(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl.[6]
G06F 3/05

(45) 공고일자 2005년06월27일
(11) 등록번호 10-0469224
(24) 등록일자 2005년01월21일

(21) 출원번호 10-1996-0048088
(22) 출원일자 1996년10월24일
(65) 공개번호 10-1997-0022701
(43) 공개일자 1997년05월30일

(30) 우선권주장 95-279412 1995년10월26일 일본(JP)

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

(72) 발명자 니시구찌 마사유끼
일본국 도쿄도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

이지마 가쯔유끼
일본국 도쿄도 시나가와구 기다시나가와 6쪼메 7방 35고 소니가부시끼가이샤내

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

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

(74) 대리인 신관호

심사관 : 이은혁

(54) 음성부호화방법및장치

요약

부호화된 데이터의 비트속도를 변화시킬 수 있는 음성부호화방법 및 장치이다. 출력벡터(x)는 제 1벡터양자화부(500)의 벡터양자화부(502)에 의해 가중벡터양자화된다. 형상지수는 출력단자(503)에서 출력되는 한편, 양자화값(X_0')은 가산기(505)에서 소스벡터(x)에서 감산된다. 결과의 양자화오차벡터(y)는 제 2벡터양자화부(510)에 의해 차원으로 분할된다. 결과의 가중벡터양자화된 형상지수는 출력단자(512_1, 512_2)로 출력된다. 양자화값(Y_1', Y_2')은 가산기(513)에 의해 차원으로 합해져서 양자화값(X_0')에 가산된다. 결과의 양자화값(X_1)이 출력된다.

대표도

도 1

명세서

도면의 간단한 설명

도 1은 본 발명을 따르는 부호화방법을 수행하기 위한 음성신호 부호화방법 및 음성신호 부호화장치(엔코더)의 기본구조를 나타내는 블록도이다.

도 2는 복호화장치(디코더)인 음성신호 복호화장치(디코더)의 기본구조를 나타내는 블록도이다.

도 3은 도 1에 나타낸 음성신호 엔코더의 보다 상세한 구조를 나타내는 블록도이다.

도 4는 도 2에 나타낸 음성신호 디코더의 보다 상세한 구조를 나타내는 블록도이다.

도 5는 LPC양자화기의 기본구조를 나타내는 블록도이다.

도 6은 LPC양자화기의 보다 상세한 구조를 나타내는 블록도이다.

도 7은 벡터양자화기의 기본구조를 나타내는 블록도이다.

도 8은 벡터양자화기의 보다 상세한 구조를 나타내는 블록도이다.

도 9는 본 발명의 음성신호 엔코더의 CELP부호화부(제 2부호화부)의 상세한 구조를 나타내는 블록회로도이다.

도 10은 도 9의 구성에 있어서의 처리흐름을 설명하기 위한 플로우차트도이다.

도 11a 및 도 11b는 서로 다른 임계값으로의 클립핑 후의 가우스잡음을 도시하고 있다.

도 12는 학습(learning)으로 형상코드북을 생성할 때의 처리흐름을 나타내는 플로우차트도이다.

도 13은 본 발명을 구체화하는 음성신호 엔코더를 이용하는 휴대단말기의 송신측의 구조를 나타내는 블록도이다.

도 14는 도 13의 대응장치인 음성신호 디코더를 이용하는 휴대단말기의 수신측의 구조를 나타내는 블록도이다.

도 15는 본 발명의 음성신호 엔코더에 있어서의 서로 다른 비트속도에 대한 출력데이터를 나타내는 표이다.

*도면의 주요부분에 대한 부호의 설명

110. 제 1부호화부 111. 역LPC필터

113. LPC분석/양자화부 114. 정현분석엔코더

115. V/UV판별부 116. 벡터양자화부

117, 118, 127. 스위치 120. 제 2부호화부

121. 잡음코드북 122. 가중합성필터

123. 감산기 124. 거리계산회로

125. 청각가중필터 132. LPC분석

133. α→LSP변환회로 134. LSP양자화기

136. LSP보간처리회로 137. LSP→α변환회로

139. 청각가중필터계산회로 141. 개루프음조탐색부

142. 제로크로싱카운터 145. 직교변환회로

146. 고정밀음조탐색부 148. 스펙트럼평가부

211. 유성음성합성부 212. 역벡터양자화기

213. LPC파라미터재생부 214. LPC합성필터

220. 무성음성합성부 221. 잡음코드북

222. 이득회로 223. 윈도우처리회로

231. 역벡터양자화기 232, 233. LSP보간처리회로

234, 235. LSP→α변환회로 236, 237. LPC합성필터

238v. 유성음용 포스트필터 238u. 무성음용 포스트필터

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 입력음성신호를 블록기준으로 분할하고 그 결과의 블록을 1단위로 하는 식으로 부호화하는 음성부호화 방법 및 장치에 관한 것이다.

지금까지 시간영역 및 주파수영역내의 신호에 대한 통계적특성과 사람의 귀에 대한 정신음향적인 특성을 개발하여 압축을 위해 (음성 및 음향신호를 포함하는) 오디오신호를 부호화하기 위한 여러가지의 부호화방법이 알려져 왔다. 그러한 부호화방법은 대체로 시간영역부호화, 주파수영역부호화 및 분석/합성부호화로 분류될 수 있다.

고효율의 음성신호부호화의 예는, 고조파부호화, 멀티밴드여기(勵起)(MBE)부호화, 서브밴드부호화(SBC), 선형예측부호화(LPC), 이산코사인변환(DCT), 변형된 이산코사인변환(MDCT) 및 고속푸리에변환(FFT) 등과 같은, 정현분석부호화(sinusoidal analysis encoding)를 포함하고 있다.

고효율의 음성신호부호화를 이용하는 음성신호부호화장치에 의하면, 정현분석부호화를 이용하여 선형예측부호화(LPC)의 오차(residual) 등과 같은 단기간의 예측 오차(residual)가 부호화되며, 결과의 스펙트럼포락선의 진폭데이터는 벡터양자화되어 코드북지수데이터를 출력하게 된다.

상술한 음성신호부호화장치에 의하면, 벡터양자화에 대한 코드북지수들을 포함하는 부호화데이터의 비트속도가 일정하게 유지된 채, 변화될 수 없다.

또한, 부호화데이터가 예를들어 M비트인 경우에는, 부호화된 데이터를 복호화하기 위한 음성신호복호화장치는 M비트의 복호화장치가 될 필요가 있다. 즉, 음성신호복호화장치에 의하면, 부호화된 데이터와 동일한 수의 비트수를 갖는 복호화된 데이터만이 얻어질 수 있는 한편, 복호화된 데이터의 비트수가 변화될 수는 없다.

발명이 이루고자 하는 기술적 과제

따라서, 본 발명은 부호화데이터의 비트속도를 변화시킬 수 있는 음성부호화방법 및 장치를 제공하기 위한 것을 목적으로 하고 있다.

본 발명을 따르는 음성부호화방법 및 장치에 의하면, 적어도 입력음성신호의 유성음부에 대하여 단기간예측 오차(residual)가 구해지고, 단기간예측 오차(residual)를 기초로 하여 정현분석 부호화파라미터가 구해진다. 이러한 정현분석 부호화파라미터는 청각가중벡터양자화에 의해 양자화된다. 입력음성신호의 유성음부는 위상재생도를 갖는 파형부호화로 부호화된다. 청각가중벡터양자화에 있어서는, 제 1벡터양자화가 우선 행해지며, 제 1벡터양자화시에 생성된 양자화에러벡터는 제 2벡터양자화에 의해 양자화된다. 이러한 식으로, 출력 부호화데이터의 비트수를 데이터송신채널의 용량에 따라 용이하게 절환시킬 수 있어, 복수의 데이터비트속도를 수습할 수 있다. 또한, 그러한 부호화데이터열을 생성하여 엔코더와 디코더간의 비트속도가 다른 경우라도 디코더측에 대하여 용이하게 수습할 수 있다.

발명의 구성 및 작용

도면을 참고로 하여, 본 발명의 적절한 실시예를 상세히 설명한다.

