의 피치검출 조건하에서 피치정보를 검출하는 피치 탐색단계와,

상기 시간축상의 현재 프레임 이외의 입력된 음성신호에 대한 프레임의 유성음/무성음 판정의 결과에 기초하여 입력 음성신호에 대한 현재 프레임의 피치를 판별하는 피치 판별 단계를 포함하여 구성된 것을 특징으로 하는 피치검출방법.

청구항 2.

제 1항에 있어서,

상기 소정의 피치검출 조건하에서 상기 피치정보를 검출하는 상기 피치탐색 단계에서, 상기 입력 음성신호의 현재 프레임의 상기 피치는 상기 시간축상의 입력 음성신호의 과거 프레임에 대한 입력 음성신호의 유성음/무성음의 판정 결과를 파라미터로 사용하여 판정되는 것으로 구성된 것을 특징으로 하는 피치검출방법

청구항 3.

제 1항에 있어서,

상기 과거 프레임로부터 검출된 피치정보가 상기 현재 프레임을 위한 최종 피치를 판정하는 정보로서 사용되는지 여부를 선택하기 위해 상기 시간축상의 현재 프레임 이외의 프레임에 대한 상기 입력 음성신호의 유성음/무성음 판정을 사용하는 선택단계를 더 포함하여 구성된 것을 특징으로 하는 피치검출방법.

청구항 4.

입력 음성신호가 시간축상에서 프레임에 의해 나누어지고, 이 프레임 기초로 부호화된 음성신호 부호화 방법에 있어서,

상기 입력 음성신호의 피치를 검출하는 단계와,

상기 입력 음성신호의 단기간 예측 잔차를 구하는 예측 부호화 단계와,

상기 예측 부호화 단계에서 구해진 상기 단기간 예측 잔차에 사인파 분석 부호화를 행하는 사인파 분석 부호화 단계와,

상기 입력 음성신호를 파형 부호화하는 파형 부호화 단계와,

상기 프레임 기초로 상기 입력 음성신호의 유성음/무성음을 판단하는 판정단계를 포함하고,

현재 프레임의 상기 입력 음성신호의 피치는 또한 상기 시간축상의 현재 부호화이외의 프레임에 대한 입력된 음성신호의 판정 결과를 사용하여 판별되는 것을 특징으로 하는 음성신호 부호화 방법.

청구항 5.

제 4항에 있어서,

상기 사인파 분석 부호화 단계에 의해 얻어진 부호화된 음성은 상기 판정 단계에서 유성음이도록 구해진 프레임을 대해 출력되고, 상기 파형 부호화에 의해 부호화된 음성은 무성음이도록 구해진 프레임에 대해 출력되는 것을 특징으로 하는 음성신호 부호화 방법.

청구항 6.

입력 음성신호가 시간축상에서 프레임에 의해 나누어지고, 이 프레임 기초로 부호화된 음성신호 부호화 장치에 있어서,

상기 입력 음성신호의 피치를 검출하는 수단과,

상기 입력 음성신호의 단기간 예측 잔차를 구하는 예측 부호화 수단과,

상기 예측 부호화 단계에서 구해진 상기 단기간 예측 잔차에 사인파 분석 부호화를 행하는 사인파 분석 부호화 수단과,

상기 입력 음성신호를 파형 부호화하는 파형 부호화 수단과,

상기 프레임 기초로 상기 입력 음성신호의 유성음/무성음을 판단하는 판정수단을 포함하고,

현재 프레임의 상기 입력 음성신호의 피치는 또한 시간축상의 현재 부호화 이외의 프레임에 대한 입력된 음성신호의 판정의 결과를 사용하여 판별되는 것을 특징으로 하는 음성신호 부호화 장치.

