(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl.[7]
G10L 19/00

(45) 공고일자 2004년07월27일
(11) 등록번호 10-0427753
(24) 등록일자 2004년04월07일

(21) 출원번호 10-1996-0047283
(22) 출원일자 1996년10월21일

(65) 공개번호 10-1998-0028284
(43) 공개일자 1998년07월15일

(30) 우선권주장
95-279410 1995년10월26일 일본(JP)
95-280672 1995년10월27일 일본(JP)
96-270337 1996년10월11일 일본(JP)

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

(72) 발명자
이지마 가쯔유끼
일본국 도꾜도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

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

마츠모또 준
일본국 도꾜도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

오모리 시로
일본국 도꾜도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

(74) 대리인
신관호

심사관 : 조광현

(54) 음성신호재생방법및장치,음성복호화방법및장치,음성합성방법및장치와휴대용무선단말장치

요약

제어된 속도로 음성신호를 재생하는 방법으로 이에의해 시간축의 레이트(rate)변환이 용이해지고, 음성을 합성하는 방법으로 이에의해 음소의 변화없이 부호화된 음성데이터에 의거하여 간단한 구조로 음조변환이 실현될 수 있는 것에 관한 것이다. 이 음성재생방법에 의하면, 부호화장치(2)가 입력음성데이터가 유성음인지 무성음인지를 판별한다. 판별결과에 의거하여, 부호화장치(2)는 유성음으로 판별된 신호부분에 대하여는 정현합성과 부호화를 수행하는 한편, 무성음으로 판별된 신호부분에 대하여는 합성에 의해 분석방법을 이용하여 폐루프탐색으로 최적벡터를 탐색하는 벡터양자화를 수행하여 부호화된 파라미터를 구한다. 복호화장치(3)는 주기변형장치(4)에서 매 소정 프레임마다 구해지는 부호화된 파라미터의 시간축을 컴팬드(compand)하여 파라미터의 출력주기를 변형하여 소정 프레임에 해당하는 다른 시간점에 관련된 변형된 부호화된 파라미터를 생성한다.
음성합성장치(6)는 변형된 부호화된 파라미터에 의거하여 유성음부분과 무성음부분을 합성한다. 이 음성합성장치에 의하면, 부호화된 비트스트림(bit stream)이나 부호화된 데이터는 부호화된 데이터출력장치(301)에 의해 출력된다. 이들 데이터중에서 적어도 스펙트럼엔벌로프의 진폭데이터와 음조데이터는 데이터변환장치(302)를 거쳐 파형합성장치(303)에 보내져서 거기서 소망하는 음조값에 따라 스펙트럼엔벌로프의 형상의 변화없이 스펙트럼엔벌로프의 진폭데이터 수가 변화된다. 파형합성장치는 변환된 스펙트럼엔벌로프데이터와 음조데이터에 의거 하여 음성파형을 합성한다.

대표도

도 1

명세서

도면의 간단한 설명

도 1은 본 발명에 따른 음성신호재생방법을 수행하기 위한 음성신호재생방법 및 음성신호재생장치의 기본구조를 나타내는 블록도이다.
도 2는 도 1에 도시된 음성신호재생장치의 부호화장치를 나타내는 개략적 블록도이다.
도 3은 부호화장치의 상세한 구조를 보여주는 블록도이다.
도 4는 도 1에 도시된 음성신호 재생장치의 복호화장치의 구조를 나타내는 개략적 블록도이다.
도 5는 복호화장치의 상세한 구조를 보여주는 블록도이다.
도 6은 복호화장치의 변형된 부호화 파라미터를 계산하기 위한 장치의 작동을 설명하는 흐름도이다.
도 7은 변형된 부호화 파라미터 계산장치에 의해 시간축상에 구해진 변형된 부호화 파라미터를 개략적으로 설명한다.
도 8은 변형된 부호화 파라미터 계산장치에 의해 수행되는 상세한 보간처리작동을 설명하는 흐름도이다.
도 9a 도 9d는 보간처리작동을 설명한다.
도 10a 도 10c는 변형된 부호화 파라미터를 계산하는 장치에 의해 수행되는 전형적인 작동을 설명한다.
도 11a 도 11c는 변형된 부호화 파라미터를 계산하는 장치에 의해 수행되는 다른 전형적인 작동을 설명한다.
도 12는 복호화장치에 의해 빨리 속도제어가 되도록 하기 위해 프레임길이가 다양하도록 된 경우의 작동을 설명한다.
도 13은 복호화장치에 의해 천천히 속도제어가 되도록 하기 위해 프레임길이가 다양하도록 된 경우의 작동을 설명한다.
도 14는 복호화장치의 또다른 상세구조를 보여주는 블록도이다.
도 15는 음성합성장치에의 응용예를 보여주는 블록도이다.
도 16은 텍스트 음성합성장치에의 응용예를 보여주는 블록도이다.
도 17은 부호화장치를 채용한 휴대용 단말기의 송신기의 구조를 설명하는 블록도이다.
도 18은 부호화장치를 채용하는 휴대용 단말기의 수신기의 구조를 설명하는 블록도이다.
* 도면의 주요부분에 대한 부호설명
1. 음성신호재생장치 2. 부호화장치
3. 주기변형장치 4. 복호화장치
5. 파라미터변형장치 6. 음성합성장치
101, 202, 203, 204, 205, 207. 입력단자
102, 103, 104, 105, 107, 201. 출력단자
109. 고역통과필터 110. 제 1부호화장치
111. 역LPC필터 113. LPC분석/양자화장치
114. 정현분석부호화장치 115. V/UV구별장치
116. 벡터양자화기 120. 제 2부호화장치
211. 유성음성합성장치 220. 무성음성합성장치

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 음성신호를 제어된 속도로 재생하는 방법 및 장치와, 그 음성을 복호화하는 방법 및 장치와 그 음성을 합성하는 방법 및 장치로 이에의해 음조변환이 단순화된 구조로 실현될 수 있는 것에 관한 것이다. 본 발명은 또한 음조변환된 음성신호를 송신 및 수신하는 휴대용 무선단말장치에 관한 것이다.
이제까지 시간영역과 주파수영역내 신호의 통계적특성과 인간 귀의 청감심리특성을 이용하여 압축하기 위하여 오디오신호(음성신호와 음향신호를 포함)를 부호화하는 여러가지 부호화방법이 알려져 왔다. 이 부호화방법은 대체적으로 시간영역부호화, 주파수영역부호화 및 분석/합성부호화로 분류된다.
음성신호의 이 고효율 부호화의 예는 고조파부호화, 다대역여기부호화(MBE), 부대역코딩(SBC), 선형예측코딩(LPC), 이산코사인변환(DCT), 변조 DCT(MDCT), 고속푸리에변환(FFT)과 같은 정현분석부호화를 포함한다.

한편, 코드여기된 선형예측(CELP)부호화로 대표되는 시간축 프로세싱에 의한 고효율 음성부호화방법은 디코더출력에 곧이어 많은 분량의 프로세싱작동의 수행 필요성 때문에 급한 시간축 변환(변형)에서 어려움을 포함하고 있다. 더욱이, 복호화에 연속하여 속도제어가 시간영역에서 수행되므로 이 방법은 비트레이트변환(bit rate conversion)에는 사용될 수 없다.
반면에, 상기 부호화방법에 의해 부호화된 음성신호를 복호화는 것이 시도된다면, 그 음성의 음소를 변화시키지 않고 음조만을 변화시키는 것이 종종 소망된다. 그러나 그 보통의 음성복호화방법으로는, 복호된 음성이 음조제어를 이용하여 음조-변환되어야 하고 따라서 복잡한 구조로 비용을 상승시킨다.

발명이 이루고자 하는 기술적 과제

따라서, 본 발명의 목적은 넓은 범위에서 소망하는 레이트(rate)로의 속도제어가 음소나 음조의 변화없이 고음질로 이루어질 수 있는 음성신호재생방법 및 장치는 제공하는 것이다.
본 발명의 다른 목적은 음조변환이나 음조제어가 단순한 구조로 달성될 수 있는 음성신호 복호화장치 및 방법과 음성신호 합성방법 및 장치를 제공하는 것이다.
본 발명의 또다른 목적은 이에의해 음조변환되거나 음조제어된 음성신호가 단순한 구성으로 송수신될 수 있도록 하는 것이다.
본 발명에 따른 음성신호 재생방법에서는 입력음성신호가 시간축상에 소정의 부호화장치의 견지에서 분할되어 부호화된 파라미터를 생성하고 이것은 보간처리되 어 소망하는 시간점을 위한 변형된 부호화된 파라미터를 생성하고, 그 다음 이러한 변형된 부호화된 파라미터에 의거하여 음성신호가 재생된다.
본 발명에 따른 음성신호 재생장치에서는, 입력음성신호가 시간축상에 소정의 부호화장치의 견지에서 분할되어 부호화된 파라미터를 생성하고 이것은 보간처리되어 소망하는 시간점을 위한 변형된 부호화된 파라미터를 생성하고, 그 다음 이러한 변형된 부호화된 파라미터에 의거하여 음성신호가 재생된다.
이 음성신호재생방법에서는 입력음성신호를 시간축상에 소정의 블록을 단위로 하여 분할하고 분할된 음성신호를 부호화된 블록의 견지에서 부호화하여 얻은 부호화된 파라미터를 이용하여, 부호화에 사용되었던 것과는 다른 블록길이로 음성이 재생된다.
본 발명에 따른 음성복호화방법 및 장치에서는, 입력부호화된 음성데이터의 고조파의 소정대역내 기본적 주파수와 개수는 변환되고, 각 입력고조파내 스펙트럼성분의 진폭을 특정하는 데이터의 수는 음조를 변형시키기 위하여 보간처리된다.
음조주파수는 고조파가 소정값으로 설정되어 있는 차원변환에 의한 부호화시에 변형된다.
이 경우, 음성압축을 위한 디코더가 동시에 텍스트음성합성을 위한 음성합성기로서 사용되어질 수 있다. 일상적인 음성발음을 위하여는 클리어(clear) 재생음성이 압축과 팽창으로 얻어지는 한편, 특별한 음성합성을 위하여는, 텍스트합성, 또는 소정의 규칙하에서의 합성이 효과적인 음성출력시스템을 구성하기 위하여 사용된다.
본 발명에 따른 음성신호 재생방법 및 장치에 의하면, 입력음성신호가 시간축상에 소정의 부호화장치의 견지에서 분할되고 그 부호화장치의 견지에서 부호화되어 부호화된 파라미터를 구하고, 그다음 이것은 보간처리되어 소망하는 시간점을 위한 변형된 부호화된 파라미터를 구한다. 그다음 이 음성신호는 변형된 부호화된 파라미터에 의거하여 재생되고 따라서 넓은 범위에 걸친 속도제어가 음소나 음조의 변화없이 고음질로 쉽게 실현될 수 있게된다.
본 발명에 따른 이 음성신호재생방법 및 장치에 의하면, 입력음성신호를 시간축상에 소정의 블록을 단위로하여 분할하고 분할된 음성신호를 부호화된 블록의 견지에서 부호화하여 얻은 부호화된 파라미터를 이용하여, 부호화에 사용되었던 것과는 다른 블록길이로 음성이 재생된다. 그 결과, 넓은 범위에 걸친 속도제어가 음소나 음조의 변화없이 고음질로 쉽게 실현될 수 있다.
본 발명에 따른 음성복호화방법 및 장치에서는, 입력부호화된 음성데이터의 고조파의 소정대역내 기본적 주파수와 개수는 변환되고, 각 입력고조파내 스펙트럼성분의 진폭을 특정하는 데이터의 수는 음조를 변형시키기 위하여 보간처리된다. 그 결과 음조는 단순화된 구조로서 소망하는 값으로 변화될 수 있다.
이 경우, 음성압축을 위한 디코더는 동시에 텍스트음성합성을 위한 합성기로서도 사용될 수 있다. 일상적인 음성발음을 위하여는, 클리어 재생음성이 압축 및 팽창으로 구해지는 한편, 특별한 음성합성을 위하여는, 텍스트합성이나 규칙하의 합성이 효율적인 음성출력시스템을 구성하기 위해 사용된다.
휴대용 단말기장치로는 음조제어된 음성신호로 음조변환된 음성신호가 단순 한 구조로 송수신될 수 있다.

발명의 구성 및 작용

도면을 참고하여 본 발명의 선택된 실시예에 따른 음성신호 재생방법 및 장치가 설명된다. 본 발명은 도 1에 도시하는 바와같이 입력음성신호를 시간축상에 부호화장치로서의 소정 개수의 프레임에 대하여 분할하고 분할된 입력음성신호를 부호화함으로써 구해지는 부호화 파라미터에 의거하여 음성신호를 재생하는 음성신호 재생장치에 구현된다.
음성신호 재생장치(1)는 선형예측부호화(LPC)파라미터, 선형스펙트럼쌍(LSP)파라미터, 음조, 유성음(V)/무성음(UV) 또는 스펙트럼진폭(Am)과 같은 부호화된 파라미터를 출력하도록 입력단자(101)에 입력되는 음성신호를 프레임을 단위로하여 부호화하는 부호화장치(2)와, 부호화 파라미터의 출력주기를 시간축 컴팬션 (compansion)에 의해 변형하는 주기변형장치(3)와를 포함한다. 또한 음성신호 재생장치는 주기변형장치(3)에 의해 변형된 주기로 출력되는 부호화된 파라미터들을 보간처리하여 소망하는 시간점들에 대한 변형된 부호화된 파라미터를 구하고, 변형된 부호화된 파라미터들에 의거하여 음성신호를 합성하여, 합성된 음성신호를 출력단자(201)로 출력하는 복호화장치(4)를 포함한

다.

부호화장치(2)가 도 2와 도 3을 참조하여 설명된다. 부호화장치(2)는 판별결과에 의거하여, 입력음성신호가 유성음인지 또는 무성음인지를 결정하고 유성음이라고 밝혀진 신호부분에 대하여는 정현합성부호화를 수행하는 한편, 무성음이라고 밝혀진 신호부분에 대하여는 합성에 의한 분석방법을 이용하여 최적벡터의 탐색에 의해 벡터 양자화를 수행하여, 부호화된 파라미터들을 구한다. 즉, 부호화장치(2)는 선형예측코딩(LPC)잔차와 같은 입력음성신호의 단기 예측잔차를 구하여 고조파부호화와 같은 정현분석부호화를 수행하는 제 1부호화장치(110)와, 입력음성신호의 위상성분을 송신함으로써 파형코딩을 수행하는 제 2부호화장치(120)와를 포함한다. 제 1부호화장치(110)와 제 2부호화장치(120)는 각각 유성음 부분(V)과 무성음 부분 (UV)을 부호화하는데 사용된다.