도 1에 있어서는, 본 발명의 음성부호화방법을 수행하기 위한 음성신호 엔코더의 기본 구조가 블록도로 나타나 있다. 음성신호 엔코더는 입력음성신호의 단기간예측 오차(residual)를 구하기 위한 수단으로써의 역LPC필터(111)와, 그러한 단기간예측 오차(residual)에서부터 정현분석 부호화파라미터를 구하기 위한 수단으로써의 정현분석엔코더(114)를 포함하고 있다. 음성신호 엔코더는 또한 정현분석 부호화파라미터에 대해 청각가중 벡터양자화를 수행하기 위한 수단으로써의 벡터양자화기(116)와, 위상전송파형 부호화를 이용해 입력음성신호를 부호화하기 위한 수단으로써의 제 2부호화부(120)를 포함하고 있다.

도 2는 도 1에 나타낸 부호화장치의 대응장치(디코더)인 음성신호 복호화장치(디코더)의 기본 구조를 나타내는 블록도이며, 도 3은 도 1에 나타낸 음성신호 엔코더의 보다 구체적인 구조를 나타내는 블록도이고, 도 4는 도 2에 나타낸 음성디코더의 보다 상세한 구조를 나타내는 블록도이다.

도 1~도 4의 블록도의 구조를 설명한다.

도 1의 음성신호 엔코더의 기본 개념은 엔코더가 고조파부호화 등과 같은 정현분석부호화를 수행하기 위한 입력음성신호의, 선형예측부호화(LPC) 오차(residual) 등과 같은 단기간예측 오차(residual)를 구하기 위한 제 1부호화부(110)와, 위상재생도를 나타내는 파형부호화를 이용해 입력음성신호를 부호화하기 위한 제 2부호화부(120)를 포함하는 것과, 제 1 및 제 2부호화부(110, 120)를 각각 입력신호의 유성음부 및 무성음부를 부호화하기 위해 사용하는 것에 있다.

제 1부호화부(110)는 LPC 오차(residual)를 고조파부호화 또는 멀티밴드부호화(MBE) 등과 같은 정현분석부호화로 부호화하는 구성을 갖는다. 제 2부호화부(120)는 분석/합성방법을 이용하여 최적의 벡터에 대해 폐루프탐색에 의한 벡터양자화를 이용하는 코드여기(勵起)선형예측(CELP)의 구성을 갖는다.

본 실시예에 있어서, 입력단자(101)에 인가되는 음성신호는 제 1부호화부(110)의 역LPC필터(111)와 LPC분석/양자화부(113)로 보내진다. LPC분석/양자화부(113)로부터 얻어지는 LPC계수 또는 소위 α-파라미터는 역LPC필터(111)에 보내져서, 상기 역LPC필터(111)에 의해 입력음성신호에 대한 선형예측 오차(residual)(LPC 오차(residual))를 도출하게 된다. LPC분석/양자화부(113)로부터는, 선형스펙트럼쌍(LSP;linear spectral pairs)가 후술한 바와같이 도출되어서 출력단자(102)로 보내진다.

역LPC필터(111)로부터의 LPC 오차(residual)는 정현분석부호화부(114)로 보내진다.

정현분석부호화부(114)는 음조검출, 스펙트럼포락선 진폭계산 및 유성음(V)/무성음(UV)판별부(115)에 의한 V/UV판별을 행한다. 정현분석부호화부(114)로부터의 스펙트럼포락선 진폭데이터는 벡터양자화부(116)로 보내진다. 스펙트럼포락선의 벡터양자화출력으로써 벡터양자화부(116)로부터의 코드북지수는 스위치(117)를 통해서 출력단자(103)로 보내지는 한편, 정현분석부호화부(114)의 출력은 스위치(118)를 통해서 출력단자(104)로 보내진다. V/UV판별부(115)로부터의 V/UV판별출력은 출력단자(105)로 보내져서, 스위칭제어신호로써 스위치(117, 118)로 보내진다. 유성음(V)신호에 대해서, 지수(index)와 음조(pitch)는 출력단자(103, 104)로 도출되도록 선택된다.

본 실시예에 있어서, 도 1의 제 2부호화부(120)는 코드여기(勵起)선형예측(CELP) 부호화구조를 가지고 있어, 잡음코드북(121)의 출력이 가중합성필터(122)에 의해 합성되는 분석/합성방법에 의한 폐루프탐색을 이용하는 시간영역파형의 벡터양자화를 수행하며, 그 결과의 가중된 음성은 감산기(123)에 인가되며, 여기서 가중된 음성과 입력단자(101)로 인가되어 가중필터(125)를 통과하는 음성신호간의 에러가 도출되어서 거리계산회로(124)로 보내져서 거리계산을 수행하게 되며, 에러를 최소화하는 벡터가 잡음코드북(121)에 의해 탐색된다. 이러한 CELP부호화는 상술한 바와같이 무성음부를 부호화하는데 사용된다. 잡음코드북(121)으로부터의 UV데이터로써의 코드북지수는 V/U판별부(115)로부터의 V/UV판별결과가 무성음(UV)을 나타낼 때 턴온되는 스위치(127)를 통해서 출력단자(107)에서 도출된다.

도 2는 도 1의 음성신호 엔코더의 대응장치로써, 본 발명을 따르는 음성복호화방법을 수행하기 위한 음성신호 디코더의 기본 구조를 나타내는 블록도이다.

도 2를 참고로 하면, 도 1의 출력단자(102)로부터의 선형스펙트럼쌍(LSPs)의 양자화출력으로써의 코드북지수는 입력단자(202)로 인가된다. 도 1의 출력단자(103, 104 및 105)의 출력들, 즉 포락선 양자화출력으로써의 지수데이터, 음조 및 V/UV판별출력이 각각 입력단자(203～205)로 인가된다. 무성음데이터용 데이터로써의 지수데이터는 도 1의 출력단자(107)에서 입력단자(207)로 인가된다.

입력단자(203)의 양자화출력으로써의 지수는 역벡터양자화용의 역벡터양자화부(212)로 보내져서 유성음성합성부(211)로 보내지게 되는 LPC 오차(residual)의 스펙트럼포락선을 구하게 된다. 유성음성합성부(211)는 정현합성을 이용해 유성음성부에 대한 선형예측부호화(LPC) 오차(residual)들을 합성한다. 유성음성합성부(211)에는 또한 입력단자(204, 205)로부터의 음조 및 V/UV판별출력이 공급된다. 유성음성합성부(211)로부터의 유성음성에 대한 LPC 오차(residual)는 LPC합성필터(214)로 보내진다. 입력단자(207)로부터의 UV데이터에 대한 지수데이터는 무성음성합성부(220)로 인가되며, 여기서 무성음부의 LPC 오차(residual)를 도출하기 위해서 잡음코드북에 대한 조회가 행해진다. 이러한 LPC 오차(residual)는 또한 LPC합성필터(214)로 보내진다. LPC합성필터(214)에서는, 유성음부의 LPC 오차(residual)와 무성음부의 LPC 오차(residual)가 LPC합성에 의해 처리된다. 선택적으로, 서로 합해진 유성음부의 LPC 오차(residual)와 무성음부의 LPC 오차(residual)가 LPC합성으로 처리될 수도 있다. 입력단자(202)로부터의 LSP지수데이터는 LPC파라미터재생부(213)로 인가되고 여기서 LPC의 α-파라미터가 도출되어 LPC합성필터(214)로 보내진다. LPC합성필터(214)에 의해 합성된 음성신호는 출력단자(201)에서 도출된다. 이제, 도 3을 참고로 하여, 도 1에 나타낸 음성신호 엔코더의 보다 상세한 구조를 설명한다. 도 3에서, 도 1에 나타낸 구성요소와 동일한 구성요소는 동일한 참조부호로 표시한다.

도 3에 나타낸 음성신호 엔코더에 있어서는, 입력단자(101)에 인가되는 음성신호를 고역통과필터(109)로 필터처리하여 불필요한 영역의 신호를 제거한 다음, LPC분석/양자화부(113)의 LPC분석회로(132)와 역LPC필터(111)로 인가한다. LPC분석/양자화부(113)의 LPC분석회로(132)는 1블록으로써 256샘플의 차수로 입력신호파형의 길이를 갖는 해밍 윈도우(Hamming window)를 적용하며, 자기상관방법을 이용하여, 소위 α-파라미터의 선형예측계수를 구한다. 데이터출력단위로써의 프레임간격은 대략 160샘플로 설정된다. 샘플링주파수(fs)가 예를들어 8kHz이면, 160샘플에 대한 1프레임간격은 20msec이다.