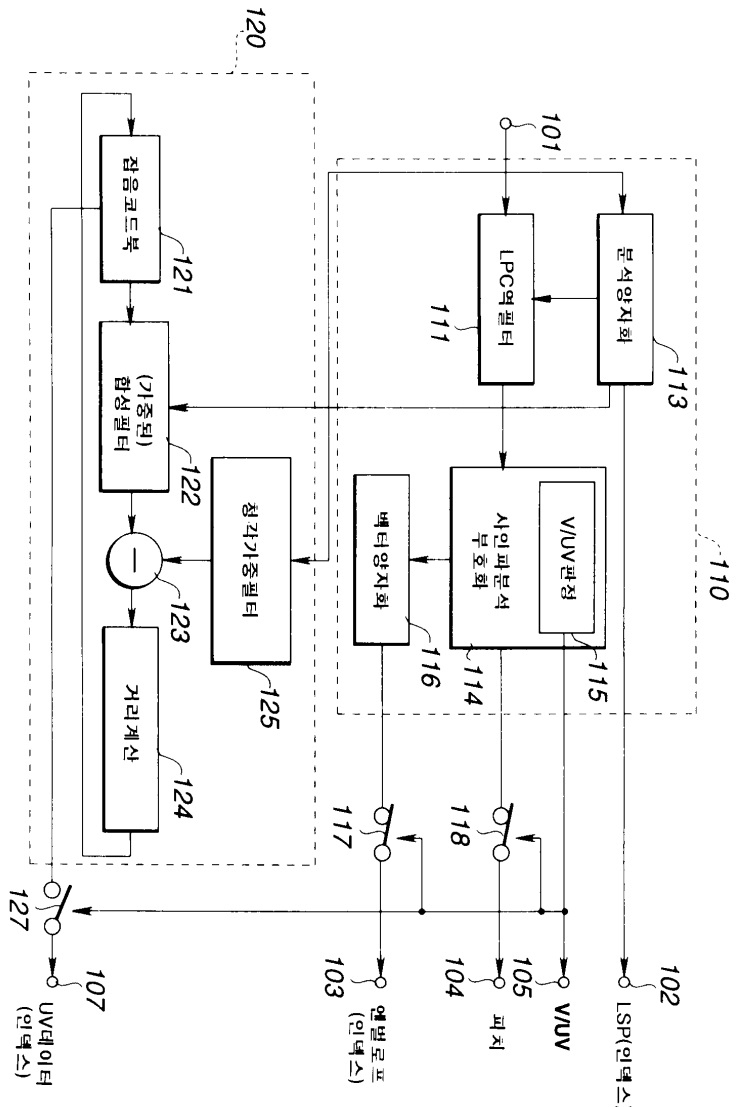
청구항 7.

제 6항에 있어서,

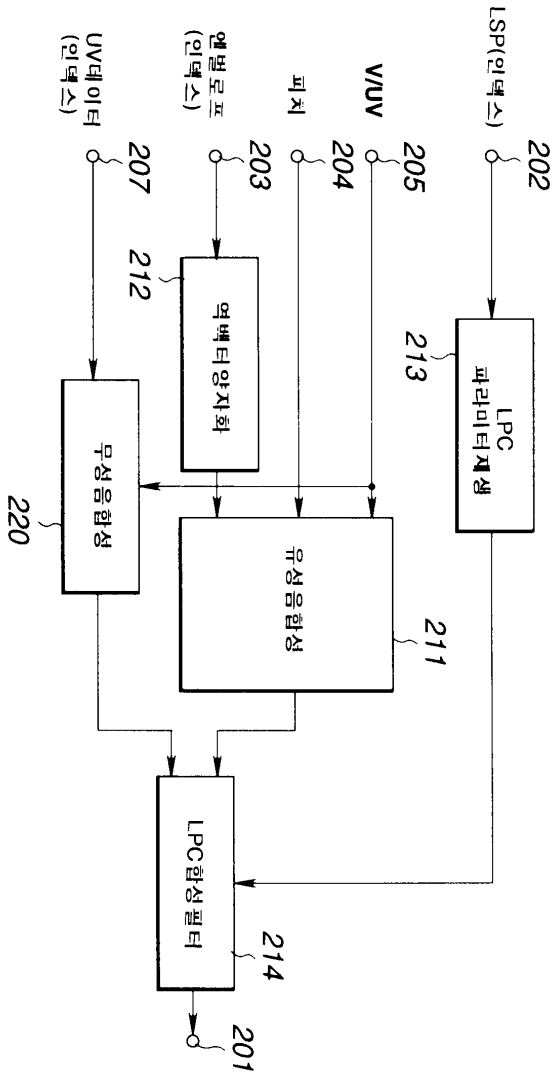
상기 사인파 분석 부호화 수단에 의해 얻어진 부호화된 음성은 상기 판정 수단에서 유성음이도록 구해진 프레임을 대해 출력되고, 상기 파형 부호화에 의해 부호화된 음성은 무성음이도록 구해진 프레임에 대해 출력되는 것을 특징으로 하는 음성신호 부호화 장치.