도 2의 실시예에서, 입력단자(101)에 공급된 음성신호는 역 LPC필터(111)와 제 1부호화장치(110)의 LPC분석/양자화장치(113)에 보내진다. LPC 분석/양자화장치 (113)로부터 구해진 LPC계수 또는 소위 α-파라미터는 역 LPC필터(111)에 보내져 역 LPC필터(111)에 의해 입력음성신호의 선형예측잔차(LPC잔차)를 추출한다. LPC 분석/양자화장치(113)로부터, 선형스펙트럼쌍(LSP)의 양자화된 출력이 후술하는 바와같이 취하여져서 출력단자(102)로 보내진다. 역 LPC필터(111)로부터의 LPC잔차는 정현분석부호화장치(114)에 보내진다. 정현분석부호화장치(114)는 음조검출, 스펙트럼 엔벌로프 진폭계산 및 V/UV선별을 유성(V)/무성(UV) 구별장치(115)에 의해 수행한다. 정현분석화장치(114)로부터의 스펙터럼 엔벌로프 진폭데이터는 벡터양자화장치(116)에 보내진다. 스펙트럼 엔벌로프의 벡터-양자화된 출력으로서의 벡터양자화장치(116)의 코드북지수는 스위치(117)를 통하여 출력단자(103)에 보내지고, 한편, 정현분석부호화장치(114)의 출력은 스위치(118)를 통하여 출력단자(104)에 보내진다. 유성(V)/무성(UV) 구별장치(115)로부터의 유성(V)/무성(UV) 구별 출력은 출력단자(105)와 스위치(117, 118)에 스위칭 제어신호로서 보내진다. 유성(V)신호 에 대하여는, 지수와 음조가 출력단자(103, 104)에서 꺼내질 수 있도록 선택된다. 벡터양자화장치(116)에서의 벡터양자화에 대하여는, 주파수축상에서 유효한 대역블록의 진폭데이터를 그 블록내 마지막 진폭데이터로부터 그 블록내 첫번째 진폭데이터까지 보간처리하기 위한 적당한 개수의 더미(dummy)데이터나 그 블록내 마지막 데이터와 첫번째 데이터를 연장시키는 더미데이터가 그 블록의 마지막과 처음에 첨부되어 N_F 까지의 데이터 개수를 향상시키도록 되어 있다. 그다음 진폭데이터의 Os-튜플(tuple) 개수가 옥타튜플 오버샘플링(octatuple oversampling)과 같은 대역제한형 Os-튜플 오버샘플링에 의해 구해진다. 진폭데이터의 Os-튜플 개수((mMx + 1) ×데이터의 Os넘버)는 선형보간에 의해 21048과 같은 더 큰 수의 N_M 으로 더욱 팽창된다. 이 N_M 넘버데이터는 데시메이션(decimation)에 의해 소정개수 (M)(44등)로 전환되고, 그다음 이 소정개수의 데이터에 대하여 벡터양자화가 수행된다.

본 실시예에서, 제 2부호화장치(120)는 부호여기된 선형예측(CELP)코딩구조를 가지고, 합성에 의한 분석방법을 채용하는 페루프탐색에 의해 시간영역 파형에 벡터양자화를 수행한다. 상술하면, 잡음코드북(121)은 가중 합성필터(122)에 의해 합성되어 가중된 합성 음성을 만들고 이는 감산기(123)에 보내지고, 여기서 가중된 합성음성과 입력단자(101)에 공급된 후 청감가중필터(125)에 의해 처리된 음성과의 사이의 에러를 구해진다. 거리계산회로(124)는 그 거리를 계산하고, 그 에러를 최소화하는 벡터가 잡음코드북(121)에서 탐색된다. 이 CELP코딩은 상술한 바와같이 무성음 부분을 부호화하는데 사용된다. 잡음코드북(121)으로부터의 UV데이터로서의 코드북지수는 유성(V)/무성(UV) 구별장치(115)로부터의 V/UV구별의 결과가 무성음 (UV)을 나타낼 때 턴온되는 스위치를 통하여 출력단자(107)에서 취해진다.

도 3을 참조하여 도 1에 도시된 음성신호 엔코더의 더욱 상세한 구조가 설명된다. 도 3에서, 도 1과 동일한 부품과 성분들에는 같은 참조번호가 부여된다.

도 3에 도시된 음성신호엔코더(2)에서, 입력단자(101)에 공급된 음성신호는 불필요한 레인지의 신호를 제거하기 위해 고역통과필터(109)에 의하여 필터된 후 LPC 분석/양자화장치(113)의 LPC분석회로(132)와 역 LPC필터(111)에 공급된다.

LPC 분석/양자화장치(113)의 LPC분석회로(132)는 256샘플의 계수로 그 입력신호파형의 길이를 한 블록으로서 가지고 해밍 윈도(Hamming window)를 응용하여, 자기-교정방법에 의해 소위 α-파라미터라고 하는 선형예측계수를 구한다. 데이터출력단위로서의 프레이밍 간격은 약 160샘플로 설정된다. 만일 샘플링주파수가 예를들어 8 kHz라면, 1-프레임간격은 20 msec나 160 샘플이다.

LPC분석회로(132)로부터의 α-파라미터는 α-LSP변환회로(133)에 보내져 선스펙트럼쌍(LSP)파라미터로 변환된다. 이것은 다이렉트형필터계수로서 구해지는 바와같이 α-파라미터를 예를들어 10, 즉 5쌍의 LSP파라미터로 변환한다. 이것은 예를들어 뉴튼-랩슨(Newton-Rhapson) 방법에 의해 수행된다. α-파라미터가 LSP파라미터로 변환되는 이유는 LSP파라미터가 α-파라미터에 비해 보간특성이 우수하기 때문이다.

α-LSP 변환회로(133)로부터의 LSP파라미터는 LSP양자화장치(134)에 의해 매트릭스(행렬)- 또는 벡터-양자화된다. 벡터 양자화 이전에 프레임 대 프레임 차이 를 취하거나, 매트릭스 양자화를 위하여 복수 프레임을 함께 수집하는 것이 가능하다. 본 케이스에서는 20 msec 마다 계산되는 LSP파라미터가 20 msec를 한 프레임으로 하여 벡터-양자화된다.

LSP양자화의 지수데이터인, 양자화장치(134)의 양자화된 출력은 단자(102)에서 복호화장치(103)로 취해지고, 한편, 양자화된 LSP벡터는 LSP보간처리회로(136)에 보내진다.

LSP보간처리회로(136)는 20msec 또는 40msec 마다 양자화된 LSP벡터를 보간처리하여 옥타튜플 레이트(octatuple rate)를 제공하도록 한다. 즉, LSP벡터는 매 2.5msec마다 갱신된다. 그 이유는, 만일 잔차파형이 고조파 부호화/복호화방법에 의해 분석/합성처리되면, 합성파형의 엔벌로프가 극도로 누그러진 파형을 나타내므로, LPC계수가 매 20 msec마다 갑작스럽게 변화하는 경우에는, 이상잡음이 생성되기 쉽기 때문이다. 즉, LPC계수가 매 2.5msec마다 점차적으로 변하게 되면, 그러한 이상잡음의 발생을 억제할 수 있을 것이다.

매 2.5msec마다 생성된 보간된 LSP벡터를 사용하는 입력음성의 역필터링을 위해, LSP파라미터는 α변환회로(137)에서 LSP에 의해 10-차수의 다이렉트형 필터의 계수로써 α-파라미터로 변환된다. α변환회로(137)에서의 LSP의 출력은 LPC역필터회로(111)로 인가되며, 여기서 역필터처리가 수행되어 매 2.5msec마다 갱신된 α-파라미터를 사용하여 양호한 출력을 생성하게 된다. 역LPC필터(111)의 출력은 고조파 부호화회로 등과 같은 정현분석부호화장치(114)의 DCT회로 등과 같은 직교변환회로(145)로 보내진다.

LPC분석/양자화장치(113)의 LPC분석회로(132)로부터의 α-파라미터는 가중필터계산회로(139)로 보내지며, 여기서 청감 가중용 데이터가 구해진다. 이러한 가중데이터는 청감가중 벡터양자화기(116), 제 2부호화장치(120)의 청감가중필터(125) 및 청감가중 합성필터(122)로 보내진다.

고조파 부호화회로의 정현분석부호화장치(114)는 고조파 부호화방법으로 역LPC필터(111)의 출력을 분석한다. 즉, 음조검출, 각 고조파의 진폭(Am)의 계산 및 유성음(V)/무성음(UV)판별이 수행되며, 음조와 함께 변환되는 진폭(Am)의 수 또는 각 고조파의 엔벌로프는 차원변환에 의해 일정하게 만들어진다.

도 3에 나타낸 정현분석부호화장치(114)의 도시예에서는, 평범한 고조파부호화를 사용하고 있다.

특히, 다대역여기(勵起)(MBE)부호화에 있어서, 유성음 및 무성음부가 주파수영역 또는 동일시점(동일한 블록이나 프레임)의 대역내에 존재하는 모델이라 가정한다. 다른 고조파부호화기술에 있어서는, 1블록 또는 1프레임내의 음성이 유성음인지 무성음인지를 훌륭하게 판정한다. 다음의 설명에 있어서, MBE부호화가 관련되는 한, 전체의 대역이 UV이면 주어진 프레임이 UV가 되도록 판정된다.

도 3의 정현분석부호화장치(114)의 개루프 음조 탐색장치(141) 및 제로크로싱계수기(142)는 입력단자(101)로부터의 입력음성신호와, 고역통과필터(HPF)(109)로부터의 신호를 각각 인가받는다. 정현분석부호화장치(114)의 직교변환회로(145)는 역LPC필터(111)로부터 LPC잔차 또는 선형예측잔차를 인가받는다. 개루프 음조탐색 장치(141)는 입력신호의 LPC잔차로 하여금 개루프에 의해 대략적인 음조탐색이 수행되도록 한다. 추출된 대충의 음조데이터는 후술하는 바와같이 폐루프에 의해 정밀음조탐색장치(146)로 보내진다. 개루프 음조 탐색장치(141)로부터는, 대충의 음조데이터와 함께 자기상관의 LPC잔차의 최대치를 표준화하여 얻어진 표준화된 자기상관(r)(p)의 최대치가 대충의 음조데이터와 함께 도출되어서 V/UV판별부(115)로 보내지게 된다.

직교변환회로(145)는 이산푸리에변환(DFT)등의 직교변환을 수행하여 시간축상의 LPC잔차를 주파수축상의 스펙트럼진폭데이터로 변환한다. 직교변환회로(145)의 출력은 정밀음조탐색장치(146) 및 스펙트럼평가부(148)로 보내져서 스펙트럼 진폭 또는 엔벌로프를 평가하게 된다.

정밀음조탐색장치(146)는 오픈루프음조탐색장치(141)에 의해 추출된 비교적 거친 음조데이터와 직교변환부(145)에 의한 DFT에 의해 얻어진 주파수영역데이터를 인가받는다. 정밀음조탐색장치(146)는 거친 음조값데이터에 대해 중심에 있는 ±몇몇 샘플에 의한 음조데이터를 0.2 0.5의 속도로 회전시켜서 결국 최적의 10진점(부동점)을 갖는 정밀음조데이터의 값에 다다르게 한다. 분석/합성방법은 전력스펙트럼이 본래의 음성의 전력스펙트럼에 밀접하게 되도록 음조를 선택하기 위한 정밀탐색기술로써 사용된다. 폐루프 정밀탐색부(146)로부터의 음조데이터는 스위치 (118)를 통해서 출력단자(104)로 보내진다.

스펙트럼 평가부(148)에 있어서, 각 고조파의 진폭 및 그 고조파의 합으로써의 스펙트럼엔벌로프는 스펙트럼진폭 및 LPC잔차의 직교변환출력으로써의 음조를 기초로 하여 평가되어 정밀음조탐색부(146), V/UV판별부(115) 및 가중벡터양자화기 (116)로 보내진다.

V/UV판별부(115)는 직교변환회로(145)의 출력, 정밀음조탐색부(146)로부터의 최적의 음조, 스펙트럼평가부(148)로부터의 스펙트럼진폭데이터, 개루프음조탐색부 (141)로부터의 표준화된 자기상관(r)(p)의 최대치, 및 제로크로싱계수기(142)의 제로크로싱카운트값을 기초로 하여 프레임의 V/UV를 판별한다. 또한, MBE용 기저대역의 V/UV판별의 경계부도 V/UV판별에 대한 조건으로 사용될 수 있다. V/UV판별부 (115)의 판별출력은 출력단자(105)에서 도출된다.

스펙트럼평가부(148)의 출력부 또는 벡터양자화기(116)의 입력부는 데이터수 변환부(일종의 샘플링속도변환을 수행하는 부분)를 부설하고 있다. 데이터수변환부는 주파수축상에 분할된 대역수와 데이터수가 음조와 다른점을 고려하여 엔벌로프의 진폭데이터(Am)를 설정하는데 사용된다. 즉, 유효대역이 3400kHz에 다다르면, 유효대역은 음조에 따라서 8 63대역으로 분할될 수 있다. 대역에서 대역까지에서 얻어진 진폭데이터(Am)의 mMx + 1의 수는 8 63의 범위로 변화된다. 따라서, 데이터수변환부는 가변수(mMx + 1)의 진폭데이터를 44데이터 등의 소정수(M)의 데이터로 변환한다.

스펙트럼평가장치(148)의 출력부 또는 벡터양자화기(116)의 입력부에 부설된 데이터수 변환장치로부터의, 44 등의, 소정수(M)의 진폭데이터 또는 엔벌로프데이터는 가중 벡터양자화를 수행함으로써, 벡터양자화장치(116)에 의해, 소정수의 데이터가 단위로써, 44데이터 등이 되는 식으로 수집된다. 이러한 가중은 가중필터계산회로(139)의 출력에 의해 인가된다. 벡터양자화기(116)로부터의 엔벌로프의 지수 는 출력단자(103)에서 스위치(117)에 의해 도출된다. 가중 벡터양자화에 앞서, 소정의 데이터수로 구성된 벡터를 위해 적절한 누설계수를 사용하여 프레임간의 차를 구하는 것이 상책이다.

제 2부호화장치(120)를 설명한다. 제 2부호화장치(120)는 소위 CELP부호화구조를 가지며 특히 입력음성신호의 무성음부를 CELP부호화하는데 사용된다. 입력음성신호의 무성음부에 대한 CELP부호화구조에 있어서, 잡음코드북 또는 소위 확률적코드북의 대표값출력으로써 무성음사운드의 LPC잔차에 상응하는 잡음출력은 이득제어회로(126)를 통해 가중합성필터(122)로 보내진다. 가중합성필터(122)는 입력잡음을 LPC합성하여 생성된 가중 무성음신호를 감산기(123)로 보낸다. 감산기(123)에는 고역통과필터(HPF)(109)를 통해서 입력단자(101)로부터 인가되어 가중필터(125)에 의해 가중된 신호가 인가된다. 그 신호와 합성필터(122)로부터의 신호와의 차이 또는 오차가 도출된다. 한편, 가중합성필터의 제로입력응답은 가중합성필터(125)의 출력에서 사전에 감산된다. 이러한 오차는 거리계산을 위한 거