LPC분석회로(132)로부터의 α-파라미터는 α→LSP변환회로(133)로 보내져서 선스펙트럼쌍(LSP) 파라미터로 변환된다. α→LSP변환회로(133)는 다이렉트형 필터계수에 의해 구해진 α-파라미터를 예를들어 10개 즉, 5쌍의 LSP파라미터로 변환한다.

이러한 변환은 예를들어 뉴튼-랩슨(Newton-Rhapson)방법에 의해 수행된다. α-파라미터가 LSP파라미터로 변환되는 이유는 LSP파라미터가 α-파라미터보다 보간특성이 우수하기 때문이다.

α→LSP변환회로(133)로부터의 LSP파라미터는 LSP양자화기(134)에 의해 행렬 또는 벡터양자화된다. 벡터양자화에 앞서 프레임 대 프레임 차를 취할 수도, 또는 행렬양자화를 행하기 위해서 복수의 프레임을 수집할 수도 있다. 본 발명의 경우에 있어서는, 매 20msec마다 계산된 2프레임(20msec)의 LSP파라미터가 수집되어 행렬양자화 및 벡터양자화로 처리된다.

양자화기(134)의 양자화된 출력, 즉 LSP양자화의 지수데이터는 단자(102)에서 도출되는 한편, 양자화된 LSP벡터는 LSP보간처리회로(136)로 보내진다.

LSP보간처리회로(136)는 매 20msec 또는 40msec마다 양자화된 LSP벡터들을 보간처리하여 8-폴드레이트(fold rate)를 제공하게 된다. 즉, LSP벡터는 매 2.5msec마다 업데이트된다. 그 이유는, 만일 오차(residual)파형이 고조파 부호화/복호화방법에 의해 분석/합성처리되면, 합성파형의 포락선이 극도로 누그러진 파형을 제공하므로, LPC계수가 매 20msec마다 갑작스럽게 변화하는 경우에는, 이상잡음(foreign noise)이 생성되기 쉽기 때문이다. 즉, LPC계수가 매 2.5msec마다 점차적으로 변하게 되면, 그러한 이상잡음의 발생을 방지할 수 있다.

매 2.5msec마다 생성된 보간처리된 LSP벡터를 사용하여 입력음성에 대한 역필터처리를 위해서, LSP파라미터는 LSP→α변환회로(137)에 예로써 10-차수의 다이렉트형 필터의 계수로써의 α-파라미터로 변환된다. LSP→α변환회로(137)의 출력은 LPC역필터회로(111)로 인가되며, 그런다음 여기서 역필터처리를 수행하여 매 2.5msec마다 업데이트된 α-파라미터를 이용하여 양호한 출력을 생성하게 된다. 역LPC필터(111)의 출력은 고조파부호화회로 등과 같은 정현분석부호화부(114)의 DCT회로 등과 같은 직교변환회로(145)로 보내진다.

LPC분석/양자화부(113)의 LPC분석회로(132)로부터의 α-파라미터는 청각가중필터계산회로(139)로 보내지며, 여기서 청각가중용 데이터가 구해진다. 이러한 가중데이터는 청각가중벡터양자화기(116)와, 제 2부호화부(120)의 청각가중필터(125) 및 청각가중합성필터(122)로 보내진다.

고조파부호화회로의 정현분석부호화부(114)는 고조파부호화방법을 이용해 역LPC필터(111)의 출력을 분석한다. 즉, 음조검출과, 각 고조파의 진폭(Am)의 계산 및 유성음(V)/무성음(UV)판별이 수행되며, 음조에 따라 변환되는 진폭(Am)의 수 또는 각 고조파의 포락선은 차원변환에 의해 일정하게 된다.

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

특히, 멀티밴드 여기(勵起)(MBE)부호화에 있어서, 유성음부 및 무성음부는 (동일한 블록이나 프레임내의) 동일시점에서의 주파수영역 또는 대역내에 존재하는 모델이라 가정한다. 다른 고조파부호화기술에 있어서는, 1블록 또는 1프레임내의 음성이 유성음인지 무성음인지를 양호하게 판정한다. 다음의 설명에 있어서, 주어진 프레임은 MBE부호화가 관련되는 한, 전체의 대역이 UV이면 UV가 되도록 판정된다.

도 3의 정현분석부호화부(114)의 개루프음조탐색부(141) 및 제로크로싱카운터(142)는 입력단자(101)로부터의 입력음성신호와, 고역통과필터(HPF)(109)로부터의 신호를 각각 인가받는다. 정현분석부호화부(114)의 직교변환회로(145)에는 역LPC필터(111)로부터 LPC 오차(residual) 또는 선형예측 오차(residual)가 인가된다. 개루프음조탐색부(141)는 입력신호의 LPC 오차(residual)를 취하여 개루프를 이용해 비교적 거친 음조탐색을 수행한다. 추출된 거친 음조데이터는 후술하는 바와같이 폐루프에 의해 고정밀음조탐색부(146)로 보내진다. 개루프음조탐색부(141)로부터는, 거친 음조데이터와 함께 LPC 오차(residual)의 자기상관최대치를 규격화하여 얻어진 규격화된 자기상관최대치(r(p))가 거친 음조데이터와 함께 도출되어서 V/UV판별부(115)로 보내진다.

직교변환회로(145)는 이산푸리에변환(DFT)등의 직교변환을 수행하여 시간축상의 LPC 오차(residual)를 주파수축상의 스펙트럼진폭데이터로 변환한다. 직교변환회로(145)의 출력은 고정밀음조탐색부(146) 및 스펙트럼평가부(148)로 보내져서 스펙트럼진폭 또는 포락선을 구하게 된다.

고정밀음조탐색부(146)에는 개루프음조탐색부(141)에 의해 추출된 비교적 거친 음조데이터와 직교변환부(145)에 의한 DFT로 얻어진 주파수영역데이터가 인가된다. 고정밀음조탐색부(146)는 거친 음조값데이터에 중심에 있는 ±6~7개의 샘플에 의한 음조데이터를 0.2~0.5의 속도로 회전시켜서, 결국 상기 음조데이터로 하여금 최적의 10진포인트(부동포인트)를 갖는 고정밀음조데이터의 값에 다다르게 한다. 분석/합성방법은 전력스펙트럼이 본래의 음성에 대한 전력스펙트럼에 근접하게 되도록, 음조를 선택하기 위한 고정밀탐색기술로써 사용된다. 폐루프 고정밀음조탐색부(146)로부터의 음조데이터는 스위치(118)를 통해서 출력단자(104)로 보내진다.

스펙트럼 평가부(148)에 있어서, 각 고조파의 진폭 및 그 고조파의 합으로써의 스펙트럼포락선은 스펙트럼진폭 및 LPC 오차(residual)의 직교변환출력으로써의 음조를 기초로 하여 구해져서, 고정밀음조탐색부(146), V/UV판별부(115) 및 청각가중벡터양자화부(116)로 보내진다.

V/UV판별부(115)는 직교변환회로(145)의 출력, 고정밀음조탐색부(146)로부터의 최적의 음조, 스펙트럼평가부(148)로부터의 스펙트럼진폭데이터, 개루프 음조탐색부(141)로부터의 규격화된 자기상관최대치(r(p)), 및 제로크로싱카운터(142)의 제로크로싱카운트값을 기초로 하여, 프레임에 대한 V/UV를 판별한다. 또한, MBE용 대역에 기준한 V/UV판별의 경계위치도 V/UV판별에 대한 조건으로 사용될 수 있다. V/UV판별부(115)의 판별출력은 출력단자(105)에서 도출된다.