도면

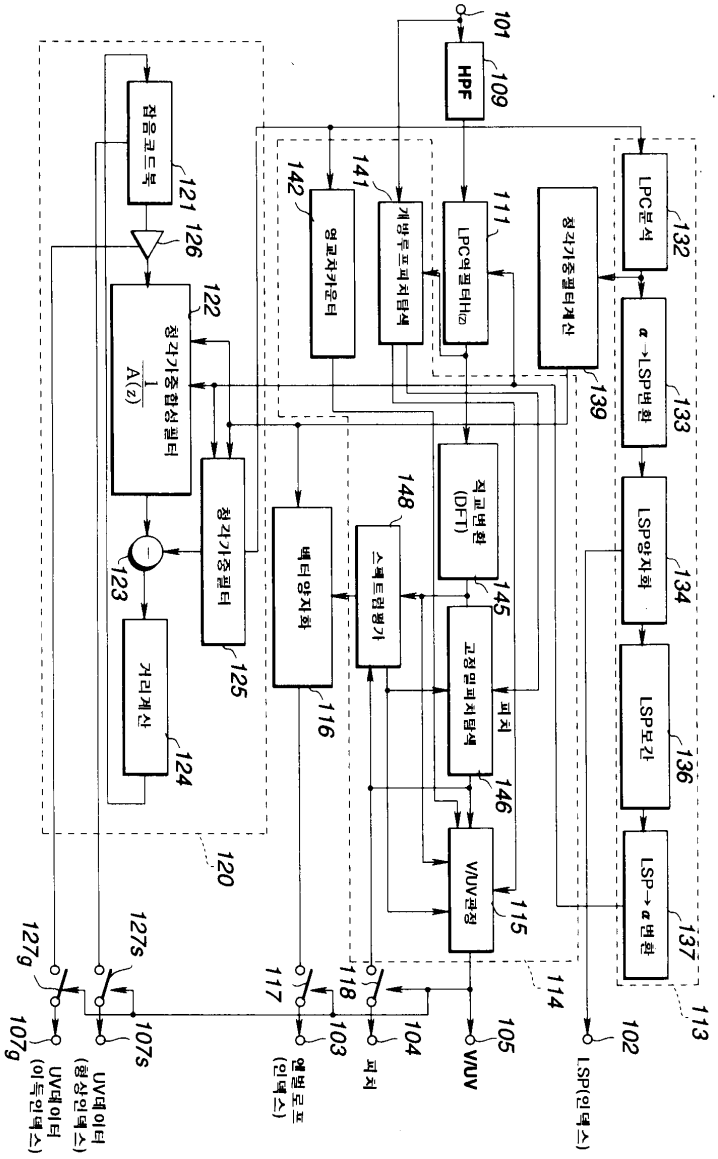
도면1


120
101
110
적응코드북
121
(가중된) 합성필터
122
청각가중필터
125
거리계산
124
123
LPC역필터
111
분석양자화
113
V/UV판정
115
사인파분석 부호화
114
벡터양자화
116
117
118
127
107
UV데이터 (인덱스)
103
에너지 (인덱스)
104
피치
105
V/UV
102
LSP(인덱스)

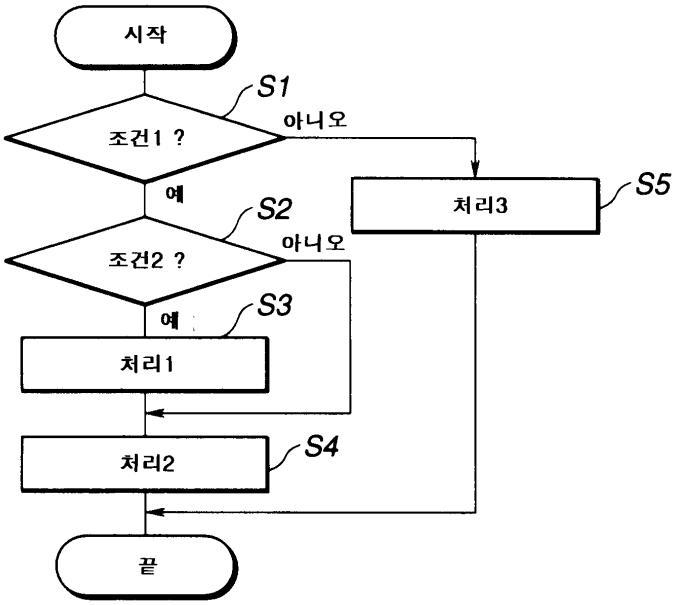
도면2


LSP(인덱스)
202
VUV
205
피치
204
엔벌로프
(인덱스)
203
UV데이터
(인덱스)
207
LPC
파라미터재생
213
역벡터양자화
212
유성음합성
211
무성음합성
220
LPC합성필터
214
201