리계산회로(124)로 인가된다. 오차를 최소화하게 될 대표적인 벡터값은 잡음코드북(121)에서 탐색된다. 이상은 분석/합성방법을 사용하는 개루프탐색을 차례로 사용하여 시간영역파형의 벡터양자화에 대한 요약이다.
CELP부호화구조를 이용하는 제 2엔코더(120)로부터의 무성음(UV)부에 대한 데이터로써, 잡음코드북(121)으로부터는 코드북에 대한 형상지수가, 이득회로(126)으로부터는 코드북에 대한 이득지수가 도출된다. 잡음코드북(121)으로부터의 UV데이터인 형상지수와, 이득회로(126)으로부터의 UV데이터인 이득지수는 스위치(127g) 를 통해 입력단자(107g)로 보내진다.
이러한 스위치(127s, 127g) 및 스위치(117, 118)는 V/UV판별부(115)로부터의 V/UV판별결과에 따라서 턴온 및 턴오프된다. 특히, 현재 송신되는 프레임의 음성신호에 대한 V/UV판별결과가 유성음(V)을 나타내는 경우에는 스위치(117, 118)가 턴온되는 한편, 현재 송신되는 프레임의 음성신호가 무성음(UV)을 나타내는 경우에는 스위치(127s, 127g)가 턴온된다.
부호화장치(2)에 의해 출력되는 부호화된 파라미터는 주기변형장치(3)에 공급된다. 주기변형장치(3)는 부호화된 파라미터의 출력주기를 시간축 압축/팽창으로 변형시킨다. 주기변형장치(3)에 의해 변형된 주기로 출력되는 그 부호화된 파라미터는 복호화장치(4)에 보내진다.
복호화장치(4)는 일 예로서 주기변형장치(3)에 의해 시간축을 따라 압축된 부호화된 파라미터를 보간하여 소정 프레임의 시점과 연합되어 있는 변형된 부호화된 파라미터를 생성하는 파라미터변형장치(5)와, 이 변형된 부호화된 파라미터에 의거하여 유성음 신호부분과 무성음 신호부분을 합성하는 음성합성장치(6)를 포함한다.
도 4와 도 5를 참조하여, 복호화장치(4)를 설명한다. 도 4에서, 주기변형장치(3)로부터의 선형스펙트럼쌍(LSPs)의 양자화된 출력데이터로서의 코드북지수데이터가 입력단자(202)에 공급된다. 양자화된 엔벌로프데이터로서 주기변형장치(3)의 출력, 즉 지수데이터와, 음조데이터 및 V/UV판별출력데이터가 각각 입력단자(203, 204, 205)에 입력된다. 또한 무성음부분에 대한 데이터로서 주기변형장치(3)로부터 의 지수데이터가 입력단자(207)에 입력된다.
양자화된 엔벌로프출력으로서 입력단자(203)로부터의 지수데이터는 LPC잔차의 스펙트럼엔벌로프를 구하기 위한 벡터양자화를 위하여 역벡터양자화기(212)에 보내진다. 유성음성합성장치(211)에 보내지기 전에 LPC잔차의 스펙트럼엔벌로프는 도 4에서 화살표(P1)에 의해 지시되는 지점 근처에서 파라미터변형장치(5)에 의해 일시적으로 취하여져 후술하는 바와같은 파라미터변형이 이루어진다. 이 지수데이터는 그다음 유성음성합성장치(211)에 보내진다.
유성음성합성장치(211)는 정현합성에 의해 유성음성신호부분의 LPC잔차를 합성한다. 각각 입력단자(204, 205)에 입력되고 도 4에서 점(P2, P3)에서 파라미터변형장치(5)에 의해 일시적으로 추출되어 파라미터변형되는 음조와 V/UV판별데이터는 동일하게 유성음성합성장치(211)에 공급된다. 유성음성합성장치(211)로부터의 유성음성의 LPC잔차가 LPC합성필터(214)에 보내진다.
입력단자(207)로부터의 UV데이터의 지수데이터는 무성음성합성장치(220)에 보내진다. UV데이터의 지수데이터는 잡음 코드북을 참고로 가지는 것에 의해 무성음성합성장치(220)에 의해 무성음성부분의 LPC잔차로 전환된다. UV데이터의 지수데이터는 도 4에서 점(P4)에 나타낸 바와같이 파라미터변형장치(5)에 의해 무성음성합성장치(220)로부터 일시적으로 추출되어 파라미터변형된다. 이렇게 파라미터변형된 LPC잔차는 또한 LPC합성필터(214)로 보내진다.
LPC합성필터(214)는 유성음성신호부분의 LPC잔차와 무성음성신호부분의 LPC잔차에 대한 독립적인 LPC합성을 수행한다. 선택적으로, 함께 합하여진 유성음성신 호부분의 LPC잔차와 무성음성부분의 LPC잔차에 대하여 LPC합성을 수행할 수도 있다.
입력단자(202)로부터의 LSP지수데이터는 LPC파라미터재생장치(213)에 보내진다. 비록 LPC의 α-파라미터는 궁극적으로는 LPC파라미터재생장치(213)에 의해 생성되지만, 화살표(P5)에 의해 지적되듯이 LSP의 역 벡터 양자화된 데이터가 파라미터변형장치(5)의해 도중에서 도출되어 파라미터변형된다.
이렇게 파라미터변형에 의해 처리된 탈양자화된 데이터는 이 LPC파라미터재생장치(213)로 되돌아가 LPC보간처리된다. 탈양자화된 데이터는 그다음 LPC의 α-파라미터로 변화되어 LPC합성필터(214)에 공급된다. LPC합성필터(214)에 의한 LPC합성에 의해 구해진 음성신호는 출력단자(201)에서 추출된다. 도 4에 도시된 음성합성장치(6)는 상술한 바와같이 파라미터변형장치(5)에 의해 계산된 변형된 부호화된 파라미터를 수신하여, 그 합성된 음성을 출력한다. 음성합성장치의 실제구조가 도 4에 해당하는 부품과 성분에 대하여는 같은 참조번호가 부여된 도 5에 도시된다.
도 5를 참조하면, 입력단자(202)에 입력되는 LSP지수데이터는 LPC파라미터 재생장치(213)내 LSPs를 위한 역벡터양자화기(231)에 보내져 LSPs(선스렉트럼쌍)로 역벡터양자화되고 이것은 파라미터변형장치(5)에 공급된다.
입력단자로부터의 스펙트럼엔벌로프의 벡터-양자화된 지수데이터(Am)는 역벡터양자화를 위하여 역벡터양자화기(212)에 보내지고 스펙트럼엔버로프의 데이터로 변하여 파라미터변형장치(5)로 보내진다.
이 입력단자(204, 205)로부터의 음조데이터와 V/UV판별데이터는 또한 파라미터변형장치(5)에 보내진다.
도 5의 입력단자(207s, 207g)에는 UV데이터로서 형상지수데이터와 이득지수데이터가 주기변형장치(3)를 통하여 도 3의 출력단자들(107s, 107g)로부터 공급된다. 그다음 형상지수데이터와 이득지수데이터는 무성음성합성장치(220)에 공급된다. 단자(207s)로부터의 형상지수데이터와 단자(207g)로부터의 이득지수데이터는 각각 무성음성합성장치(220)의 잡음코드북(221) 및 이득회로(222)에 공급된다. 잡음코드북(221)으로부터 독출되는 대표값출력은 무성음의 LPC잔차에 해당하는 잡음신호성분이고, 이것은 이득회로(222)내 소정 이득의 진폭이 된다. 그 결과신호는 파라미터변형장치(5)에 공급된다.
파라미터변형장치(5)는 부호화장치(2)에 의해 출력되고 주기변형장치(3)에 의해 출력주기가 변형된 부호화된 파라미터를 보간처리하여 변형된 부호화된 파라미터를 생성하고 이것은 음성합성장치(6)에 공급된다. 주기변형장치(3)는 부호화된 파라미터의 속도변형을 한다. 이것은 디코더 출력후 속도변형의 조작을 제거하여 본 음성신호 재생장치(1)가 유사한 알고리즘과는 다른 고정레이트(rate)를 다룰 수 있게한다.
도 6 및 도 8의 흐름도를 참조하여 주기변형장치(3)와 파라미터변형장치(5)의 작동을 설명한다.

도 6의 단계(S1)에서, 주기변형장치(3)는 LSPs, 음조, 유성음/무성음(V/UV), 스펙트럼엔벌로프(Am) 또는 LPC잔차와 같은 부호화된 파라미터들을 수신한다. LSPs, 음조, 유성음/무성음(V/UV), 스펙트럼엔벌로프(Am) 및 LPC잔차는 각각 $l_{sp}[n][p]$, $P_{ch}[n]$, $vu_{v}[n]$, $a_{m}[n][k]$ 및 $r_{es}[n][i][j]$로 나타내어진다.
궁극적으로 파라미터 변형장치(5)에 의해 계산되는 변형된 부호화된 파라미터들은 $mod\_l_{sp}[m][p]$, $mod\_P_{ch}[m]$, $mod\_vu_{v}[m]$, $mod\_a_{m}[m][k]$ 및 $mod_r_{es}[m][i][j]$로 나타내어지고, 여기서 k와 p는 각각 고조파 및 LSP 차수의 수를 의미한다. n과 m의 각각은 시간축변환의 이전과 이후의 시간-영역 지수데이터에 해당하는 프레임수를 의미한다. 한편, n과 m의 각각은 20msec의 간격을 가지는 프레임의 지수를 의미하고 i와 j는 각각 서브프레임수와 샘플수를 의미한다.
단계(S2)에서 도시하는 바와같이, 주기변형장치(3)는 그다음 초기 시간지속을 N_{1}으로 설정하고, 변화후 그 시간지속을 나타내는 프레임의 수를 N_{2}로 설정한다. 단계(S3)에 나타내는 바와같이, 주기변형장치는 그다음 음성(N_{2})으로 음성(N_{1})의 시간축압축으로 진행한다. 즉, 주기변형장치(3)에서 시간축압축율(spd)는 $0 \leq n < N_{1}$ 및 $0 \leq m < N_{2}$ 일 때 $spd = N_{2}/N_{1}$으로 된다.
그다음 파라미터변형장치(5)는 시간축변형후의 시간축의 지수에 해당하는 프레임의 개수에 상응하는 m을 2로 설정한다.
파라미터변형장치(5)는 그다음 두 프레임(fro, fr1)과 그 두 프레임(fro, fr1)사이의 차이(좌측 및 우측) 및 그 비율(m/spd)을 구한다.
파라미터(l_{sp}, P_{ch}, vu_{v}, a_{m} 및 r_{es})가 *으로 주어진다면, mod*[m]은 일반식
$mod\_*[m] = *[m/spd]$ ($0 \leq m < N_{2}$ 인 경우)
으로 나타내어질 수 있다. 그러나, m/spd는 정수가 아니기 때문에 m/spd에서 변형된 부호화된 파라미터는 두 프레임($f_{r0} = [m/spd]$ 및 $f_{r1} = f_{0} + 1$)으로부터 보간처리에 의해 생성된다.
프레임(f_{r0}), m/spd, 프레임(f_{r1}) 사이에는 도 7에 도시된 관계, 즉,
좌측 = $m/spd - f_{r0}$
우측 = $f_{r1} - m/spd$
이 유지된다.
도 7에서 m/spd에 대한 부호화된 파라미터 즉, 변형된 부호화된 파라미터들은 단계(S6)에서 도시된 바와같이 보간처리에 의해 구해질 수 있다.
그 변형된 부호화된 파라미터는 간단히
$mod_*[m] = *[f_{r0}] \times$우측 $+ *[f_{r1}] \times$좌측
에 의한 선형보간에 의해 구해진다.
그러나, 두 프레임(f_{r0}, f_{r1}) 사이의 보간처리로는, 만일 그 두 프레임이 V/UV에 대하여 다르다면, 즉 두 프레임중 하나는 V이고 다른 것은 UV라면, 상기 일반식은 사용되어 질 수 없다. 따라서, 파라미터변형장치(5)는 도 8의 단계(S11)에 의해 나타내어지는 바와같이, 두 프레임(f_{r0}, f_{r1})의 유성음(V) 또는 무성음(UV) 특성에 따라 부호화된 파라미터를 구하는 방법을 변화시킨다.
먼저, 두 프레임(f_{r0}, f_{r1})의 유성음(V) 또는 무성음(UV) 특성이 단계(S11)에서 도시되는 바와 같이 판별된다. 두 프레임(f_{r0}, f_{r1})이 모두 유성음(V)아라면, 단계는 (S12)로 진행하여, 모든 파라미터가 선형보간되어
$mod_P_{ch}[m] = P_{ch}[f_{r0}] \times$우측 $+ P_{ch}[f_{r1}] \times$좌측
$mod_a_{m}[m][k] = a_{m}[f_{r0}][k] \times$우측 $+ a_{m}[f_{r1}][k] \times$좌측
(여기서, $0 \leq k < 1$, L은 고조파의 최대가능넘버)
으로 나타내어진다. 고조파가 없는 경우에는 $a_{m}[n][k]$에 대하여 0이 삽입된다. 고조파 넘버가 프레임들(f_{r0}, f_{r1}) 사이에서 다른 경우에는, 빈자리에 0이 삽입된다. 대안적으로, 디코더측상의 다수의 데이터변환기를 통과하기 이전에 $0 \leq k < L$ (L = 43)과 같은 고정수가 사용되어질 수 있다.
$mod_l_{sp}[m][p] = l_{sp}[f_{r0}][p] \times$우측 $+ l_{sp}[f_{r1}][p] \times$좌측
여기서, $0 \leq p < P$, P는 LSPs의 차수를 나타내고 일반적으로 10이다.
$mod\_vu_{v}[m] = 1$
V/UV판별에서, 1과 0은 각각 유성음(V)과 무성음(UV)을 의미한다.
만일, 단계(S11)에서 두 프레임(f_{r0}, f_{r1})이 모두 유성음(V)이 아니라고 판단되면, 단계(S13)에서 두 프레임(f_{r0}, f_{r1})이 모두 무성음(UV)인가가 판단된다. 단계 (S13)에서의 판정결과가 그렇다이면, 즉, 두 프레임 모두 무성음이라면, 단계(S14) 에서 나타내는 바와 같이 보간처리장치(5)는 r_{es}의 이전과 이후의 80 샘플을 m/spd를 센터로 하고 p_{ch}를 최대값으로 하여 슬라이스한다.
그 결과, 단계(S14)에서 좌측<우측 이라면, m/spd를 센터로 하여 r_{es}의 이전과 이후의 80 샘플이 슬라이스되고 도 9a에 나타내는 바와같이 mod r_{es}에 삽입된다. 즉,

(j = 0; j < FRM × (1/2 - m/spd + f_{r0}); j++ {mod

$r_{es}[m][0][j] = r_{es}[f_{r0}][0][j + (m/spd - f_{r0}) \times FRM];$};

(j = FRM × (1/2 - m/spd + f_{r0}); j < FRM/2; j++) {mod

$r_{es}[m][0][j] = r^{es}[m][0][j] = r_{es}[f_{r0}][1][j - FRM \times (1/2 - m/spd +$

$f_{r0})];$};

(j = 0; j < FRM × (1/2 - m/spd + f_{r0}); j++) {mod

$r_{es}[m][1][j] = r_{es}[f_{r0}][1][j + (m/spd - f_{r0}) \times FRM];$};

(j = FRM × (1/2 - m/spd + f_{r0}); j = FRM/2; j++) {mod

$r_{es}[m][1][j] = r_{es}[f_{r0}][0][j + FRM \times (1/2 - m/spd + f_{r0})];$};

(여기서 FRM은 예를들어 10이다.)

한편, 단계(S14)에서 좌측≥우측 이라면, 보간처리장치(5)는 m/spd를 센터로 하여 r_{es} 의 이전과 이후의 80 샘플을 슬라이스하여 도 9b에 나타내는 바와같이 mod_r_{es} 를 생성한다.

단계(S13)에서의 조건이 충족되지 않으면, 처리는 단계(S15)로 진행하여 프레임(f_{r0})이 유성음(V)이고 프레임(f_{r1})이 무성음(UV)인지를 판별한다. 판별결과가 그렇다이면, 즉, 프레임(f_{r0})이 유성음(V)이고 프레임(f_{r1})이 무성음(UV)이라면, 처 리공정은 단계(S16)로 진행한다. 만일 판별결과가 아니다이면, 즉, 프레임(f_{r0})이 무성음(UV)이고 프레임(f_{r1})이 유성음(V)이라면, 처리공정은 단계(S17)로 진행한다.

단계(S15) 이후의 처리에서는 두 프레임(f_{r0}, f_{r1})이 V/UV에 대하여, 즉 유성음(V)에 대하여 무성음(UV)와 같이 다르다. 이것은 만일 V/UV에 대하여 다른 두 프레임(f_{r0}, f_{r1}) 사이에서 파라미터들이 보간처리되면, 보간결과는 무의미해진다는 사실을 고려하고 있다.

단계(S16)에서, 좌측(=m/spd-f_{r0})의 사이즈와 우측(,= f_{r1} - m/spd)의 사이즈가 서로 비교되어 프레임(f_{r0})의 사이즈가 m/spd에 근접하는지를 판별한다.