스펙트럼평가부(148)의 출력부 또는 벡터양자화부(116)의 입력부는 데이터수 변환부(일종의 샘플링속도변환을 수행하는 장치부)를 부설하고 있다. 데이터수 변환부는 주파수축상에 분할된 대역수와 데이터수가 음조와 다른점을 고려하여, 포락선의 진폭데이터(Am)를 설정하기 위해 사용되는 것이다. 즉, 유효대역이 3400kHz에 다다르면, 유효대역은 음조에 따

라서 8~63대역으로 분할될 수 있다. 대역에서 대역까지에서 얻어진 진폭데이터(Am)의 mMx + 1의 수는 8~63의 범위로 변화된다. 따라서, 데이터수 변환부는 변수(mMx + 1)의 진폭데이터를 44데이터 등의 소정수(M)의 데이터로 변환한다.

스펙트럼평가부(148)의 출력부 또는 벡터양자화부(116)의 입력부에 부설된 데이터수 변환부로부터의, 44 등의, 소정수(M)의 진폭데이터 또는 포락선데이터는 가중벡터양자화를 수행함으로써 벡터양자화부(116)에 의해 44데이터 등의 소정수의 데이터를 단위로 하는 식으로 수집된다. 이러한 가중은 청각가중필터계산회로(139)의 출력에 의해 인가된다. 벡터양자화부(116)로부터의 포락선의 지수는 스위치(117)에 의해 출력단자(103)에서 도출된다. 가중 벡터양자화에 앞서, 소정수의 데이터로 이루어진 벡터를 위해서는 적절한 누설계수를 사용하여 프레임간의 차를 구하는 것이 상책이다.

제 2부호화부(120)를 설명한다. 제 2부호화부(120)는 소위 CELP부호화구조를 가지며 특히 입력음성신호의 무성음부를 부호화하는데 사용된다. 입력음성신호의 무성음부에 대한 CELP부호화구조에 있어서, 잡음코드북 또는 소위 확률적코드북의 대표값출력으로써 무성음의 LPC 오차(residual)에 상응하는 잡음출력은 이득제어회로(126)를 통해 청각가중합성필터(122)로 보내진다. 청각가중합성필터(122)는 입력잡음을 LPC합성하여 생성된 가중 무성음신호를 감산기(123)로 보낸다. 감산기(123)에는 고역통과필터(HPF)(109)를 통해서 입력단자(101)로부터 인가되어 청각가중필터(125)에 의해 청각가중된 신호가 인가된다. 그 신호와 합성필터(122)로부터의 신호와의 차 또는 에러가 도출된다. 한편, 청각가중합성필터의 제로입력응답은 청각가중합성필터(125)의 출력에서 사전에 감산된다. 이러한 에러는 거리계산을 위한 거리계산회로(124)로 인가된다. 에러를 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 각각 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.

도 4는 도 2에 나타낸 음성신호디코더의 보다 상세한 구조를 나타낸다. 도 4에서는, 도 2에 나타낸 구성소자를 표시하는데 동일한 참조부호를 사용하였다.

도 4에서, 도 1 및 도 3의 출력단자(102)에 상응하는 LSP의 벡터양자화출력, 즉 코드북지수는 입력단자(202)로 인가된다.

LSP지수는 선스펙트럼쌍(LSP)데이터로 역벡터양자화되도록 LPC파라미터생성부(213)용 LSP의 역벡터양자화기(231)로 보내져서, 보간처리를 위해 LSP보간처리회로(232, 233)로 인가된다. 결과의 보간처리된 데이터는 LSP→α변환회로(234, 235)에 의해 α-파라미터로 변환되어서, LPC합성필터(214)로 보내진다. LSP보간처리회로(232)와 LSP→α변환회로(234)는 유성음(V)사운드용으로 설계되는 한편, LSP보간처리회로(233)와 LSP→α변환회로(235)는 무성음(UV)용으로 설계된다. LPC합성필터(214)는 유성음성부의 LPC합성필터(236)를 무성음성부의 LPC합성필터(237)에서 분리해낸다. 즉, LPC계수보간처리가 유성음성부와 무성음성부에 대해 독립적으로 수행되어, 완전히 다른 특성의 LSPs에 대한 보간처리에 의해 유성음성부에서 무성음성부로의 과도부분 혹은 그 반대로의 과도부분내에 역효과가 발생하는 것을 방지하게 된다.

도 4의 입력단자(203)로는 도 1 및 도 3의 엔코더의 단자(103)의 출력에 상응하여 가중된 스펙트럼포락선(Am)에 상응하는 코드지수데이터가 인가된다. 입력단자(204)로는 도 1 및 도 3의 단자(104)로부터의 음조데이터가 인가된다. 입력단자(205)로는 도 1 및 도 3의 단자(105)로부터의 V/UV판별데이터가 인가된다.

입력단자(203)로부터의 스펙트럼포락선(Am)의 벡터양자화 지수데이터는 역벡터양자화용의 역벡터양자화기(212)로 인가되며, 여기서 데이터수변환에 대한 역변환이 수행된다. 결과의 스펙트럼포락선데이터는 정현합성회로(215)로 보내진다.

프레임간의 차가 부호화중에 스펙트럼의 벡터양자화에 앞서 구해지게 되면, 프레임간의 차는 스펙트럼포락선데이터를 생성하는 역벡터양자화 이후에 복호화된다.

정현합성회로(215)로는 입력단자(204)로부터의 음조데이터와 입력단자(205)로부터의 V/UV판별데이터가 인가된다. 정현합성회로(215)로부터는 도 1 및 도 3에 나타낸 LPC역필터의 출력에 상응하는 LPC 오차(residual)데이터가 도출되어 가산기(218)로 보내진다.

역벡터양자화기(212)의 포락선데이터와 입력단자(204, 205)로부터의 음조데이터 및 V/UV판별데이터는 유성음부(V)에 대해 잡음가산을 위한 잡음합성회로(216)로 인가된다. 잡음합성회로(216)의 출력은 가중중첩-가산회로(217)를 통해서 가산기(218)로 인가된다. 특히, 유성음의 LPC합성필터에 대한 입력으로써 여기(勵起)가 정현파합성에 의해 생성되는 경우에는, 남성음성등의 저음조의 음에서는 답답한 느낌이 생성되고, 유성음과 무성음간의 음질이 급격히 변환함에 따라서 부자연스런 청감이 생성되는 사실을 고려하여, 잡음을 LPC 오차(residual)신호의 유성음성부에 가산한다. 그러한 잡음은 유성음성부의 LPC합성필터입력 즉, 여기(勵起)와 결부시켜서 볼때, 음조, 스펙트럼포락선의 진폭, 프레임의 최대진폭 또는 오차(residual)신호레벨 등의 음성부호화데이터와 관계된 파라미터를 고려하고 있다.

가산기(218)의 가산출력은 LPC합성필터(214)의 유성음용 합성필터(236)에 인가되고, 여기서 LPC합성이 수행되어 시간파형데이터를 이루게 된 다음, 유성음성용 포스트필터(238v)에 의해 필터처리되어 가산기(239)로 보내진다.

도 3의 출력단자(107s, 107g)로부터의 UV데이터로써는, 형상지수와 이득지수가 도 4의 입력단자(207s, 207g)로 인가된 다음, 무성음성합성부(220)로 보내진다. 단자(207s)로부터의 형상지수는 무성음성합성부(220)의 잡음코드북(221)으로 인가되는 한편, 단자(207g)로부터의 이득지수는 이득회로(222)로 인가된다. 잡음코드북(221)에서 독출된 대표값출력은 무성음에 대한 LPC 오차(residual)에 상응하는 잡음신호성분이다. 이러한 대표값출력은 이득회로(222)에서 소정의 이득진폭이 되고, 유성음부에 대해 완만한 접속을 위해 윈도우처리되도록 윈도우처리회로(223)로 보내진다.

윈도우처리회로(223)의 출력은 LPC합성필터(214)의 무성음(UV)용 합성필터(237)로 보내진다. 합성필터(237)로 인가된 데이터는 LPC합성처리되어 무성음부에 대한 시간파형데이터가 된다. 무성음부에 대한 시간파형데이터는 가산기(239)에 인가되기 전에 무성음용 포스트필터에 의해 필터처리된다.

가산기(239)에서는, 유성음성용 포스트필터(238v)로부터의 시간파형신호와 무성음성용 포스트필터(238u)로부터의 시간파형신호데이터가 서로 합해지고, 그 결과의 합산데이터는 출력단자(201)로 도출된다.