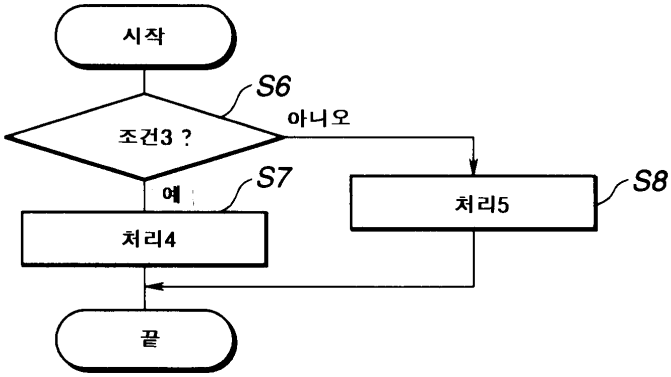
도면3


101
HPF
109
LPC분석
132
α→LSP변환
133
LSP양자화
134
LSP보간
136
LSP→α변환
137
113
LSP(인덱스)
102
청각가중필터계산
139
LPC역필터H(z)
111
개방루프피치탐색
141
영교차카운터
142
직교변환 (DFT)
145
스펙트럼평가
148
고정밀피치탐색
146
피치
V/UV판정
115
114
벡터양자화
116
청각가중필터
125
청각가중합성필터
122
1/A(z)
123
거리계산
124
잡음코드북
121
126
120
127g
127s
107g
107s
UV데이터 (이득인덱스)
UV데이터 (형상인덱스)
117
118
103
104
105
V/UV
피치
엘벌로프 (인덱스)