그 프레임(f_{r0})이 m/spd에 가까워진다면, 변환된 부호화된 파라미터가 프레임 (f_{r0})의 파라미터를 이용하여 설정되어, 단계(S18)에 도시된 바와같이,

$mod\_p_{ch}[m] = p_{th}[f_{r0}]$

$mod\_a_m[m][k] = a_m[f_{r0}][k] \quad (0 \le k < L)$

$mod\_l_{sp}[m][p] = l_{sp}[f_{r0}][p] \quad (0 \le p < I)$

$mod\_vu_u[m] = 1$

이다.

단계(S16)에서의 판정결과가 아니오 라면, 좌측≥우측, 따라서 프레임(f_{r1})은 가까워지고, 따라서 처리과정은 음조를 최대화하기 위하여 단계(S19)로 이전된다. 또한, 도 9c에 도시하는 바와같이, 프레임(f_{r1})의 r_{es} 가 직접적으로 사용되고, mod_r_{es} 로서 설정된다. 즉, mod_r_{es} [m][i][j] = r_{es} f_{r1} [i][j]이다. 그 이유는 유성음의 프레임(f_{r0})에 대하여 LPC잔차(r_{es})가 전송되지 않기 때문이다.

단계(S17)에서, 단계(S16)와 동일한 판정이 단계(S15)에서 주어진 판정에 의거하여 그 두 프레임(f_{r0}, f_{r1})이 각각 무성음(UV)과 유성음(V)이라고 주어진다. 즉, 좌측(=m/spd-f_{r0})의 사이즈와 우측(f_{r1} - m/spd)의 사이즈가 서로 비교되어 프레임 (f_{r0})의 사이즈가 m/spd에 근접하는지를 판별한다.

프레임(f_{r0})이 가까워지면, 처리공정은 음조를 최대화하기 위해 단계(S18)로 이동한다. 또한, 프레임(f_{r0})의 r_{es} 가 직접적으로 사용되고, mod_r_{es} 로서 설정된다. 즉, mod_r_{es} [m][i][j] = r_{es} f_{r0} [i][j]이다. 그 이유는 유성음의 프레임(f_{r1})에 대하여 LPC잔차(r_{es})가 전송되지 않기 때문이다.

단계(S17)에서의 판정결과가 아니오 라면, 좌측≥우측, 따라서 프레임(f_{r0})은 m/spd에 가까워지고, 따라서 처리과정은 단계(S21)로 진행하고 변형된 부호화된 파라미터가 프레임(f_{r1})의 파라미터를 이용하여 설정되고 따라서,

$mod\_p_{ch}[m] = p_{ch}[f_{r1}]$

$mod\_a_m[m][k] = a_m[f_{r1}][k] \quad (0 \le k < L)$

$mod\_l_{sp}[m][p] = l_{sp}[f_{r1}][p] \quad (0 \le p < I)$

$mod\_vu_u[m] = 1$

이다.

이렇게하여, 보간처리장치(5)는 도 8에 상세히 나타내는 바와같이, 두 프레임(f_{r0}, f_{r1})의 V/UV특성에 따라 다른 처리를 도 6의 단계(S6)의 보간처리를 위하여 제공한다. 단계(S6)에서의 보간처리의 종료이후에는, 처리공저이 단계(S6)로 이동하여 m의 값을 증가시킨다. 단계(S5, S6)의 처리는 m의 값이 N_2 와 같을때까지 반복된다.

주기변형장치(3)와 파라미터변형장치(5)의 작동이 도 10을 참조하여 총괄적으로 설명된다. 도 10a를 참조하면, 부호화장치(2)에 의해 매 20msec의 주기로 추출되는 부호화 파라미터의 주기는 도 10a에 도시되는 바와같이 주기변형장치(3)에 의해 시간축 압축으로 15msec로 변형된다. 두 프레임(f_{r0}, f_{r1})의 V/UV의 상태에 응답하는 그 보간처리조작에 의해, 파라미터변형장치(5)는 도 10c에 도시되는 바와같이 매 20msec 마다 변형된 부호화된 파라미터를 계산한다

주기변형장치(3)와 파라미터변형장치(5)에 의한 작동은 그 순서가 반대로 될 수도 있다. 즉, 도 11a에 도시된 부호화된 파라미터가 먼저 도 11b에 도시된 바와같이 보간처리된 후, 도 11c에 도시된 바와같이 압축되어 변형된 부호화된 파라미터를 계산하여도 좋다.

도 5로 되돌아가서, 파라미터계산장치(5)에 의해 계산된 LSP데이터상의 변형된 부호화된 파라미터들($mod\ l_{sp}[m][p]$)은 LSP보간처리를 위하여 LSP보간처리회로 (232v, 232u)에 보내진다. 결과데이터는 LSP에 의해 변환되어 선형예측코딩(LPC)를 위한 α-파라미터로 변환되기 위하여 α변환회로(234v, 234uv)로 보내지고, 그 LPC 는 LPC합성필터(214)로 보내진다. LSP보간처리회로(232v)와 α변환회로(234v)에 대한 LSP는 유성음(V)신호부를 위하여 사용되고, LSP보간처리회로(232u)와 α변환회로(234u)에 대한 LSP는 무성음(UV)신호부를 위하여 사용된다. LPC합성필터(214)는 유성음부를 위한 LPC합성필터(236)와 무성음부를 위한 LPC합성필터(237)로 구성되어 있다. 즉, 그렇지 않다면 발생할 수 있는 유성음부에서 무성음부로의 전이영역이나 무성음부에서 유성음부로의 전이영역에서 전적으로 다른 특성의 LSP보간처리에 의해 생성될 수 있는 악효과를 방지하기 위하여, LPC계수보간처리는 유성음부와 무성음부에 대하여 독립적으로 수행된다.

파라미터변형장치(5)에 의해 구해지는 바와같이, 스펙트럼 엔벌로프데이터 ($mod\_a_m[m][k]$)에 관한 변형된 부호화된 파라미터는 유성음성합성장치(211)의 정현합성회로(215)에 보내진다. 이 유성음성합성장치(211)는 또한 파라미터변형장치(5)에 의해 계산되는 바와같이, 음조에 관한 변형된 부호화된 파라미터($mod\_p_{ch}[m]$)와 V/UV결정데이터에 관한 변형된 부호화된 파라미터($mod\_vu_v[m]$)을 공급받는다. 이 정현합성회로(215)로부터 도 3의 LPC역필터(111)의 출력에 해당하는 LPC잔차데이터가 추출되어 가산기(218)에 보내진다.

파라미터변형장치(5)에 의해 구해지는 바와같이, 스펙트럼 엔벌로프 데이터에 관한 변형된 부호화된 파라미터($mod\_a_m[m][k]$), 음조에 관한 변형된 부호화된 파라미터($mod\_p_{ch}[m]$), 및 V/UV결정데이터에 관한 변형된 부호화된 파라미터 ($mod\_vu_v[m]$)가 잡음합성회로(216)에 보내져 유성음(V)부에 대하여 잡음을 부가한다. 잡음합성회로(216)의 출력은 오버랩과 가산회로(217)를 거쳐 가산기(218)에 보내진다. 상술하면, 음조스펙트럼 엔벌로프진폭, 프레임내 최대진폭 또는 잔차신호레벨과 같은 부호화된 음성데이터로부터 유도된 파라미터들을 고려한 잡음가, 여기인 유성음성의 LPC합성필터의 입력이 정현합성에 의해 생성된다면, 남성음성과 같은 로우-음조음성에서 '스터프된(stuffed)'감흥이 생성되고, 한편, 음질은 V와 UV음성부 사이에서 갑자기 변화하고 따라서 부자연스러운 느낌을 생성한다는 것을 고려하여, LPC합성필터입력의 LPC잔차신호의 즉 여기(excitation)인 유성음부분에 부가된다.

가산기(218)의 합출력은 시간파형데이터가 LPC합성에 의해 생성되는 유성음성을 위한 합성필터(236)에 보내진다. 또한 그 결과의 시간파형데이터는 포스트 (post)필터(238v)에 의해 필터링되어 가산기(239)에 공급된다.

상술한 바와같이, LPC합성필터(214)는 V(236)을 위한 합성필터와 UV(237)를 위한 합성필터로 구분된다는 것을 주목한다. 합성필터가 이와같이 구분되어 있지 않으면, 즉, LSPs가 V, UV신호부분 사이의 구별없이 20샘플마다 또는 2.5msec마다 계속하여 보간처리된다면, 전적으로 다른 특성의 LSPs가 V에서 UV로 UV에서 U로 보간처리되어 외계의 음을 생성한다. 이러한 악영향을 방지하기 위하여, LPC합성필터는 V를 위한 필터와 UV를 위한 필터와 구분되어 있고 V, UV에 대하여 독립적으로 LPC계수를 보간처리한다.

파라미터변형장치(5)에 의해 계산되는 바와같이, LPC잔차에 대한 변형된 부호화된 파라미터들($mod\ r_{es}[m][i][j]$)는 윈도윙회로(windowing circuit)(223)에 보내져서 유성음성부분과의 접합부분을 자연스럽게 하기위한 윈도윙처리를 한다.

윈도윙회로(223)의 출력은 LPC합성필터(214)의 UV를 위한 합성필터(237)에 유성음성합성장치(220)의 출력으로서 보내진다. 합성필터(237)는 그 데이터에 대하여 LPC합성을 수행하여 무성음부에 대한 시간파형데이터를 제공하고, 이것은 무성음(238u)을 위한 포스트필터에 의해 필터링되어 가산기(239)에 공급된다.

가산기(239)는 유성음부를 위한 포스트필터(238v)로부터의 유성음부의 시간파형신호를 무성음부를 위한 포스트필터(238u)로부터의 무성음부의 시간파형데이터에 가산하고, 그 결과데이터를 출력단자(201)에서 출력한다.

본 음성신호 재생장치(1)에 의하면, 변형된 부호화된 파라미터들의 배열 $mod\_*[m]$ ($0 \le m < N_2$)이 고유의 배열 $*[n]$ ($0 \le n < N_1$) 대신에 이 방법으로 복호화된다. 복호화도중의 프레임간격은 종래와같이 20msec와같이 고정될 수 있다. 이러한 경우, 시간축압축과 그 결과로서의 재생레이트의 속도향상이 $N_2 < N_1$에 대하여 실현될 수 있는 한편, $N_2 > N_1$에 대하여는 시간축팽창과 그결과의 재생레이트의 속도다운이 실현될 수 있다.

본 시스템에 의하면, 궁극적으로 구하여진 파라미터열이 20msec의 고유간격으로 배열되어 복호화되고 따라서, 쉽게 서택적인 소고향상이 실현될 수 있다. 또한, 속도향상과 속도다운이 어떤 구별없이 같은 처리장치에 의해 실현될 수 있다.

따라서, 고체-상태 기록의 내용이 실시간속도의 2배속도로 재생될 수 있다. 재생속도가 증가하여도 음조나 음소는 불변하기 때문에 재생속도가 월등히 향상된 재생에도 불구하고 기록내용이 분명하게 구별될 수 있다.

$N_2 < N_1$, 즉, 재생속도가 낮아진다면, 무성음 프레임의 경우에 같은 LPC잔차들(r_{es})로부터 복수의 파라미터들(mod_r_{es})이 생성되기 때문에 재생음이 부자연스럽게 된다. 이러한 경우, 이러한 부자연스러움을 어느정도 제거하기 위하여 그 파라미터(mod_r_{es})에 적당한 양의 잡음이 부가될 수 있다. 잡음을 부가하는 대신에, 그 파라미터들(mod_r_{es})을 적합하게 생성된 가우스잡음(Gaussian noise)나 코드북으로부터 랜덤하게 선택된 여기벡터로 치환하는 것이 또한 채용될 수 있다.