상술한 음성신호 엔코더는 요구된 음질에 따라서 서로 다른 비트속도의 데이터를 출력할 수 있다. 즉, 출력데이터는 다양한 비트속도로 출력될 수 있다. 예를들어, 저속의 비트속도가 2kbps이고 고속의 비트속도가 6kbps이면, 출력데이터는 도 15에 나타낸 비트속도를 갖는 비트속도데이터이다.

출력단자(104)로부터의 음조데이터는 언제나 8비트/20msec의 비트속도로 유성음성용으로 출력되며, 출력단자(105)로부터의 V/UV판별출력은 언제나 1비트/20msec로 출력되게 된다. 출력단자(102)로부터 출력된 LSP양자화용 지수는 32비트/40msec와 48비트/40msec사이로 절환된다. 반면에, 출력단자(103)에 의해 유성음성(V)이 출력되는 동안의 지수는 15비트/20msec와 87비트/20msec사이로 절환된다.

출력단자(107s, 107g)로부터 출력된 무성음(UV)용 지수는 11비트/10msec와 23비트/5msec사이로 절환된다. 유성음성(UV)용 출력데이터는 2kbps에 대해서는 40비트/20msec이고 6kbps에 대해서는 120kbps/20msec이다. 반면에, 무성음성(UV)용 출력데이터는 2kbps에 대해서는 39비트/20msec이고 6kbps에 대해서는 117kbps/20msec이다.

LSP양자화용 지수, 유성음성(V)용 지수 및 무성음성(UV)용 지수를 이하 적합한 부분의 구성과 결부시켜 설명한다.

도 5 및 도 6을 참고로 하여, LSP양자화기(134)에 있어서의 행렬양자화 및 벡터양자화를 상세히 설명한다.

LPC분석회로(132)로부터의 α-파라미터는 LSP파라미터로 변환하기 위한 α→LSP변환회로(133)로 보내진다. P-차수의 LPC분석이 LPC분석회로(132)에서 수행되면, Pα-파라미터가 계산된다. 이러한 Pα-파라미터는 버퍼(610)에 유지되어 있는 LSP파라미터로 변환된다.

버퍼(610)는 2프레임의 LSP파라미터를 출력한다. 이 2프레임의 LSP파라미터는 제 1행렬양자화기(620_1)와 제 2행렬양자화기(620_2)로 이루어진 행렬양자화기(620)에 의해 행렬-양자화된다. 2프레임의 LSP파라미터는 제 1행렬양자화기(620_1)에서 행렬-양자화되며, 결과의 양자화에러도 또한 제 2행렬양자화기(620_2)에서 행렬-양자화된다. 행렬양자화는 시간축과 주파수축의 양측으로의 상관율을 촉진한다.

행렬양자화기(620_2)로부터의 두개의 프레임에 대한 양자화에러는 제 1벡터양자화기(640_1)와 제 2벡터양자화기(640_2)로 이루어진 벡터양자화부(640)로 입력된다. 제 1벡터양자화기(640_1)는 두개의 벡터양자화부(650, 660)로 이루어지는 한편, 제 2벡터양자화기(640_2)는 두개의 벡터양자화부(670, 680)로 이루어져 있다. 행렬양자화부(620)로부터의 양자화에러는 제 1벡터양자화기(640_1)의 벡터양자화부(650, 660)에 의해 프레임기준으로 양자화된다. 그 결과의 양자화에러벡터는 또한 제 2벡터양자화기(640_2)의 벡터양자화부(670, 680)에 의해 벡터양자화된다. 상술한 벡터양자화는 주파수축을 따라 상관율을 촉진한다.

상술한 바와같이 행렬양자화를 수행하는 행렬양자화부(620)는 제 1행렬양자화단계를 수행하기 위한 적어도 하나의 제 1행렬양자화기(620_1)와 제 1행렬양자화에 의해 생성된 양자화에러를 행렬양자화하기 위한 제 2행렬양자화단계를 수행하기 위한 적어도 하나의 제 2행렬양자화기(620_2)를 포함하고 있다. 상술한 바와같이 벡터양자화를 수행하는 벡터양자화부(640)는 제 1벡터양자화단계를 수행하기 위한 적어도 하나의 제 1벡터양자화기(640_1)와 제 1벡터양자화에 의해 생성된 양자화레러를 벡터양자화하기 위한 제 2벡터양자화단계를 수행하기 위한 적어도 하나의 제 2벡터양자화기(640_2)를 포함하고 있다.

이하, 행렬양자화 및 벡터양자화를 상세히 설명한다.

버퍼(600)에 저장되어 있는 2개의 프레임에 대한 LSP파라미터, 즉 10×2행렬은 제 1행렬양자화기(620_1)로 보내진다. 제 1행렬양자화기(620_1)는 2프레임의 LSP파라미터를 LSP파라미터 가산기(621)를 통해 가중거리계산부(623)로 인가하여 최소값의 가중거리를 구한다.

제 1행렬양자화기(620_1)에 의한 코드북탐색동안의 왜곡측정(d_{MQ1})은 다음 수학식 1로 주어진다.

$$d_{MQ1}(X_1, X_1') = \sum_{t=0}^{1} \sum_{i=1}^{P} w(t,i)(x_1(t,i) - x_1'(t,i))^2$$

여기서, X_1는 LSP파라미터이고 X_1'는 양자화값인 한편, t 및 i는 P-차원수가 된다.

주파수축 및 시간축으로의 무게제한을 고려하지 않은 상태에서의 무게(W(t, i))는 수학식 2로 주어진다.

$$w(t,i) = \frac{1}{x(t,i+1) - x(t,i)} + \frac{1}{x(t,i) - x(t,i-1)}$$

여기서 x(t, 0) = 0이고, x(t, p+ 1)는 t에 관계없이 π이다.

수학식 2의 무게는 또한 하류측의 행렬양자화 및 벡터양자화에도 사용된다.

계산된 가중거리는 행렬양자화용 행렬양자화기(MQ_1)(622)로 보내진다. 이 행렬양자화에 의해 출력된 8-비트의 지수는 신호절환기(690)로 보내진다. 행렬양자화에 의한 양자화값은 가산기(621)에서 2프레임에 대한 LSP파라미터에서 감산된다. 가중거리계산부(623)는 매 2프레임마다의 가중거리를 연속적으로 계산하며, 그 결과 행렬양자화가 행렬양자화부(622)에서 수행된다. 또한, 가중거리를 최소화하는 양자화값이 선택된다. 가산기(621)의 출력은 제 2행렬양자화기(620)의 가산기(631)로 보내진다.

제 1행렬양자화기(620_1)와 비슷하게, 제 2행렬양자화기(620_2)도 행렬양자화를 수행한다. 가산기(621)의 출력은 가산기(631)를 통해 가중거리 계산부(633)로 보내지며, 여기서 최소의 가중거리가 계산된다.

제 2행렬양자화기(620_2)에 의한 코드북 탐색동안의 왜곡측정(d_{MQ2})은 수학식 3에 의해 주어진다.

$$d_{MQ2}(X_2, X_2') = \sum_{t=0}^{1} \sum_{i=1}^{P} w(t,i)(x_2(t,i) - x_2'(t,i))^2$$

여기서 X_2 및 X_2'는 각각 제 1행렬양자화기(620_1)로부터의 양자화에러 및 양자화값이다.

가중거리는 행렬양자화를 위한 행렬양자화부(MQ_2)(632)로 보내진다. 행렬양자화에 의해 출력된 8-비트의 지수는 가산기(631)에서 2프레임의 양자화에러에서 감산된다. 가중거리 계산부(633)는 가산기(631)의 출력을 이용하여 가중거리를 계속해서 계산한다. 가중거리를 최소화하는 양자화값이 선택된다. 가산기(631)의 출력은 제 1벡터양자화기(640_1)의 가산기(651, 661)로 한 프레임씩 차례대로 보내진다.

제 1벡터양자화기(640_1)는 한 프레임씩 차례대로 벡터양자화를 수행한다. 가산기(631)의 출력은 한 프레임씩 차례대로 가산기(651, 661)를 통해서 각각의 가중거리 계산부(653, 663)로 보내져서, 최소의 가중거리를 계산하게 된다.