도면4


시작
S1
조건1 ?
아니오
예
S2
조건2 ?
아니오
예
S3
처리1
처리2
S4
처리3
S5
끝

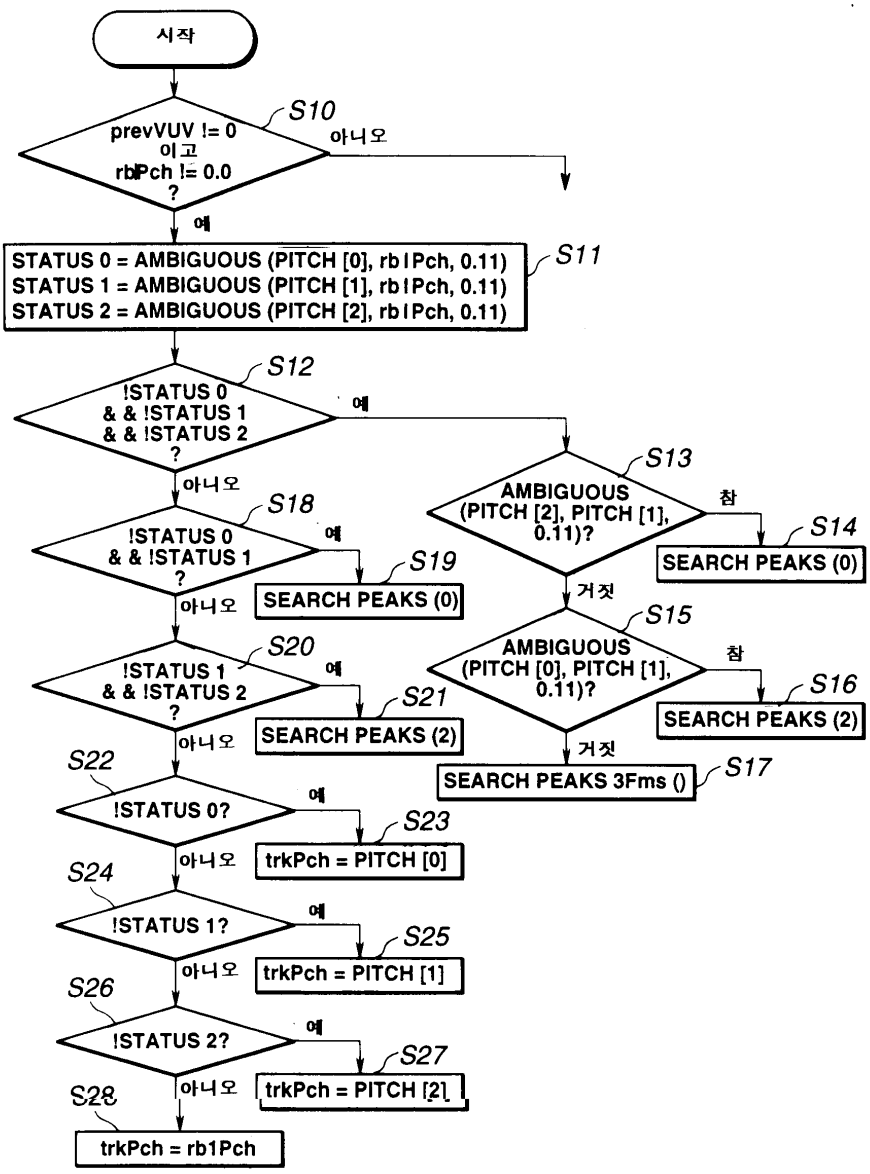
도면5


시작
S6
조건3 ?
아니오
예
S7
처리4
처리5
S8
끝

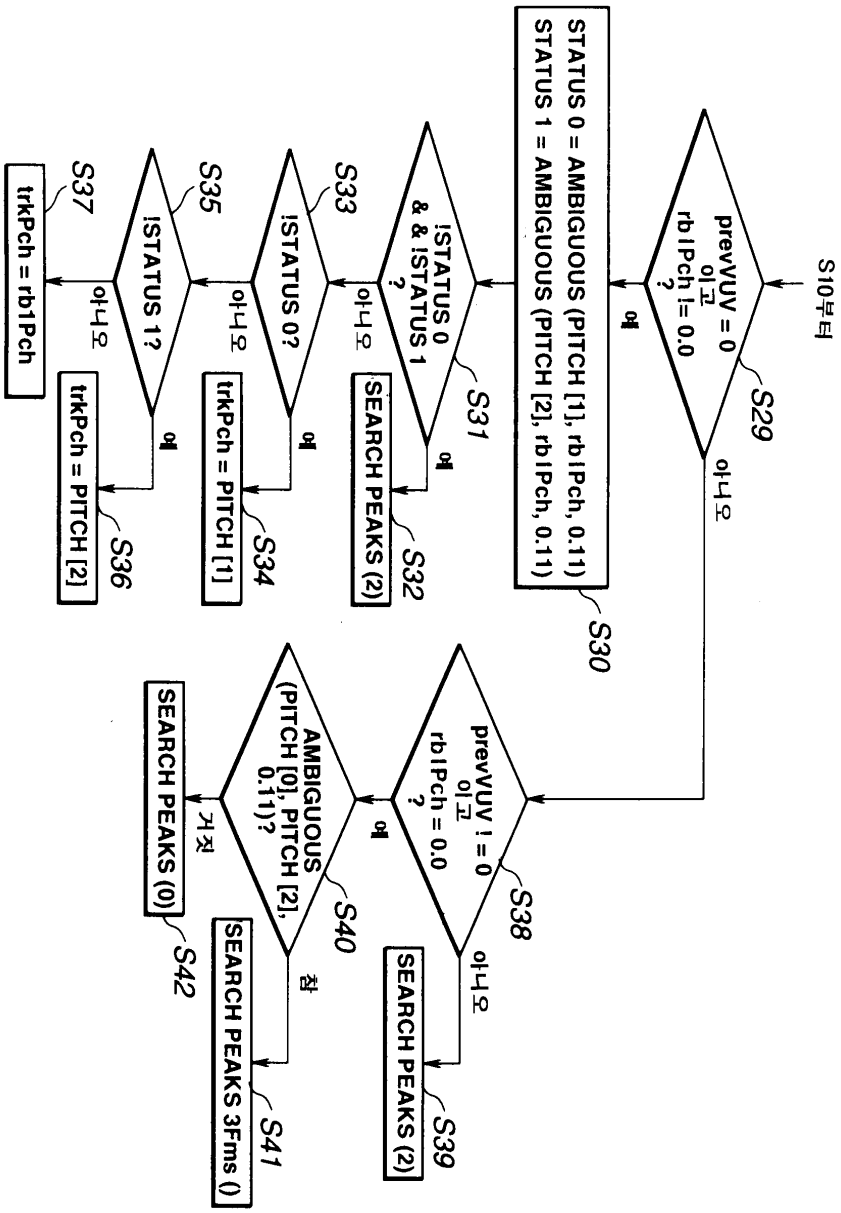
도면6

<table>
<tr><td></td><td colspan="3">2Kbps</td><td colspan="3">6Kbps</td></tr>
<tr><td>V/UV판정출력</td><td colspan="3">1bit / 20msec</td><td colspan="3">1bit / 20msec</td></tr>
<tr><td>LSP양자화인덱스</td><td colspan="3">32bits / 40msec</td><td colspan="3">48bits / 40msec</td></tr>
<tr><td rowspan="4">유성음(V)용</td><td colspan="2">피치데이터</td><td>8bits / 20msec</td><td colspan="2">피치데이터</td><td>8bits / 20msec</td></tr>
<tr><td colspan="2">인덱스15bit/20msec</td><td colspan="2">인덱스87bit/20msec</td></tr>
<tr><td>형상
(제1단)
이득</td><td>5+5bits / 20msec</td><td rowspan="2">5bits / 20msec</td><td>형상
(제1단)
이득
인덱스
(제2단)</td><td>5+5bits / 20msec</td></tr>
<tr><td colspan="2"></td><td></td><td>72bits / 20msec</td></tr>
<tr><td rowspan="4">무성음(UV)용</td><td colspan="2">인덱스11bit/10msec</td><td colspan="2">인덱스23bit/5msec</td></tr>
<tr><td>형상
(제1단)
이득</td><td>7bits / 10msec</td><td rowspan="2">4bits / 10msec</td><td>형상
(제1단)
이득
형상
(제2단)
이득</td><td>9bits / 5msec</td></tr>
<tr><td colspan="2"></td><td>6bits / 5msec</td></tr>
<tr><td colspan="2"></td><td colspan="2">5bits / 5msec</td><td>3bits / 5msec</td></tr>
<tr><td>유성음용
무성음용</td><td colspan="3">40bits / 20msec
39bits / 20msec</td><td colspan="3">120bits / 20msec
117bits / 20msec</td></tr>
</table>