상술한 음성신호 재생장치(1)에 의하면, 부호화장치(2)로부터의 부호화된 파라미터들의 출력주기의 시간축이 주기변형장치(3)에 의해 압축되어 재생속도를 향상시킨다. 그러나 프레임길이는 복호화장치(4)에 의해 변할 수 있도록 되어 있어서 재생속도를 제어할 수 있도록 한다.
이 경우, 프레임길이가 다양하게 되기 때문에, 복호화장치(4)의 파라미터변형장치(5)에 의한 파라미터생성 이전과 이후에 프레임수(n)는 변하지 않는다.
또한, 파라미터변형장치(5)는 주제의 프레임이 유성음인지 무성음인지에 무관하게, 파라미터들 $l_{sp}[n][p]$, $vu_v[n]$ 을 각각 $mod\_l_{sp}[n][p]$, $mod\_vu_v[n]$ 으로 변형시킨다.
만일 $mod_vu_v[n]$이 1 이라면, 즉, 주제의 프레임이 유성음(V)이라면, 파라미터들 $p_{ch}[n]$, $a_m[n][k]$는 각각 $mod\_p_{ch}[n]$, $mod_a_m[n][k]$로 변형된다.
만일 $mod_vu_v[n]$이 0 이라면, 즉, 주제의 프레임이 무성음(UV)이라면, 파라미터 $r_{es}[n][i][j]$는 $mod_r_{es}[n][i][j]$로 변형된다.
파라미터변형장치(5)는 $l_{sp}[n][p]$, $p_{ch}[n]$, $vu_v[n]$, 및 $a_m[n][k]$를 직접적으로 $mod\_l_{sp}[n][p]$, $mod\_p_{ch}[n]$, $mod\_vu_v[n]$, 및 $mod\_a_m[n][k]$ 으로 변형시킨다. 그러나, 파라미터변형장치는 속도(spd)에 따라 잔차신호 $r_{es}[n][i][j]$를 변화시킨다.
속도(spd)가 < 1.0이면, 즉, 속도가 빠르면, 원음의 잔차신호는 도 12에 도시되는 바와같이, 중간부분에서 슬라이스된다. 원래 프레임길이가 orgFrmL이면, (ó rgFrmL - frmL)/2 ≤j ≤(orgFrmL + frmL)/2가 원래프레임 $r_{es}[n][i]$에서 슬라이스되어 $mod_r_{es}[n][i]$를 만든다. 또한 원래프레임의 선단에서 슬라이싱이 되는 것도 가능하다.
속도(spd)가 > 1.0이면, 즉, 속도가 느리면, 원래프레임이 사용되고 잡음성분과 함께 부가된 원래프레임이 어떤 결손부분에 사용된다. 적당하게 생성된 잡음과 함께 부가되는 복호화된 여기벡터가 또한 사용될 수 있다. 가우시언잡음이 생성되어 여기벡터로서 사용되어, 같은 파형의 프레임의 연속에 의해 생성되는 이상한 느낌을 감소시키도록 할 수도 있다. 상기 잡음성분은 또한 원래프레임의 양단에 부 가될 수도 있다.
따라서, 프레임길이의 변화에 의해 속도제어를 변화시키도록 구성된 음성신호 재생장치(1)의 경우에서는, 음성합성장치(6)가 구비되어 LSP보간처리장치(232v, 232u), 정현합성장치(215), 및 윈도잉장치(223)가 시간축팽창에 의해 속도를 제어하기 위하여 다른 동작을 수행하도록 디자인되어 있다.
LSP보간처리장치(232v)는 주제의 프레임이 유성음(V)인 경우에는, frmL/p ≤20의 관계를 만족하는 가장 작은 정수(p)를 구한다. LSP보간처리장치(232u)는 주제의 프레임이 무성음(UV)인 경우에는, frmL/p≤80의 관계를 만족하는 가장 작은 정수(p)를 구한다. LSP보간처리를 위한 부프레임(subl[i][j])의 범위가 다음의 식에 의해 결정된다.
nint(frmL/p ×i) ≤j ≤nint(frmL/p ×(i+1) (0≤i<p-1)
상기 식에서, nint(x)는 소수 첫째자리에서 반올림하여 x에 가장 가까운 정수로 환원하는 함수이다. 유성음 및 무성음 모두에 대하여 frmL이 20이나 80 보다 적으면, p=1이다.
예를들어, i번째 서브프레임에 대하여, 서브프레임의 중앙이 frmL ×(2i + 1)/2p이기 때문에, 동시계류중인 우리의 일본국 출원번호 6-198451에 개시된 바와같이, LSP는 frmL ×(2p -2i -1)/(20:frmL×(2i+1)/2p의 속도로 보간처리된다.
선택적으로, 부프레임의 수는 고정될 수 있고, 각 부프레임의 LSPs가 항상 같은 비율로 보간될 수 있다. 정현합성장치(223)는 프레임길이(frmL)와 일치하도록 윈도 길이를 변형시킨다.
상술한 음성신호 재생장치에 의하면, 부호화된 파라미터의 출력주기가 시간축상에 팽창되어 있는 그 부호화된 파라미터가 주기변형장치(3)에 의해 변형되고, 파라미터변형장치(5)에 의해 음조나 음소의 변화없이 재생속도가 변화된다. 그러나 주기변형장치(3)를 생략하고 도 14에 도시된 복호화장치(8)의 데이터변환장치(270)의 수에 의해 부호화장치(2)로부터의 부호화된 데이터를 처리하여 음소의 변화없이 음조를 변화시키는 것이 가능하다. 도 14에서 도 4와 같은 부품과 성분들은 같은 참조번호로 나타내어진다.
복호화장치(8)의 기본 개념은 부호화장치(2)로부터 입력되는 부호화된 음성데이터의 고조파의 기본주파수와 소정대역내 진폭데이터의 수를 데이터변환수단에 의해 작동되는 데이터변환장치(270)의 수에 의해 변환시켜 음소의 변화없이 단지 음조만을 변화시키는 것이다. 데이터변환장치(270)의 수는 각 입력 고조파에서 스펙트럼성분의 사이즈를 특정하는 데이터의 수를 변화시킴으로써 음조를 변화시킨다.
도 14를 참조하면, 도 2와 도 3의 출력단자(102)의 출력에 해당하는 LSPs의 벡터 양자화된 출력 또는 코드북지수가 입력단자(2020)에 공급된다.
LSP지수데이터는 LPC파라미터재생장치(213)의 역벡터양자화기(321)에 보내져 선스펙트럼쌍(LSPs)으로 역벡터양자화된다. 이 LSPs는 LSP보간처리회로(323, 233)로 보내져서 보간처리되고 그후 LSP를 α변환회로(234, 235)로 공급하여 선형예측코드의 α-파라미터로 변환한다. 이들 α-파라미터들은 LPC합성필터(214)로 보내진다. LSP보간처리회로(232)와 α변환회로(234)에 대한 LSP가 유성음(V)신호부에 대 하여 사용되고, LSP보간처리회로(233)와 α변환회로(235)에 대한 LSP가 무성음(UV)신호부에 대하여 사용된다. LPC합성필터(214)는 유성음부를 위한 LPC합성필터(236)와 무성음부를 위한 LPC합성필터(237)와로 구성된다. 즉, 유성음부와 무성음부에 대하여 독립적으로 LPC계수 보간처리가 수행되어, 그렇지 않으면 발생할 유성음부로부터 무성음부로의 전이영역과 무성음부에서 유성음부로의 전이영역에서 전적으로 다른 특성의 LSPs보간에 의해 생성되는 악영향을 방지한다.
도 14의 입력단자(203)에는 도 2 및 도 3에서 도시된 엔코더의 단자(103)의 출력에 해당하는 스펙트럼 엔벌로프(Am)의 가중된 벡터양자화된 코드지수데이터가 공급된다. 입력단자(205)에는 도 2 및 도 3의 단자(105)로부터 V/UV결정데이터가 공급된다.
입력단자(203)로부터의 스펙트럼엔벌로프(Am)의 벡터양자화된 지수데이터는 역벡터양자화기(212)에 보내져 역벡터양자화된다. 역벡터양자화된 엔벌로프의 진폭데이터의 수는 예를들어 44의 소정값으로 고정된다. 기본적으로 데이

터의 수는 음조데이터에 해당하는 고조파의 수를 부여하기 위하여 변환된다. 본 실시예에서와 같이 음조를 변화시키는 것이 소망된다면, 역벡터양자화기(212)로부터의 엔버로프데이터는 데이터변환장치(270)의 수로 보내져서 예를들어 소망하는 음조값에 의존하는 보간에 의해 진폭데이터의 수를 변화시킨다.
데이터변환장치(270)의 수에는 또한 입력단자(204)로부터의 음조데이터가 공급되어 부호화시간에서 음조가 소망하는 음조로 변화되어 출력된다. 진폭데이터와 변형된 음조데이터는 유성음성합성장치(211) 정현합성회로(215)에 보내진다. 합성 회로(215)에 공급되는 진폭데이터의 수는 데이터변환장치(270)의 수로부터의 LPC잔차의 스펙트럼 엔벌로프의 변형된 음조에 해당한다.
데이터변환장치(270)의 수에 의해 LPC잔차의 스펙트럼 엔벌로프의 진폭데이터의 수를 변환시키는 보간처리방법은 여러가지가 있다. 예를들어, 주파수축상의 효과적인 대역블록의 진폭데이터를 그블록의 마지막 진폭데이터에서부터 그 블록내 제 1진폭데이터나 좌측단(제 1데이터)과 우측단(마지막데이터)에 걸친 더미데이터까지 보간처리하기 위한 더미데이터의 적당한 수가 그 블록의 진폭데이터에 첨부되어 N_F 에 대한 데이터의 수를 향상시킨다. 그다음, 진폭데이터의 Os-튜플 수는 옥타튜플 오버샘플링(octatuple oversampling)과 같은 대역제한형 Os-튜플 오버샘플링에 의해 구하여진다. 진폭데이터의 Os-튜플 수((mMx + 1)×데이터의Os 수)는 선형보간에 의해 더욱 팽창하여 2048과 같은 더 큰 N_M 의 수로 된다. 이 N_M 수 데이터는 데시메이션에 의해 소정수의 M(44 등)으로 변환되고 그다음 소정수의 데이터상에 벡터양자화가 수행된다.
데이터변환장치(270)의 수에서의 예시적인 작동으로서, 음조지연 L에서 Fx 까지에 대하여 주파수 $F_0 = f_s/L$인 경우가 설명된다(fs는 fs = 8kHz = 8000Hz와 같은 샘플링주파수).
이 경우, 음조주파수 F_0 = 8000/L, 4000 Hz로 설정된 n=L/2 고조파가 있다. 통상의 3400 Hz의 음성영역에서, 고조파의 개수는 (L/2)×(3400/4000). 이것은 벡터양자화 이전에 상기 데이터수 변환이나 차원변환에 의해 예를들어 44로 변환된 다. 단순히 음조를 변화시키고자 하는 경우에는 양자화를 수행할 필요가 없다.
역벡터양자화 이후에는, 고조파의 수, 44는 데이터변환장치(270)의 개수에 의한 차원변환에 의해 소망하는 개수, 즉, 소망하는 음조주파수(Fx)로 변화될 수 있다. 음조주파수 Fx(Hz)에 해당하는 음조지연(Lx)는 Lx=8000/Fx이고 따라서 3400 Hz로 설정된 고조파의 개수는 (Lx/2)×(3400/4000)=(4000/Fx)×(3400/4000)=3400/ Fx, 즉, 3400/Fx이다. 즉, 이것은 데이터변환장치(270)의 개수에서 데이터변환의 개수나 차원변환에 의해 44에서 3400/Fx으로 변환을 수행하는데 충분하다.
스펙트럼데이터의 벡터양자화 이전에 부호화시에 프레임-프레임 사이의 차이가 발견되면, 그 프레임-프레임 사이의 차이는 역벡터양자화이후 복호화된다. 데이터변환의 수는 그다음 스펙트럼 엔벌로프 데이터를 생성하기 위하여 수행된다.
정현합성회로(215)에는 데이터변환장치(270)의 수로부터 LPC잔차의 스펙트럼 엔벌로프 진폭데이터와 음조데이터 뿐만아니라, 입력단자(205)로부터 V/UV결정데이터가 공급된다. 정현합성회로(215)로부터, LPC잔차데이터가 취하여져 가산기(218)로 보내진다.
역벡터양자화기(212)로부터의 엔버로프데이터, 입력단자(205)로부터의 음조데이터 및 입력단자(205)로부터의 V/UV결정데이터가 잡음가산회로(216)에 보내져서 유성음(V)에 대하여 잡음이 부가된다. 상술하면, 음조 스펙트럼 엔벌로프 진폭, 그 프레임내 최대진폭, 또는 잔차신호레벨과 같은 부호화된 음성데이터로부터 도출되는 파라미터들을 고려하는 잡음이 LPC합성필터입력, 즉 여기를 위하여 LPC잔차신호의 유성음부에 가산되며, 이때 유성음성의 LPC합성필터에로의 입력이 정현합성에 의해 생성된다면, 남성음성과 같은 저음에서 '스터프된' 느낌이 생성되는 한편, V와 UV음성부분의 사에에서 음질이 급격히 변하고 따라서 부자연스러운 느낌을 생성된다는 것을 고려한다.
가산기(218)의 출력은 유성음성을 위한 합성필터(236)로 보내져서, LPC합성에 의해 시간파형데이터가 생성된다. 또한, 결과의 시간파형데이터는 유성음 데이터를 위한 포스트필터(238v)에 의해 필터링되고 그후 가산기(239)에 공급된다.
도 14의 입력단자(207s, 207g)에는 형상지수데이터와 이득지수데이터가 UV데이터로서 도 3의 출력단자(107s, 107g)로부터 주기변형장치(3)를 통하여 공급된다. 형상지수데이터와 이득지수데이터는 그 후 무성음성합성장치(220)에 보내진다. 단자(207s)로부터의 형상지수데이터와 단자(207g)로부터의 이득지수데이터는 각각 무성음성합성장치(220)의 잡음코드북(221)과 이득회로(222)에 공급된다. 잡음코드북 (221)으로부터 독출되는 대표값출력은 무성음성의 LPC잔차에 해당하는 잡음신호성분이고 이는 이득회로(222)에서 소정의 이득의 진폭이 된다. 소정의 이득진폭의 대표값출력은 윈도잉회로(223)에 보내져서 유성음신호부와 연결부를 부드럽게 하는 위도잉을 수행한다.
위도잉회로(223)의 출력은 무성음성합성장치(220)의 출력으로서 LPC합성필터 (214)의 무성음(UV)부를 위한 합성필터(237)에 보내진다. 윈도잉회로(223)의 출력은 LPV합성에 의해 합성필터(237)에 의해 처리되어 무성음성신호부의 시간-영역 파형신호를 만들고, 이는 그다음 무성음성부(238u)를 위한 포스트필터에 의해 필터링되고 그다음 가산기(239)에 보내진다.
가산기(239)는 유성음성을 위한 포스트필터(238v)로부터의 유성음성신호부를 위한 시간-영역 파형신호와 무성음성을 위한 포스트필터(238u)로부터의 무성음성신호부를 위한 시간-영역 파형데이터에 합산한다. 그 결과의 합계신호가 출력단자 (201)에서 출력된다.
상기에서, 스펙트럼 엔벌로프의 형상의 변화없이 고조파의 개수를 변화시킴으로써 음성의 음조변화없이 음조가 변할 수 있다. 따라서, 음성패턴의 부호화된 데이터, 즉, 부호화된 비트스트림(bitstream)이 이용가능하면, 그 음조는 선택적으로 합성을 위하여 변화할 수 있다.
도 15를 참조하면, 도 2 및 도 3의 엔코더에 의한 부호화에 의해 구해진 부호화된 비트스트림이나 부호화된 데이터가 부호화된 데이터출력장치(301)에 의해 출력된다. 이들 데이터중 적어도 음조데이터와 스펙트럼 엔벌로프 데이터는 데이터변환장치(302)를 통하여 파형합성장치(303)에 보내진다. 유성음/무성음(V/UV)결정데이터와 같은 음조변환과

무관한 데이터는 직접 파형합성장치(303)에 보내진다.
파형합성장치(303)는 스펙트럼 엔벌로프 데이터나 음조데이터에 의거하여 음성파형을 합성한다. 물론, 도 4 또는 도 5에 도시된 합성장치의 겨우에는, 출력장치(301)로부터 LSP데이터나 CELP데이터가 또한 취하여져 상술한 바와같이 공급된다.
도 15의 구성에 있어서, 적어도 음조데이터나 스펙트럼 엔벌로프 데이터는 상술한 바와같이 소망하는 음조에 따라 데이터변환장치(302)에 의해 변환되고 그후 파형합성장치(303)에 공급되어 거기서 변환된 데이터로부터 음성파형이 합성된다. 따라서, 음소의 변화없이 음조가 변화된 음성신호가 출력단자(304)에서 출력될 수 있다.
상술의 기술은 획일적으로 음성합성을 하거나 텍스트를 합성하는데 사용될 수 있다.
도 16은 음성텍스트합성에 응용된 본 발명의 일 예를 보여준다. 본 실시예에서, 상술의 압축을 위한 음성부호화용 디코더는 동시에 텍스트음성합성기로서 이용될 수 있다. 도 16의 실시예에서, 음성데이터의 재생이 공동으로 사용된다. 도 6에서, 음성규칙합성기와 상술한 바와같이 음조변형을 위한 데이터변환을 가지는 음성합성기가 규칙에 따른 음성합성장치(300)에 포함되어 있다. 텍스트분석장치(310)로부터의 데이터는 규칙에 따른 음성합성장치(300)로 공급되고 그로부터 소망하는 음조를 가지는 합성된 음성이 출력되어 전환스위치(330)의 고정접점에 보내진다. 음성재생장치(320)는 종종 압축되어 ROM과 같은 메모리에 기억되어 있는 음성데이터를 독출하여 팽창을 위해 그 데이터를 복호화한다. 복호화된 데이터는 전환스위치(330)의 다른 고정접점(b)에 보내진다. 합성된 음성신호와 재생된 음성신호 중 하나가 전환스위치(330)에 의해 선택되어 출력단자(340)에 출력된다.
도 16에 도시된 장치는 예를들어 차량을 위한 네비게이션장치에 사용되어질 수 있다. 이러한 경우, 음성재생기(320)로부터 고음질과 고명료함을 가지는 재생된 음성이 '우측으로 회전하십시요'와 같은 지시를 나타내는 일상적인 음성으로 사용되어지는 한편, 규칙에 따른 음성합성생성기(300)로부터의 합성된 음성이 너무 분량이 많아 ROM에 음성정보로서 저장될 수 없는 빌딩이나 지역과 같은 특정한 지시 의 음성을 위하여 사용될 수 있다.
본 발명은 같은 하드웨어가 컴퓨터음성합성기(300) 및 음성재생기(320)를 위하여 사용되어질 수 있다는 부가적인 장점을 가지고 있다.
본 발명은 상술한 실시예에 국한되지 않는다. 예를들어, 상기 하드웨어로서 기술된 도 1 및 도 3의 음성분석측(엔코더)의 구성이나 도 14의 음성합성측(디코더)의 구성은 예를들어 디지털신호처리기(DSP)를 이용하는 소프트웨어 프로그램에 의해 실현되어도 좋다. 복수 프레임의 데이터는 함께 다루어지거나 벡터양자화 대신에 행렬양자화에 의해 양자화되어도 좋다. 또한 본 발명은 다양한 음성분석/합성방법에 응용되어질 수 있다. 또한 본 발명은 송신이나 기록/재생에만 국한되지 않고, 음조변환, 속도나 레이트변환, 규칙에 따른 음성합성이나 잡음억제와 같은 다양한 용도에 응용되어 질 수 있다.
상술한 신호 부호화 및 신호 복호화장치는 예를들어 도 14에 도시된 휴대용통신단말이나 휴대용전화기세트에서 채용되는 음성부호로서 이용될 수도 있다.
도 17은 도 2와 도 3에 도시된 바와같이 구성된 음성부호화장치(160)를 채용하는 휴대용단말기의 송신측을 보여준다. 마이크로폰(161)에 의해 수집된 음성신호는 증폭기(162)에 의해 증폭되고 아나로그/디지털(A/D)변환기(163)에 의해 디지털신호로 변환되어 도 1 및 도 3에 도시된 바와같은 구성의 음성부호화장치(160)에 보내진다. A/D변환기(163)로부터의 디지털신호는 입력단자(101)에 입력된다. 음성부호화장치(160)는 도 1 및 도 3을 참조하여 설명된 바와같이 부호화를 수행한다. 도 1 및 도 2의 출력단자의 출력신호는 음성부호화장치(160)의 출력신호로서 송신 채널부호화장치(164)에 보내지고 그다음 공급된 신호에 대하여 채널코딩을 수행한다. 송신채널부호화장치(164)의 출력신호는 변조를 위하여 변조회로에 보내지고 그로부터 디지털/아나로그(D/A)변환기(166) 및 RF증폭기(167)를 거쳐서 안테나(168)에 공급된다.
도 18은 도 5 및 도 14에 도시된 바와같이 구성된 음성복호화장치(260)를 채용하는 휴대용 단말기의 수신측을 보여준다. 도 14의 안테나(261)에 의해 수신된 음성신호는 RF증폭기(262)에 의해 증폭되어, 아나로그/디지털(A/D)변환기(263)를 통하여 복조회로(264)에 보내지고, 그로부터 복조된 신호가 송신채널복호화장치 (265)에 보내진다. 복호화장치(265)의 출력신호는 도 5 및 도 14에 도시된 바와같이 구성된 음성복호화장치(260)에 공급된다. 음성복호화장치(260)는 도 5 및 도 14와 관련하여 설명된 바와같이 신호를 복호화한다. 도 2 및 도 4의 출력단자(201)에서의 출력신호는 음성복호화장치(260)의 신호로서 디지털/아나로그(D/A)변환기 (266)에 보내진다. D/A변환기(266)로부터의 아나로그음성신호는 스피커(268)에 보내진다.