양자화에러(X_2)와 양자화에러(X_2')간의 차는 (10×2)행렬이다. 만일 그 차가 $X_2 - X_2' = [x_{3-1}, x_{3-2}]$와 같이 나타나는 경우, 제 1벡터양자화기($640_1$)의 벡터양자화부(652, 662)에 의한 코드북탐색동안의 왜곡측정(d_{VQ1}, d_{VQ2})은 수학식 4 및 5에 의해 주어진다.

$$d_{VQ1}(x_{3-1}, x_{3-1}') = \sum_{i=1}^{P} w(0,i)(x_{3-1}(0,i) - x_{3-1}'(0,i))^2$$

$$d_{VQ2}(x_{3-2}, x_{3-2}') = \sum_{i=1}^{P} w(1,i)(x_{3-2}(1,i) - x_{3-2}'(1,i))^2$$

가중거리는 벡터양자화를 위한 벡터양자화부(VQ_1)(652)와 벡터양자화부(VQ_2)(662)로 보내진다. 이러한 벡터양자화에 의해 출력된 각 8-비트의 지수는 신호절환기(690)로 보내진다. 양자화값은 2프레임의 입력양자화에러벡터로부터 가산기(651, 661)에 의해 감산된다. 가중거리 계산부(653, 663)는 가산기(651, 661)의 출력을 이용하여, 가중거리를 연속적으로 계산하여, 가중거리를 최소화하는 양자화값을 선택한다. 가산기(651, 661)의 출력은 제 2벡터양자화기(640_2)의 가산기(671, 681)로 보내진다.

$$\underline{x}_{4-1} = x_{3-1} - \underline{x}_{3-1}'$$
$$\underline{x}_{4-2} = x_{3-2} - \underline{x}_{3-2}'$$

에 대한, 제 2벡터양자화기(640_2)의 벡터양자화기(672, 682)에 의한 코드북 탐색동안의 왜곡측정(d_{VQ3}, d_{VQ4})은 수학식 6, 7에 의해 주어진다.

$$d_{VQ3}(\underline{x}_{4-1}, \underline{x}_{4-1}') = \sum_{i=1}^{P} w(0,i)(x_{4-1}(0,i) - x_{4-1}'(0,i))^2$$

$$d_{VQ4}(\underline{x}_{4-2}, \underline{x}'_{4-2}) = \sum_{i=1}^{P} w(1,i)(x_{4-2}(1,i) - x'_{4-2}(1,i))^2$$

이러한 가중거리는 벡터양자화를 위해 벡터양자화기(VQ_3)(672)와 벡터양자화기(VQ_4)(682)로 보내진다. 벡터양자화로부터의 8-비트의 지수데이터는 가산기(671, 681)에 의해 2프레임에 대한 입력 양자화에러벡터로부터 감산된다. 가산거리계산부(673, 683)는 가산기(671, 681)의 출력을 이용하여 가중거리를 연속해서 계산하여, 가중거리를 최소화하는 양자화값을 선택하게 된다.

코드북 학습동안에는, 학습(learning)은 각각의 왜곡측정을 기초로 범용의 로이드 알고리듬(Lloyd algorithm)에 의해 수행된다.

코드북 탐색동안의 왜곡측정과 학습동안의 왜곡측정은 서로 다른 값이 될 수도 있다.

행렬양자화부(622, 632) 및 벡터양자화부(652, 662, 672 및 682)로부터의 8-비트의 지수데이터는 신호절환기(690)에 의해 절환되어 출력단자(691)에서 출력된다.

특히, 제 1행렬양자화단계를 수행하는 제 1행렬양자화기(620_1)의 출력과, 제 2행렬양자화단계를 수행하는 제 2행렬양자화기(620_2)의 출력과, 제 1벡터양자화단계를 수행하는 제 1벡터양자화기(640_1)의 출력은 저속의 비트속도로 도출되는 한편, 그러한 낮은 비트속도로의 출력은 제 2벡터양자화단계를 수행하는 제 2벡터양자화기(640_2)의 출력에 합해져서, 그 결과의 합이 고속의 비트속도로 도출된다.

상기 결과 합은 2kbps 및 6kbps에 대해 각각 32비트/40msec의 지수 및 48비트/40msec의 지수를 출력한다.

행렬양자화부(620)와 벡터양자화부(640)는 LPC계수를 나타내는 파라미터의 특성에 따라 주파수축 및/또는 시간축으로 제한된 가중(weighting)을 수행한다.

LSP파라미터의 특성에 따라 주파수축으로 제한되는 가중을 우선 설명한다. 만일 차수 P = 10이면, LSP파라미터(X(i))는 다음과 같이 저, 중, 고범위의 3가지 범위로 그룹지워진다.

$$L_1 = \{X(i) \mid 1 \le i \le 2\}$$

$$L_2 = \{X(i) \mid 3 \le i \le 6\}$$

$$L_3 = \{X(i) \mid 7 \le i \le 10\}$$

만일, 그룹(L_1, L_2 및 L_3)의 가중(weighting)이 각각 1/4, 1/2 및 1/4이면, 주파수축으로만 제한되는 가중은 다음의 수학식 8, 9 및 10으로 주어진다.

$$w'(i) = \frac{w(i)}{\sum_{j=1}^{2} w(j)} \times \frac{1}{4}$$

$$w'(i) = \frac{w(i)}{\sum_{j=3}^{6} w(j)} \times \frac{1}{2}$$

$$w'(i) = \frac{w(i)}{\sum_{j=7}^{10} w(j)} \times \frac{1}{4}$$

각 LSP파라미터의 가중은 각각의 그룹에서만 수행되며, 그러한 무게는 각 그룹에 대한 가중에 의해 제한된다.

시간축방향으로 볼때, 각각의 프레임의 총합은 반드시 1이므로, 시간축방향으로의 제한은 프레임기준이다. 시간축방향으로만 제한되는 무게는 다음의 수학식 11로 주어진다.

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{10} \sum_{s=0}^{1} w(j,s)}$$

여기서 $1 \le i \le 10$ 이고, $0 \le t \le 1$이다.

이 수학식 11에 의해, 주파수축방향으로는 제한되지 않는 가중이 t = 0 및 t = 1의 프레임수를 갖는 두개의 프레임사이에서 수행된다. 시간축방향으로만 제한되는 이러한 가중은 행렬양자화로 처리된 두개의 프레임사이에서 수행된다.

학습하는 동안의, 총수(T)를 갖는 학습데이터로써 사용되는 총 프레임은 다음 수학식 12에 따라서 가중된다.

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{10} \sum_{s=0}^{T} w(j,s)}$$

여기서 $1 \le i \le 10$이고 $0 \le t \le T$이다.

주파수축방향 및 시간축방향으로 제한되는 가중을 설명한다. 만일 차수 P = 10이면, LSP파라미터(X(i, t))는 다음과 같이 저, 중, 고범위의 3가지 범위로 그룹지워진다.

$$L_1 = \{X(i, t) \mid 1 \le i \le 2, 0 \le t \le 1\}$$

$$L_2 = \{X(i, t) \mid 3 \le i \le 6, 0 \le t \le 1\}$$

$$L_3 = \{X(i, t) \mid 7 \le i \le 10, 0 \le t \le 1\}$$

만일, 그룹(L_1, L_2 및 L_3)의 가중(weighting)이 각각 1/4, 1/2 및 1/4이면, 주파수축으로만 제한되는 가중은 다음의 수학식 13, 14 및 15로 주어진다.

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{2}\sum_{s=0}^{1} w(j,s)} \times \frac{1}{4}$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=3}^{6}\sum_{s=0}^{1} w(j,s)} \times \frac{1}{2}$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=7}^{10}\sum_{s=0}^{1} w(j,s)} \times \frac{1}{4}$$

이들 수학식 13, 14 및 15에 의해, 매 3프레임마다 주파수축으로 제한되고 행렬양자화로 처리되는 두개의 프레임을 가로지르는 가중이 수행된다. 이는 코드북탐색 및 학습중에 효과적이다.