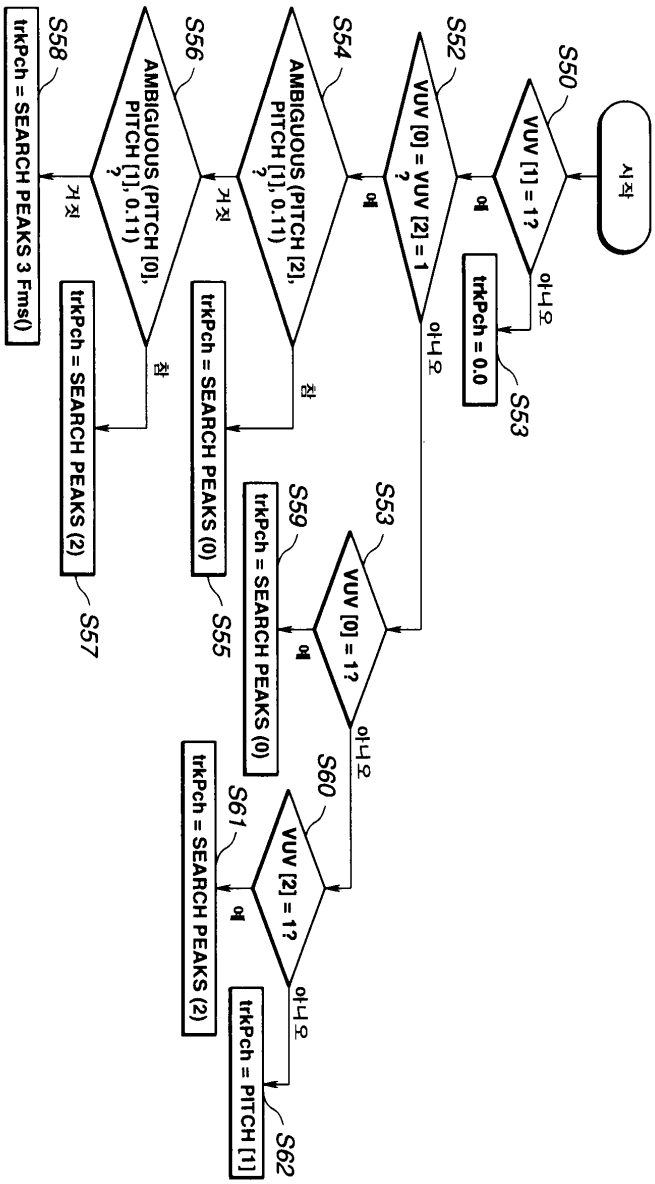
도면7


시작
S10
prevVUV != 0
이고
rb1Pch != 0.0
?
아니오
예
STATUS 0 = AMBIGUOUS (PITCH [0], rb1Pch, 0.11)
STATUS 1 = AMBIGUOUS (PITCH [1], rb1Pch, 0.11)
STATUS 2 = AMBIGUOUS (PITCH [2], rb1Pch, 0.11)
S11
S12
!STATUS 0
& & !STATUS 1
& & !STATUS 2
?
예
아니오
S13
AMBIGUOUS
(PITCH [2], PITCH [1],
0.11)?
참
S14
SEARCH PEAKS (0)
거짓
S15
AMBIGUOUS
(PITCH [0], PITCH [1],
0.11)?
참
S16
SEARCH PEAKS (2)
거짓
SEARCH PEAKS 3Fms ()
S17
S18
!STATUS 0
& & !STATUS 1
?
예
S19
SEARCH PEAKS (0)
아니오
S20
!STATUS 1
& & !STATUS 2
?
예
S21
SEARCH PEAKS (2)
아니오
S22
!STATUS 0?
예
S23
trkPch = PITCH [0]
아니오
S24
!STATUS 1?
예
S25
trkPch = PITCH [1]
아니오
S26
!STATUS 2?
예
S27
trkPch = PITCH [2]
아니오
S28
trkPch = rb1Pch