(57) 청구의 범위

청구항 1.
입력음성신호를 소정의 부호화장치의 견지에서 시간축상에 분할하고 분할된 입력음성신호를 부호화함으로써 구해지는 부호화된 파라미터들에 의거하여 음성신호를 재생하는 방법에 있어서,
상기 부호화된 파라미터들을 보간처리하여 소망하는 시간점에 대한 변형된 부호화된 파라미터들을 구하고, 이 변형된 부호화된 파라미터들에 의거하여 음성신호를 재생하는 것을 특징으로 하는 음성신호 재생방법.

청구항 2.
제 1항에 있어서,
상기 부호화된 파라미터들은 고조파코딩에 대하여 구하여진 것을 특징으로 하는 음성신호 재생방법.

청구항 3.
제 1항에 있어서,

상기 입력음성신호가 유성음인지 무성음인지를 판정하고, 판정결과에 의거하여 유성음으로 판정된 입력음성신호의 부분은 정현합성부호화에 의해 부호화되고 무성음이라고 판정된 입력음성신호의 부분은 합성방법에 의한 분석을 이용하여 최적벡터의 폐루프탐색에 의하는 벡터양자화에 의해 양자화되는 것을 특징으로 하는 음성신호재생방법.

청구항 4.

제 1항에 있어서,
하나의 부호화장치에서 구해지는 부호화된 파라미터들의 시간축을 다른것으로 컴팬딩(companding)하여 상기 부호화된 파라미터들의 출력주기를 변형하는 주기변형단계와,
변형된 파라미터들을 보간처리하여 상기 부호화장치와 일치하는 시간 포인트와 관련된 변형된 부호화된 파라미터를 구하는 보간처리단계와,
상기 변형된 부호화된 파라미터에 의거하여 상기 유성음 및 무성음부분을 합성하는 음성합성단계를 더 포함하는 것을 특징으로 하는 음성신호재생방법.

청구항 5.

제 1항에 있어서,
상기 무성음부분의 합성시에 여기신호에 잡음성분이 부가되고, 상기 잡음성분은 상기 여기신호를 대체하거나 코드북으로부터 무작위로 선택된 여기벡터가 사용되는 것을 특징으로 하는 음성신호재생방법.

청구항 6.

입력음성신호를 소정의 부호화장치의 견지에서 시간축상에 분할하고 분할된 입력음성신호를 부호화함으로써 구해지는 부호화된 파라미터들에 의거하여 음성신호를 재생하는 장치에 있어서,
상기 부호화된 파라미터들은 보간처리되어 소망하는 시간점에 대한 변형된 부호화된 파라미터들을 구하고, 음성신호가 이 변형된 부호화된 파라미터들에 의거하여 재생되는 것을 특징으로 하는 음성신호 재생장치.

청구항 7.

제 6항에 있어서,
상기 부호화된 파라미터들은 고조파코딩에 대하여 구하여진 것을 특징으로 하는 음성신호 재생장치.

청구항 8.

제 6항에 있어서,
상기 입력음성신호가 유성음인지 무성음인지를 판정하고, 판정결과에 의거하여 유성음으로 판정된 입력음성신호의 부분은 정현합성부호화에 의해 부호화되고 무성음이라고 판정된 입력음성신호의 부분은 합성방법에 의한 분석을 이용하여 최적벡터의 폐루프탐색에 의하는 벡터양자화에 의해 양자화되는 것을 특징으로 하는 음성신호 재생장치.

청구항 9.

제 6항에 있어서,
하나의 부호화장치에서 구해지는 부호화된 파라미터들의 시간축을 다른것으로 컴팬딩(companding)하여 상기 부호화된 파라미터들의 출력주기를 변형하는 주기변형단계와,
변형된 파라미터들을 보간처리하여 상기 부호화장치와 일치하는 시간 포인트와 관련된 변형된 부호화된 파라미터를 구하는 보간처리단계와,
상기 변형된 부호화된 파라미터에 의거하여 상기 유성음 및 무성음부분을 합성하는 음성합성단계와,
를 더 포함하는 것을 특징으로 하는 음성신호 재생장치.

청구항 10.

제 6항에 있어서,
상기 무성음부분의 합성시에 여기신호에 잡음성분이 부가되고, 상기 잡음성분은 상기 여기신호를 대체하거나 코드북으로부터 무작위로 선택된 여기벡터가 사용되는 것을 특징으로 하는 음성신호 재생장치.

청구항 11.

입력음성신호를 소정의 부호화장치의 견지에서 시간축상에 분할하고 분할된 입력음성신호를 부호화함으로써 구해지는 부호화된 파라미터들에 의거하여 음성신호를 재생하는 방법에 있어서,
상기 부호화된 파라미터들을 이용하여 그 시간에서의 길이와 다른 길이의 블록을 가지고 음성이 재생되는 것을 특징으로 하는 음성신호재생방법.

청구항 12.

제 11항에 있어서,
상기 부호화된 파라미터는 고조파코딩이나 LPC잔차의 고조파코딩의 제 1단계와 파형부호화의 제 2단계에 의해 구해지는 것을 특징으로 하는 음성신호재생방법.

청구항 13.

제 11항에 있어서,
입력음성신호가 유성음인지 무성음인지를 판정하고, 판정결과에 의거하여 유성음으로 판정된 입력음성신호의 부분은 고조파코딩이나 LPC잔차의 고조파코딩에 의해 부호화되고 무성음이라고 판정된 입력음성신호의 부분은 합성방법에 의한 분석을 채용하는 LPC잔차의 시간파형의 벡터양자화에 의해 양자화되는 것을 특징으로 하는 음성신호 재생방법.

청구항 14.

제 12항에 있어서,

스펙트럼 엔벌로프를 나타내는 LSPs의 보간처리의 서브-프레임 길이가 디코더에 의해 지시되는 재생속도에 응답하여 변형되는 것을 특징으로 하는 음성신호 재생방법.

청구항 15.

제 11항에 있어서,
무성신호부분을 위한 여기신호의 길이가 감소될 때는 고유의 여기신호가 사용되고, 그 길이가 부족할 때는 잡음성분이 부가된 고유의 여기신호나, 잡음성분이나 여기신호코드북으로부터 무작위로 선택된 여기벡터가 사용되도록 무성신호부분을 위한 여기신호의 길이가 변형된 것을 특징으로 하는 음성신호재생방법.

청구항 16.

기본주파수와 입력 부호화된 음성데이터의 고조파의 소정 대역내 그 수를 변환하는 단계와,
각 입력고조파내 스펙트럼성분의 사이즈를 특정하는 데이터의 수를 보간 및 변형하여 합성된 음성의 음조를 변형하는 단계를
포함하는 것을 특징으로 하는 음성복호화방법.

청구항 17.

제 16항에 있어서,
상기 보간처리는 대역제한형 오버샘플링 필터를 이용하여 수행되는 것을 특징으로 하는 음성복호화방법.

청구항 18.

기본주파수와 입력 부호화된 음성데이터의 고조파의 소정 대역내 그 수를 변환하는 수단과,
각 입력고조파내 스펙트럼성분의 사이즈를 특정하는 데이터의 수를 보간 및 변형하여 합성된 음성의 음조를 변형하는 수단을 포함하는 것을 특징으로 하는 음성복호화장치.

청구항 19.

제 18항에 있어서,
상기 보간처리는 대역제한형 오버샘플링 필터를 이용하여 수행되는 것을 특징으로 하는 음성복호화장치.

청구항 20.

고조파의 진폭데이터를 출력하도록 소정의 규칙에 따라 정규음성을 합성하는 정규음성합성단계와,
입력데이터의 고조파의 기본적 주파수와 소정대역내 진폭의 수를 변환하는 많은 데이터변환단계와,
합성된 음성의 음조를 변형하도록 각 입력 고조파내 스펙트럼성분의 사이즈 를 특정하는 데이터를 보간처리하는 단계를
포함하는 것을 특징으로 하는 음성합성방법.

청구항 21.

제 20항에 있어서,
상기 보간처리는 대역제한형 오퍼샘플링필터를 이용하여 수행되는 것을 특징으로 하는 음성합성방법.

청구항 22.

고조파의 진폭데이터를 출력하도록 텍스트에 따라 정규음성을 합성하는 정규음성합성수단과,
입력데이터의 고조파의 기본적 주파수와 소정대역내 진폭의 수를 변환하는 많은 데이터변환수단과,
합성된 음성의 음조를 변형하도록 각 입력 고조파내 스펙트럼성분의 사이즈를 특정하는 데이터를 보간처리하는 수단을 포함하는 것을 특징으로 하는 음성합성장치.

청구항 23.

제 22항에 있어서,
상기 보간처리는 대역제한형 오퍼샘플링필터를 이용하여 수행되는 것을 특징으로 하는 음성합성장치.

청구항 24.

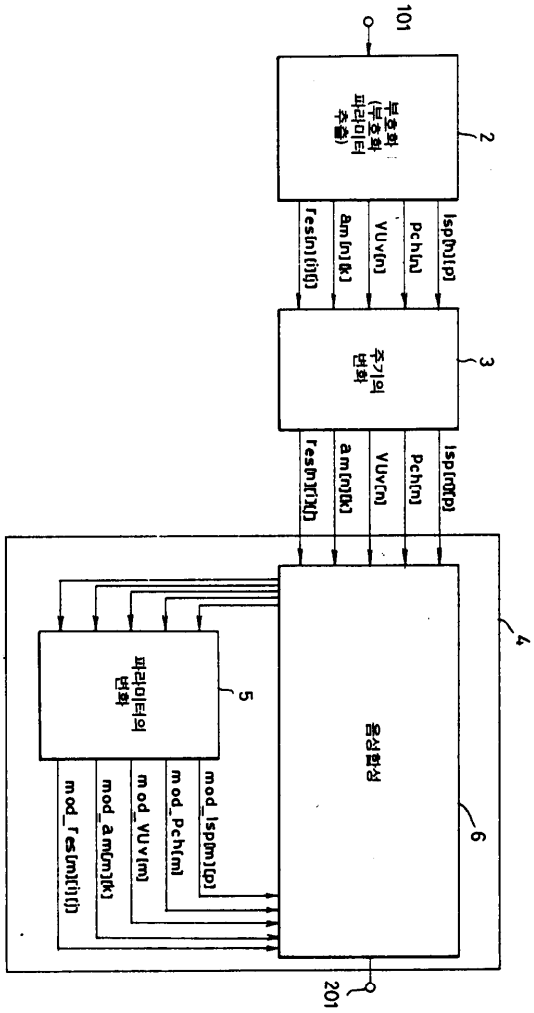
수신신호를 증폭하는 증폭수단과,
상기 증폭신호를 A/D변화하고 이어서 복조하는 복조수단과,
상기 복조된 신호를 채널-코딩하는 전송로 복호화수단과,
상기 전송로 복호화수단의 출력을 음성-복호화하는 음성복호화수단과,
상기 음성복호화수단으로부터 복호화된 음성신호를 D/A변환하여 아나로그 음성신호를 생성하는 D/A변환수단을 포함하는 것을 특징으로 하는 휴대용 무선단말기장치.

청구항 25.

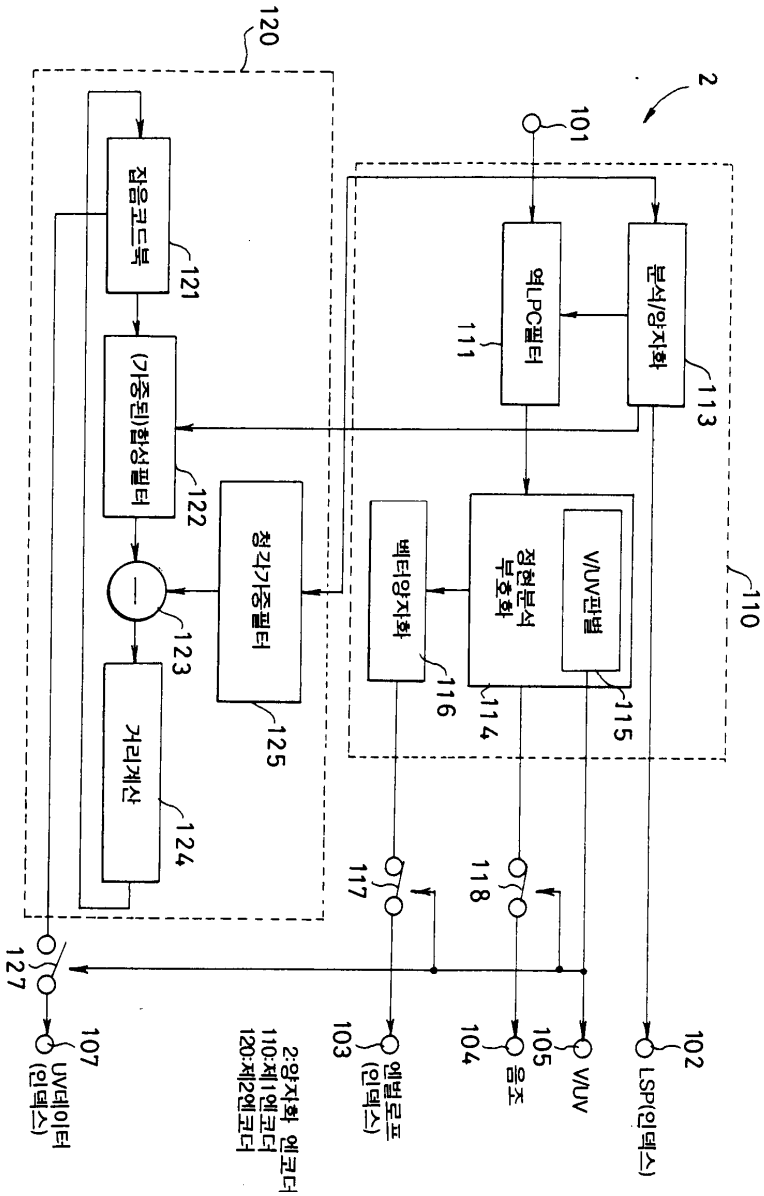
제 24항에 있어서,
상기 음성복호화수단은 입력데이터의 고조파의 기본적주파수와 소정대역내 진폭의 수를 변환하는 변환수단과,
각 입력고조파내 스펙트럼성분의 사이즈를 특정하는 데이터를 보간하여 합성된 음성의 음조를 변형시키는 수단을 포함하는 것을 특징으로 하는 휴대용 무선단말장치.