학습동안의, 가중은 전체데이터의 총프레임용이다. LSP파라미터(X(i, t))는 다음과 같은 저, 중, 고범위으로 그룹지워진다.

$$L_1 = \{X(i, t) \mid 1 \le i \le 2, 0 \le t \le T\}$$

$$L_2 = \{X(i, t) \mid 3 \le i \le 6, 0 \le t \le T\}$$

$$L_3 = \{X(i, t) \mid 7 \le i \le 10, 0 \le t \le T\}$$

만일, 그룹(L_1, L_2 및 L_3)의 가중(weighting)이 각각 1/4, 1/2 및 1/4이면, 주파수축으로만 제한되는 그룹(L_1, L_2 및 L_3)의 가중(weighting)은 다음의 수학식 16, 17 및 18로 주어진다.

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=1}^{2}\sum_{s=0}^{T} w(j,s)} \times \frac{1}{4}$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=3}^{6}\sum_{s=0}^{T} w(j,s)} \times \frac{1}{2}$$

$$w'(i,t) = \frac{w(i,t)}{\sum_{j=7}^{10}\sum_{s=0}^{T} w(j,s)} \times \frac{1}{4}$$

이들 수학식 16, 17 및 18에 의해, 주파수축으로의 3범위에 대한 가중이 수행될 수 있어, 시간축방향으로의 총 프레임을 가로지르게 된다.

또한, 행렬양자화부(620)와 벡터양자화부(640)는 LSP파라미터에 있어서의 변화도에 따라서 가중을 수행한다. 전체 음성프레임중에서 하위 프레임을 나타내는 V→UV 또는 UV→V로의 과도영역에 있어서, LSP파라미터는 자음과 모음사이의 주파수응답에 있어서의 차이로 인해 엄청나게 변한다. 따라서, 수학식 19에 나타낸 가중은 과도영역상에 강세를 가중하여 위치시키기 위해서 가중(W'(i, t))에 의해 승산될 수도 있다.

$$wd(t) = \sum_{i=1}^{10} |x_1(i,t) - x_1(i,t-1)|^2$$

이 수학식 19대신에 다음의 수학식 20을 사용할 수도 있다.

$$wd(t) = \sum_{i=1}^{10} \sqrt{|x_1(i,t) - x_1(i,t-1)|}$$

따라서, LSP양자화부(134)는 2단의 행렬양자화 및 2단의 벡터양자화를 실행하여 다양한 출력지수의 비트수를 제공하게 된다.

벡터양자화부(116)의 기본 구조를 도 7에 나타낸 한편, 도 7에 나타낸 벡터양자화부(116)에 대한 보다 상세한 구조는 도 8에 나타내었다. 이제, 벡터양자화부(116)에서 스펙트럼포락선(Am)에 대한 가중벡터양자화의 도시적인 구조를 설명한다.

우선, 도 3에 나타낸 음성신호 부호화장치에 있어서, 스펙트럼평가부(148)의 출력측 또는 벡터양자화부(116)의 입력측상의 스펙트럼포락선의 일정수의 진폭데이터를 제공하기 위한 데이터수변환용의 도시적구성을 설명한다.

그러한 데이터수변환을 위한 방법에는 여러가지 방법이 고려될 수 있다. 본 실시예에 있어서는, 블록의 최종데이터로부터 블록의 최초데이터까지의 값을 보간처리하는 더미데이터(dummy data) 또는 그외 블록내의 최종데이터 또는 최초데이터를 반복하는 데이터 등의 소정의 데이터를 주파수축상의 유효대역의 1블록의 진폭데이터에 부가하여 데이터수를 N_F로 강화시키며, Os배(8배)수와 같은 수의 진폭데이터를 예를들어, FIR필터에 의한 제한된 대역폭형의 8폴드오버샘플링 등의 Os폴드를 이용하여 구한다. (mMx + 1) ×Os진폭데이터는 2048등의 보다 큰 N_M수로의 확장을 위해 선형보간된다. 이 N_M데이터는 상술한 44데이터 등의 소정수(M)의 데이터로의 변환을 위해 서브샘플처리된다.

실제, 최종으로 요구되는 M데이터를 공식화하기 위해서는 필요한 데이터만을 상술한 N_M데이터를 구하지 않고서 오버샘플링처리 및 선형보간처리를 이용해 계산한다.

도 7의 가중벡터양자화를 수행하기 위한 벡터양자화부(116)는 적어도 제 1벡터양자화단계를 수행하기 위한 제 1벡터양자화부(500)와, 제 1벡터양자화부(500)에 의한 제 1벡터양자화동안에 생성된 양자화에러벡터를 양자화시키기 위한 2벡터양자화단계를 수행하기 위한 제 2벡터양자화부(510)를 포함하고 있다. 이러한 제 1벡터양자화부(500)를 소위 제 1단 벡터양자화부라 하는 한편, 제 2벡터양자화부(510)를 소위 제 2단 벡터양자화부라 한다.

스펙트럼 평가부(148)의 출력벡터(x), 즉 소정수(M)를 갖는 포락선데이터는 제 1벡터양자화부(500)의 입력단자(501)로 입력된다. 이 출력벡터(x)는 벡터양자화부(502)에 의한 가중 벡터양자화로 양자화된다. 따라서 벡터양자화부(502)에 의해 출력된 형상지수는 출력단자(503)로 출력되는 한편, 양자화값(x_0')은 출력단자(504)로 출력되어 가산기(505, 513)로 보내진다. 가산기(505)는 소스벡터(x)에서 양자화값(x_0')을 감산하여 다차수의 양자화에러벡터(y)를 산출한다.

양자화에러벡터(y)는 제 2벡터양자화부(510)내의 벡터양자화부(511)로 보내진다. 이 제 2벡터양자화부(511)는 복수의 벡터양자화부 또는 도 7에 나타낸 두개의 벡터양자화기(511_1, 511_2)로 이루어진다. 양자화에러벡터(y)는 두개의 벡터양자화기(511_1, 511_2)에서 가중벡터양자화에 의해 양자화되도록 차원적으로 분할된다. 이러한 벡터양자화기(511_1, 511_2)에 의해 출력된 형상지수는 출력단자(512_1, 512_2)에서 출력되는 한편, 양자화값(y_1', y_2')은 차원적인 방향으로 접속되어 가산기(513)로 보내진다. 가산기(513)는 양자화값(y_1', y_2')을 양자화값(x_0')에 더하여 양자화값($\underline{x}_1'$)을 생성하여 출력단자(514)로 출력한다.

따라서, 제 1벡터양자화부(500)에 의한 제 1벡터양자화단계의 출력은 저속의 비트속도로 도출되는 반면, 제 1벡터양자화단계 및 제 2양자화부(510)에 의한 제 2양자화단계의 출력은 고속의 비트속도로 출력된다.

특히, 벡터양자화부(116)내의 제 1벡터양자화부(500)내에 있는 벡터양자화기(502)는 도 8에 나타낸 바와같이 44차수의 2단구조 등의 L차수의 것이다.

즉, 이득(g_i)으로 승산된 32의 코드북사이즈를 갖는 44차수의 벡터양자화 코드북의 출력벡터의 합은 44차수의 스펙트럼 포락선벡터(x)의 양자화값(x_0')으로 사용된다. 따라서, 도 8에 나타낸 바와같이, 두개의 코드북은 CB0, CB1인 한편, 출력벡터는 s_{1i}, s_{1j}(여기서 $0 \leq i$이고 $j \leq 31$)이다. 한편, 이득코드북(CB_g)의 출력은 g_1(여기서, $0 \leq 1 \leq 31$이고, g_1는 척도이다)이다. 최종의 출력(x_0')은 $g_1(s_{1i} + s_{1j})$이다.

상기 LPC 오차(residual)의 MBE분석에 의해 얻어져서 소정의 차수로 변환되는 스펙트럼포락선은 x이다. x가 얼마나 효과적으로 양자화되는지는 굉장히 중요한 것이다.

양자화에러에너지(E)는 다음 수학식 21로 정의된다.

$$E = \| W \{Hx - Hg_1((s_{0i} + s_{1j})\}\|^2$$

$$= \| WH \{x - \{x - g_1(s_{0i} + s_{1j})\}\|^2$$

여기서, H는 LPC합성필터의 주파수축상의 특성을 나타내고, W는 주파수축상에 청각가중을 위한 특성을 나타내는 가중용 행렬이다.