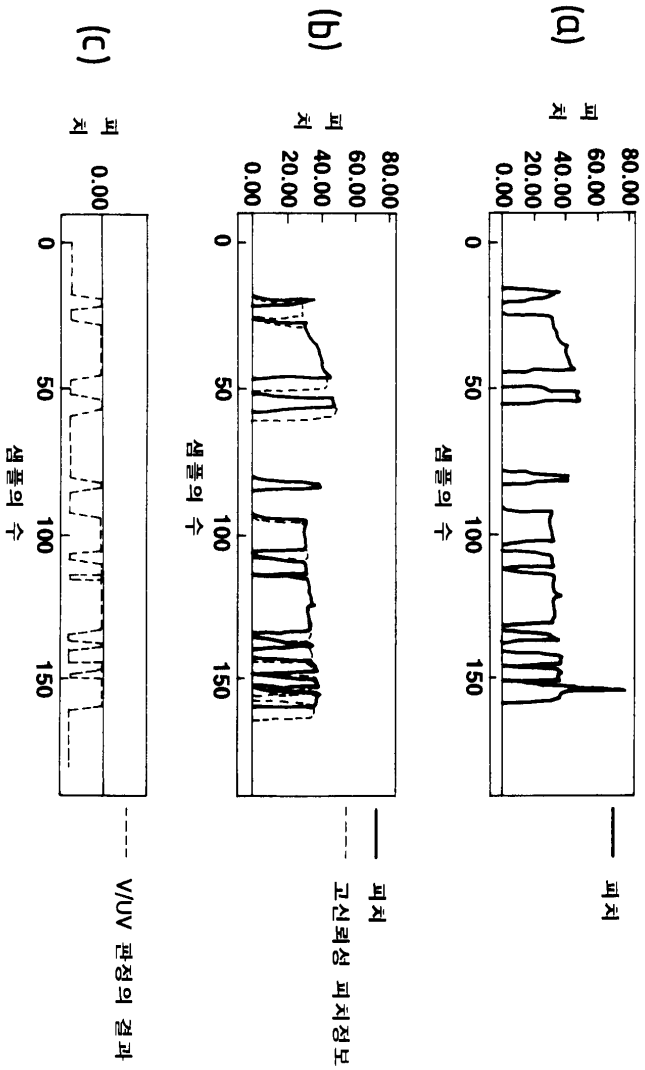
도면8


S10부터
prevVUV = 0 이고 rb1Pch != 0.0 ?
S29
예
아니오
STATUS 0 = AMBIGUOUS (PITCH [1], rb1Pch, 0.11)
STATUS 1 = AMBIGUOUS (PITCH [2], rb1Pch, 0.11)
S30
!STATUS 0 && !STATUS 1 ?
S31
예
아니오
SEARCH PEAKS (2)
S32
!STATUS 0?
S33
예
아니오
trkPch = PITCH [1]
S34
!STATUS 1?
S35
예
아니오
trkPch = PITCH [2]
S36
trkPch = rb1Pch
S37
prevVUV != 0 이고 rb1Pch = 0.0 ?
S38
예
아니오
SEARCH PEAKS (2)
S39
AMBIGUOUS (PITCH [0], PITCH [2], 0.11)?
S40
참
거짓
SEARCH PEAKS 3Fms ()
S41
SEARCH PEAKS (0)
S42

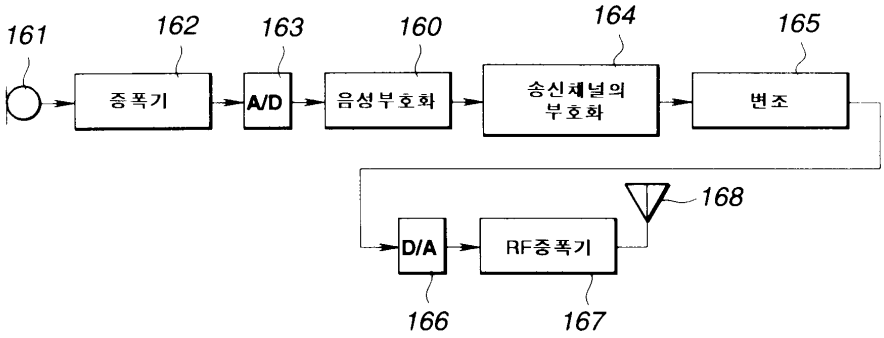
도면9


시작
S50
VUV [1] = 1?
예
아니오
trkPch = 0.0
S53
S52
VUV [0] = VUV [2] = 1 ?
예
아니오
S54
AMBIGUOUS (PITCH [2], PITCH [1], 0.11) ?
거짓
참
S56
AMBIGUOUS (PITCH [0], PITCH [1], 0.11) ?
거짓
참
trkPch = SEARCH PEAKS 3 Fms()
S58
trkPch = SEARCH PEAKS (2)
S57
trkPch = SEARCH PEAKS (0)
S55
S53
VUV [0] = 1?
예
아니오
S59
trkPch = SEARCH PEAKS (0)
S60
VUV [2] = 1?
예
아니오
S61
trkPch = SEARCH PEAKS (2)
trkPch = PITCH [1]
S62

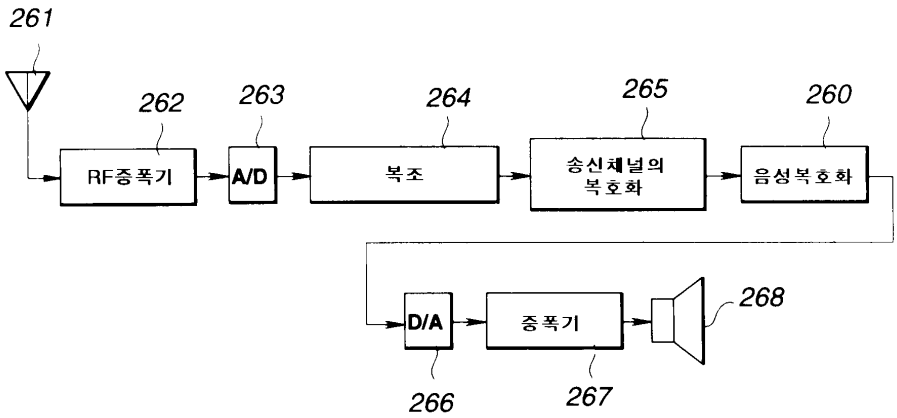
도면10


(a)
(b)
(c)
피치
0.00
20.00
40.00
60.00
80.00
0
50
100
150
샘플의 수
피치
고신뢰성 피치정보
V/UV 판정의 결과

도면11


161
162
163
160
164
165
168
증폭기
A/D
음성부호화
송신채널의 부호화
변조
D/A
RF증폭기
166
167

도면12


261
262
RF증폭기
263
A/D
264
복조
265
송신채널의
복호화
260
음성복호화
D/A
266
증폭기
267
268