도면

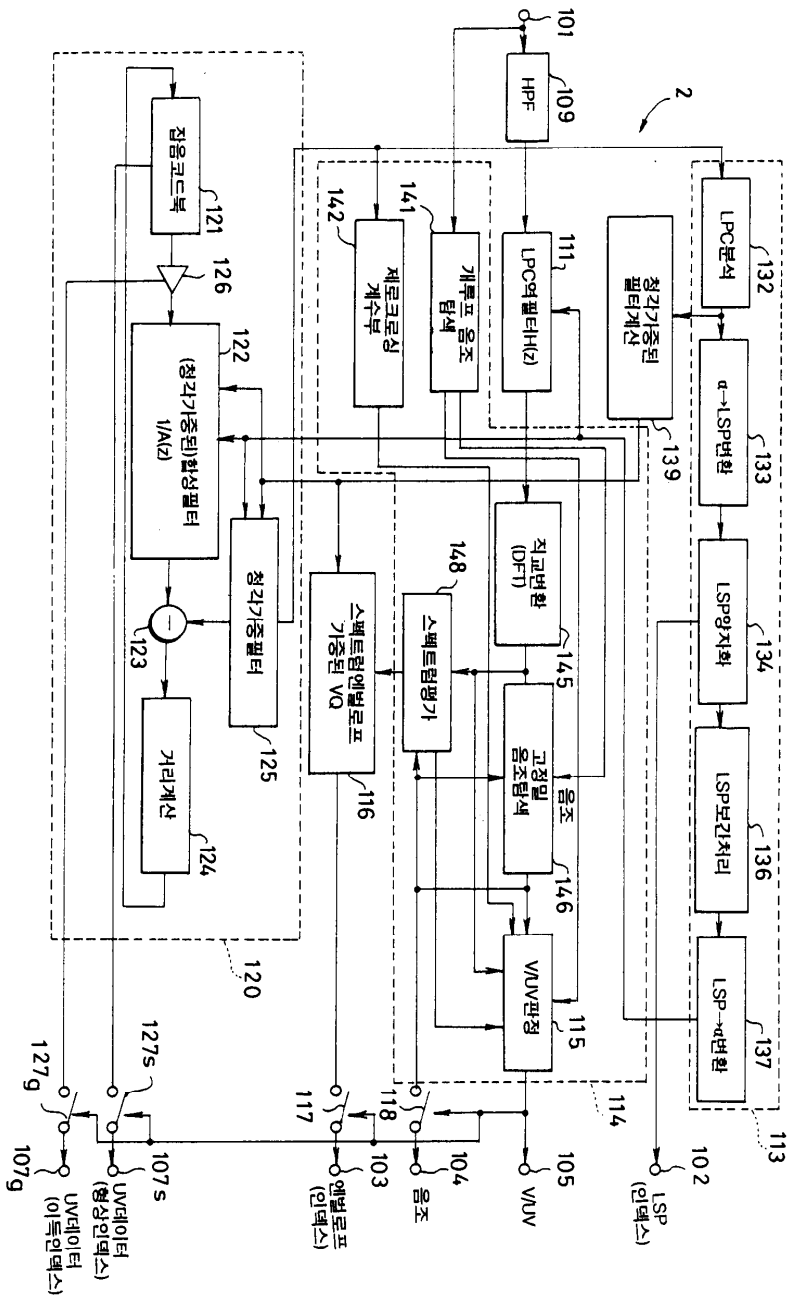
도면1


101
부호화
(부호화
파라미터
추출)
2
lsp[n][p]
Pch[n]
VUv[n]
am[n][k]
res[n][i][j]
주기의
변화
3
lsp[n][p]
Pch[n]
VUv[n]
am[n][k]
res[n][i][j]
4
음성합성
6
파라미터의
변화
5
mod_lsp[m][p]
mod_Pch[m]
mod_VUv[m]
mod_am[m][k]
mod_res[m][i][j]
201

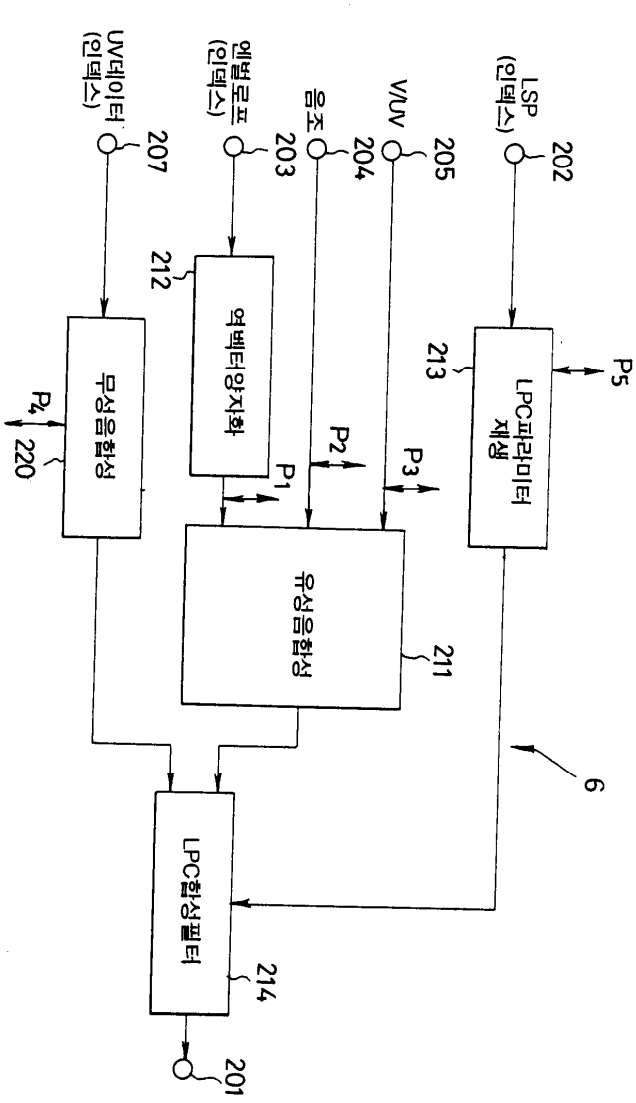
도면2


2
101
110
분석/양자화
113
역LPC필터
111
V/UV판별
정현분석
부호화
114
115
벡터양자화
116
120
잡음코드북
121
(가중된)합성필터
122
청각가중필터
125
123
거리계산
124
127
117
118
107
UV데이터
(인덱스)
2:양자화 엔코더
110:제1엔코더
120:제2엔코더
103
엔벌로프
(인덱스)
104
음조
105
V/UV
102
LSP(인덱스)

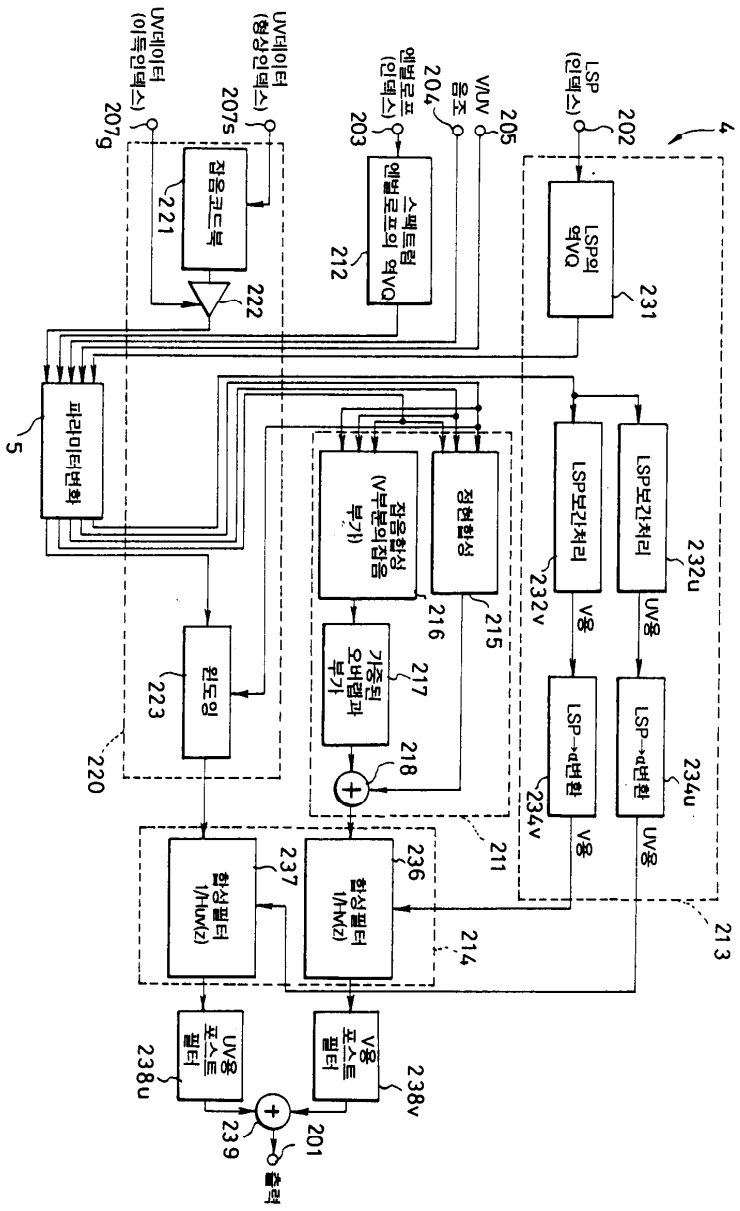
도면3


101
HPF
109
2
LPC분석
132
청각가중된 필터계산
139
LPC역필터H(z)
111
개루프 음조 탐색
141
제로코로싱 계수부
142
α→LSP변환
133
LSP양자화
134
LSP보간처리
136
LSP→α변환
137
113
114
직교변환 (DFT)
145
스펙트럼평가
148
고정밀 음조탐색
146
음조
V/UV판정
115
V/UV
105
음조
104
LSP (인덱스)
102
스펙트럼엔벨로프 가중된 VQ
116
엔벨로프 (인덱스)
103
117
118
127s
127g
107s
107g
UV데이터 (형상인덱스)
UV데이터 (이득인덱스)
120
적응코드북
121
126
(청각가중된)합성필터 1/A(z)
122
청각가중필터
125
123
거리계산
124

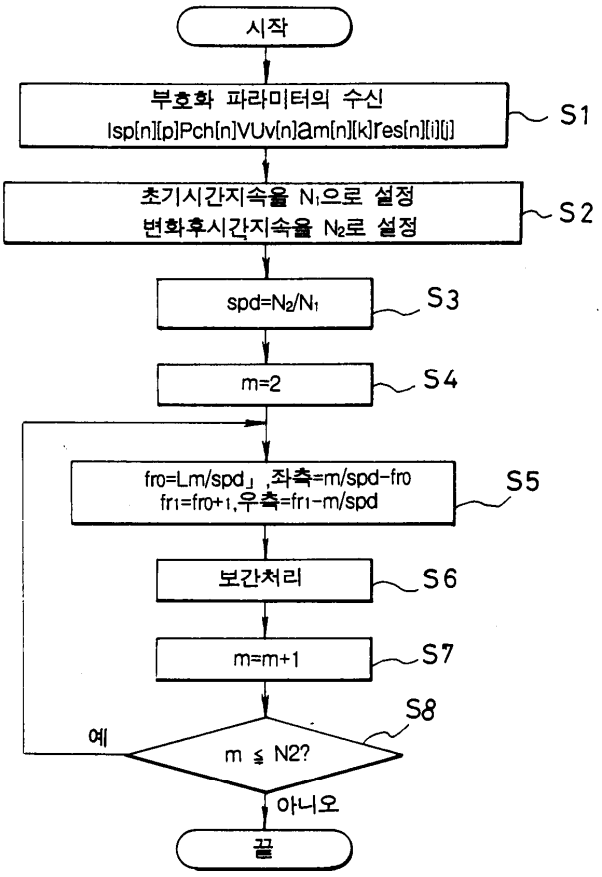
도면4


LSP (인덱스)
202
V/UV
205
204
음조
203
엔벌로프 (인덱스)
UV데이터 (인덱스)
207
212
역벡터양자화
무성음합성
P4
220
P1
P2
P3
213
LPC파라미터 재생
P5
유성음합성
211
6
LPC합성필터
214
201

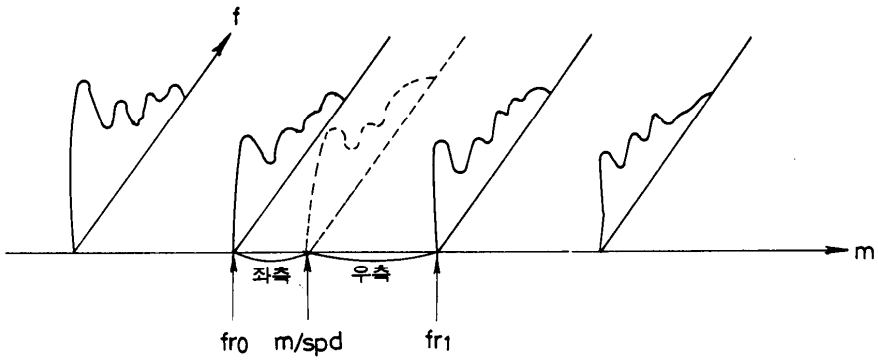
도면5


4
LSP (인덱스)
202
LSP의 역VQ
231
엔벌로프 (인덱스)
203
스펙트럼 엔벌로프의 역VQ
212
V/UV 음조
204
205
UV데이터 (형상인덱스)
207s
잡음코드북
221
222
UV데이터 (이득인덱스)
207g
파라미터변화
5
LSP보간처리
232v
LSP보간처리
232u
V용
UV용
LSP→α변환
234v
LSP→α변환
234u
V용
UV용
213
정현합성
215
잡음합성 (V부분의잡음 부가)
216
가중된 오버랩과 부가
217
218
211
윈도잉
223
220
합성필터 1/A(z)
236
합성필터 1/Ahuv(z)
237
214
V용 포스트 필터
238v
UV용 포스트 필터
238u
239
201
출력

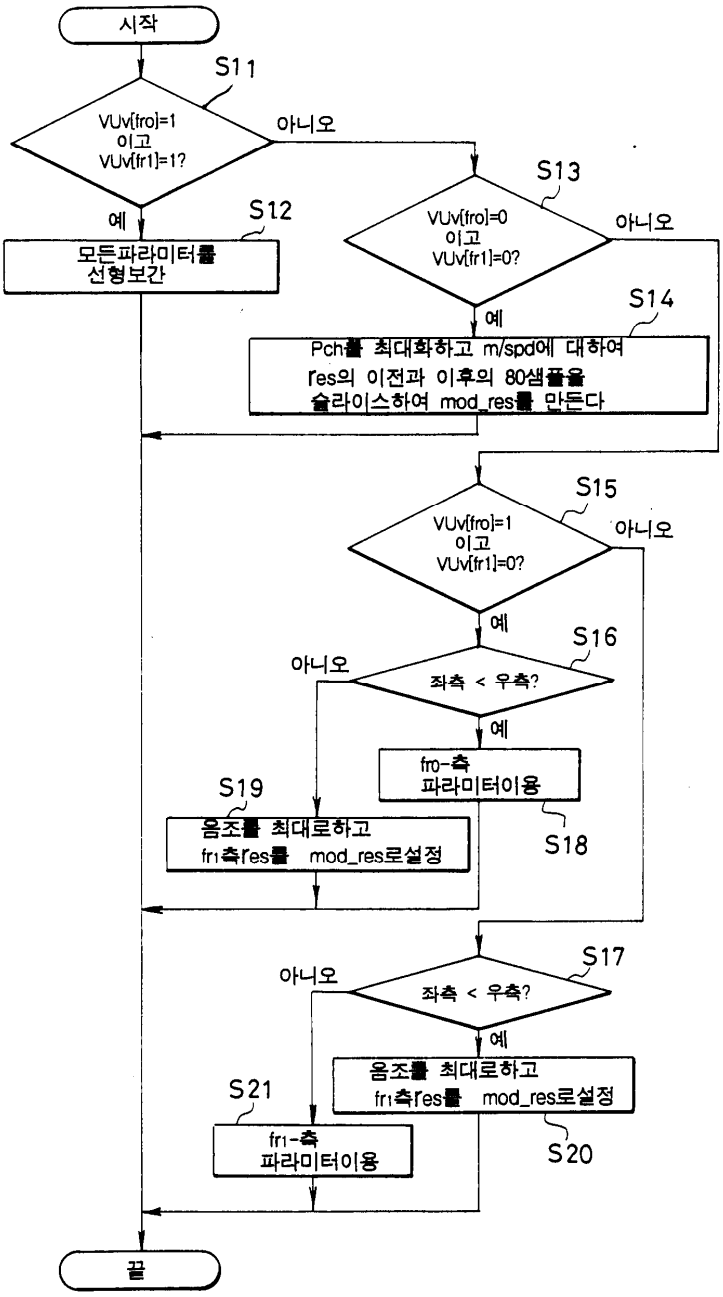
도면6


시작
부호화 파라미터의 수신
lsp[n][p]Pch[n]VUv[n]am[n][k]res[n][i][j]
S1
초기시간지속율 N1으로 설정
변화후시간지속율 N2로 설정
S2
spd=N2/N1
S3
m=2
S4
fr0=Lm/spd」,좌측=m/spd-fr0
fr1=fr0+1,우측=fr1-m/spd
S5
보간처리
S6
m=m+1
S7
S8
예
m ≦ N2?
아니오
끝

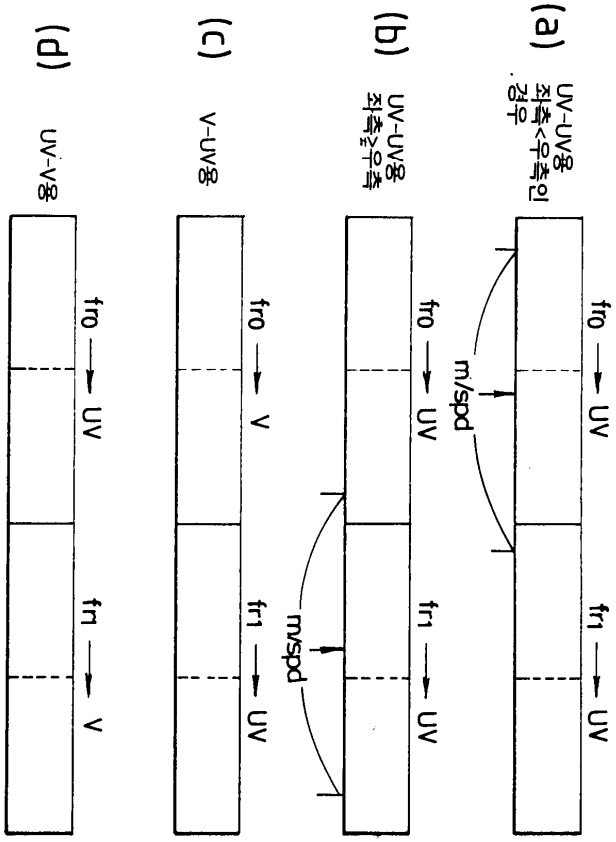
도면7


f
m
좌측
우측
fr0
m/spd
fr1

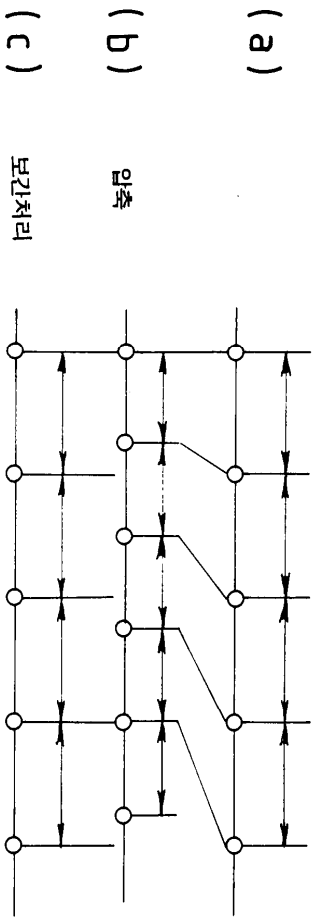
도면8


시작
S11
VUv[fro]=1
이고
VUv[fr1]=1?
아니오
예
S12
모든파라미터를
선형보간
S13
VUv[fro]=0
이고
VUv[fr1]=0?
아니오
예
S14
Pch를 최대화하고 m/spd에 대하여
res의 이전과 이후의 80샘플을
슬라이스하여 mod_res를 만든다
S15
VUv[fro]=1
이고
VUv[fr1]=0?
아니오
예
S16
아니오
좌측 < 우측?
예
fro-측
파라미터이용
S18
S19
음조를 최대로하고
fr1측res를 mod_res로설정
S17
아니오
좌측 < 우측?
예
음조를 최대로하고
fr1측res를 mod_res로설정
S20
S21
fr1-측
파라미터이용
끝

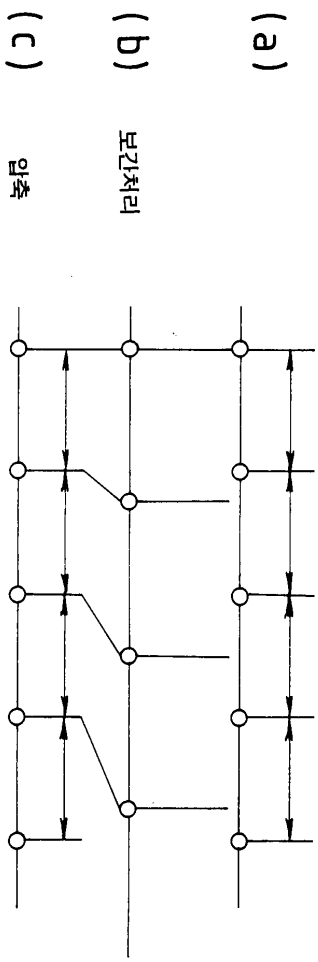
도면9


(a)
UV-UV용
좌측<우측인
경우
fr0 → UV
m/spd
fr1 → UV
(b)
UV-UV용
좌측≥우측
fr0 → UV
mspd
fr1 → UV
(c)
V-UV용
fr0 → V
fr1 → UV
(d)
UV-V용
fr0 → UV
fr1 → V

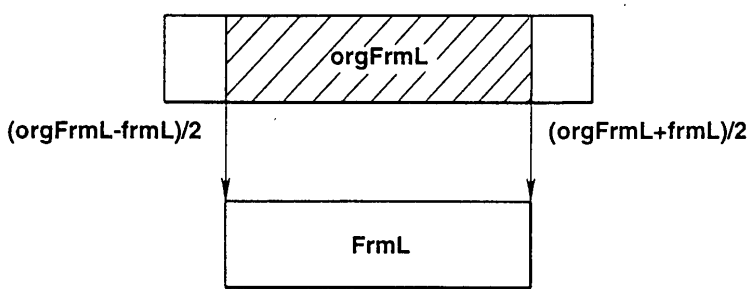
도면10



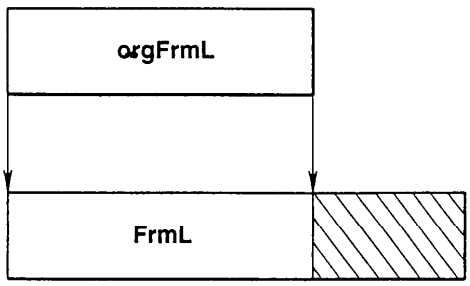
도면11



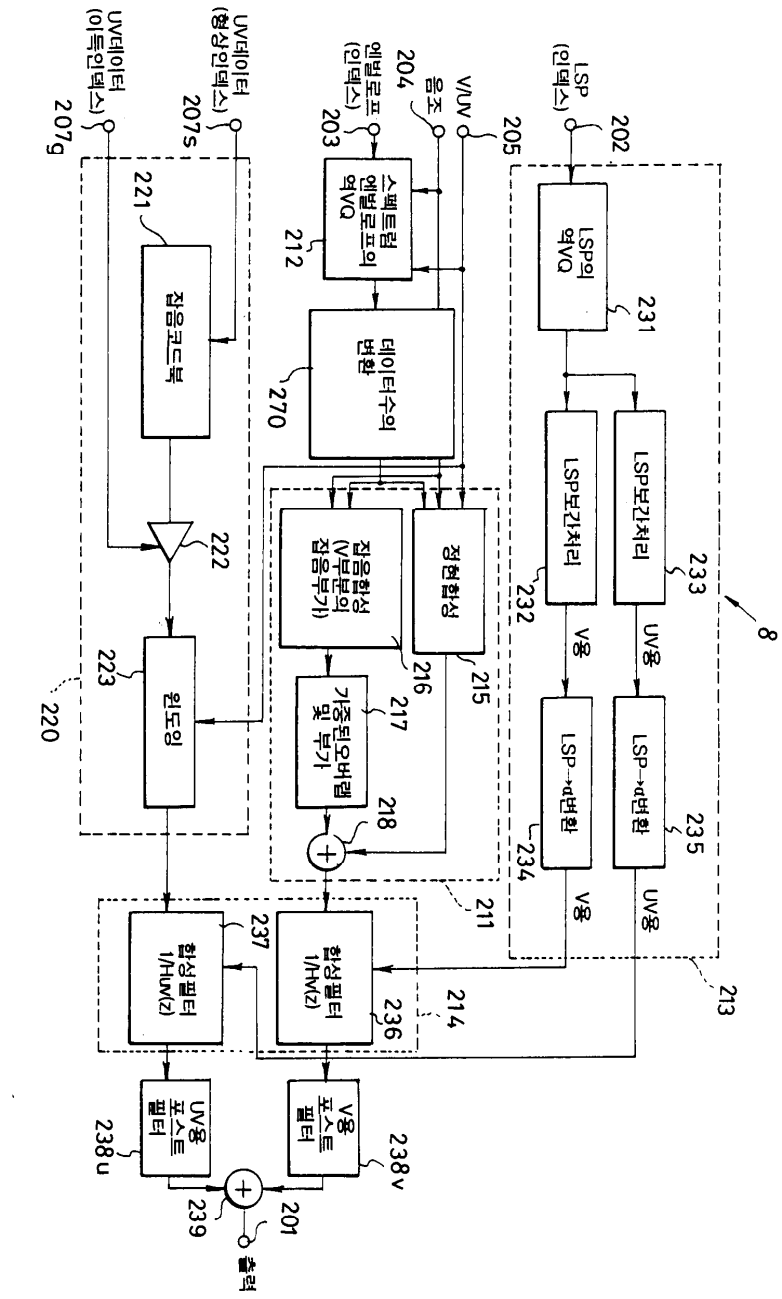
도면12



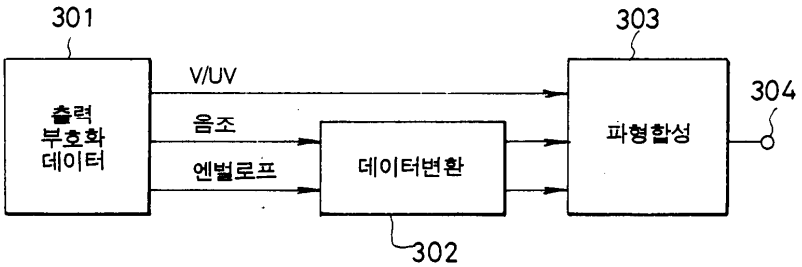
도면13



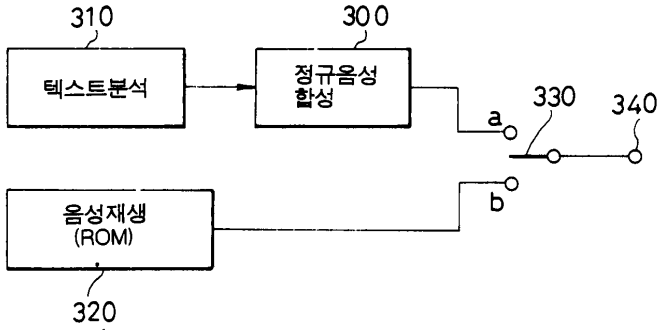
도면14


UV데이터 (이득인덱스)
207g
UV데이터 (형상인덱스)
207s
엔벌로프 (인덱스)
203
음조
204
V/UV
205
LSP (인덱스)
202
스펙트럼 엔벌로프의 역VQ
212
LSP의 역VQ
231
잡음코드북
221
데이터수의 변환
270
LSP보간처리
232
LSP보간처리
233
8
222
잡음합성 (V부분의 잡음부가)
216
정현합성
215
V용
UV용
223
윈도잉
220
가중된오버랩 및 부가
217
218
LSP→α변환
234
LSP→α변환
235
211
V용
UV용
합성필터 1/Huv(z)
237
합성필터 1/Hv(z)
236
214
213
UV용 포스트 필터
238u
V용 포스트 필터
238v
239
201
출력

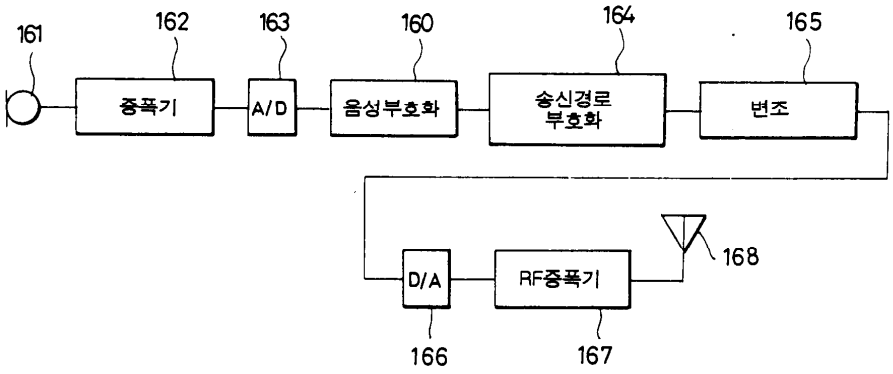
도면15


301
출력 부호화 데이터
V/UV
음조
엔벌로프
데이터변환
302
303
파형합성
304

도면16



도면17



도면18