현재프레임의 LPC분석의 결과에 의한 α-파라미터가 $\alpha_i(1 \leq i \leq P)$로 표시되면, L차수의 값, 예를들어 44차수에 상응하는 포인트가 수학식 22의 주파수응답에서 샘플처리된다.

$$H(z) = \frac{1}{1 + \sum_{i=1}^{P} \alpha_i z^{-i}}$$

계산을 위해, 0s을 1, α_1, α_2, $\cdots\alpha_p$의 열에 이어 채워넣어서 1, α_1, α_2, $\cdots\alpha_p$, 0, 0, …, 0의 열로 하여, 예를들어, 256포인트 데이터를 얻는다. 그런다음, 256포인트 FFT에 의해, $(r_e^2 + im^2)^{1/2}$를 0~π까지의 범위와 연합된 포인트에 대해 계산하여, 그 결과의 역을 구한다. 이러한 역을 44포인트 등의 L포인트에서 서브샘플처리하여, 다각원소로써 이러한 L포인트를 갖는 행렬을 형성한다.

$$H = \begin{bmatrix} h(1) & & 0 \\ & h(2) & \\ & & \ddots \\ 0 & & h(L) \end{bmatrix}$$

청각가중행렬(W)은 수학식 23으로 주어진다.

$$W(z) = \frac{1 + \sum_{i=1}^{P} \alpha_i \lambda_b^i z^{-i}}{1 + \sum_{i=1}^{P} \alpha_i \lambda_a^i z^{-i}}$$

여기서, α_i는 LPC분석의 결과이며, λa, λb는 상수로써 λa = 0.4이고 λb = 0.9이다.

행렬(W)은 상기 수학식 23의 주파수응답으로부터 계산될 수도 있다. 예를들어, FFT를 1, α1λb, $\alpha 2\lambda b^2$, $\cdots \alpha p \lambda b^p$, 0, 0, …, 0의 256포인트 데이터로 수행하여 0에서 π까지의 영역에 대해 $(r_e^2[i] + Im^2[i])^{1/2}$(여기서 $0 \leq i \leq 128$)를 얻는다.

수학식 23의 주파수응답은 다음식에 의해 얻어질 수도 있다.

$$w_0[i] = \frac{\sqrt{re^2[i] + im^2[i]}}{\sqrt{re'^2[i] + im'^2[i]}}$$

여기서, $0 \le i \le 128$이다. 이것은 예를들어, 44차수의 벡터의 연합포인트마다 다음 방법으로 구해진다. 보다 명확히 말하면, 선형보간처리가 사용될 수 있다. 그렇지만, 다음예에 있어서는, 가장 근접한 포인트를 대신 사용한다.

즉,

$\omega[i] = \omega 0[\mathrm{nint}\{128i/L\}]$, 여기서 $1 \le i \le L$이다.

본 수학식에서 nint(X)는 값을 X로 근접시키는 함수이다.

H, h(1), h(2), …h(L)에 대해서도 동일한 방법을 이용해 구한다.

즉,

$$H = \begin{bmatrix} h(1) & & 0 \\ & h(2) & \\ & & \ddots \\ 0 & & h(L) \end{bmatrix} \quad W = \begin{bmatrix} w(1) & & 0 \\ & w(2) & \\ & & \ddots \\ 0 & & w(L) \end{bmatrix}$$

$$WH = \begin{bmatrix} h(1)w(1) & & & 0 \\ & h(2)w(2) & & \\ & & \ddots & \\ 0 & & & h(L)w(L) \end{bmatrix}$$

또다른 예로써, H(z)W(z)를 우선 구한 다음 주파수응답을 구하여 FFT의 회수를 감소시킨다.

$$H(z)W(z) = \frac{1}{1 + \sum_{i=1}^{P} \alpha_i z^{-i}} \cdot \frac{1 + \sum_{i=1}^{P} \alpha_i \lambda_b^i z^{-i}}{1 + \sum_{i=1}^{P} \alpha_i \lambda_a^i z^{-i}}$$

즉, 수학식 25의 분모는

$$\left(1 + \sum_{i=1}^{P} \alpha_i z^{-i}\right)\left(1 + \sum_{i=1}^{P} \alpha_a^i \lambda_a^i z^{-i}\right) = 1 + \sum_{i=1}^{2P} \beta_i z^{-i}$$

로 확장된다. 예를들어, 256포인트 데이터를 1, β_1, β_2, …, β_{2P}, 0, 0, …, 0의 열을 사용하여 생성한다. 그런다음, 256포인트 FFT를 진폭의 주파수응답이

$$rms[i] = \sqrt{re''^2[i] + im''^2[i]}$$

이 되는 상태로 수행한다. 여기서, $0 \leq i \leq 128$이다. 이로부터,

$$wh_0[i] = \frac{\sqrt{re^2[i] + im^2[i]}}{\sqrt{re''^2[i] + im''^2[i]}}$$

여기서 $0 \leq i \leq 128$이다. 이것을 L차원벡터의 각각의 상응하는 포인트마다 구한다. 만일 FFT의 포인트의 수가 작은 경우에는, 선형보간처리가 사용될 수 있다. 그렇지만 최대근사치는 여기서,

$$wh[i] = wh_0[n\ int(\frac{128}{L} \cdot i)]$$

에 의해 구해지며, 여기서 $1 \leq i \leq L$이다. 이러한 대각원소를 갖는 행렬은 W'이다.

$$W' = \begin{bmatrix} wh(1) & & 0 \\ & wh(2) & \\ & & \ddots \\ 0 & & wh(L) \end{bmatrix}$$

수학식 26은 수학식 24와 동일한 행렬을 나타내고 있다.

선택적으로, $|H(\exp(j\omega))W(\exp(j\omega))|$는 wh[i]로 사용되기 위해 $\omega = i/L\lambda$에 대해 수학식 25에서 직접 구해진다. 또한 선택적으로, 수학식 25의 임펄스응답은 64와 같은 적정한 길이로 구해져서 FFT되어 wh[i]로 사용될 진폭주파수특성을 구하게 된다.

이러한 행렬 즉, 가중합성필터의 주파수응답을 이용하여 수학식 21을 다시 쓰면, 우리는 수학식 27을 얻는다.

$$E = \|W'(x - g_1((s_{0i} + s_{1i}))\|^2$$

형상코드북과 이득코드북을 학습하는 방법을 설명한다.

왜곡에 대한 기대치는 코드벡터($s0_c$)가 CB0용으로 선택되는 모든 프레임(k)에 대해 최소화된다. 만일 그러한 M프레임이 존재하는 경우에는, 수학식 28이 최소화되면 충분하다.

$$J = \frac{1}{M}\sum_{k=1}^{M} \|W_k'(x - g_k(s_{0c} + s_{ik})\|^2$$

수학식 28에서, W_k', X_k, g_k 및 s_{ik}는 각각 k'번째 프레임에 대한 가중, k'번째 프레임으로의 입력, k'번째 프레임의 이득 및 k'번째 프레임에 대한 코드북(CB1)의 출력을 나타낸다.

수학식 28을 최소화해보면,

$$J=\frac{1}{M}\sum_{k=1}^{M}\{(x_k^T-g_k(s_{0c}^T+s_{1k}^T))W_k'^TW_k'(x_k-g_k(s_{0c}+s_{1k}))\}$$
$$=\frac{1}{M}\sum_{k=1}^{M}\{x_k^TW_k'^TW_k'x_k-2g_k(s_{0c}^T+s_{1k}^T)W_k'^TW_k'x_k$$
$$+g_k^2(s_{0c}^T+s_{1k}^T)W_k'^TW_k'(s_{0c}+s_{1k})\}$$
$$=\frac{1}{M}\sum_{k=1}^{M}\{x_k^TW_k'^TW_k'x_k-2g_k(s_{oc}^T+s_{1k}^T)W_k'^TW_k'x_k$$
$$+g_k^2s_{oc}^TW_k'^TW_k's_{oc}+2g_k^2s_{oc}^TW_k'^TW_k's_{1k}+g_k^2s_{1k}^TW_k'^TW_k's_{1k}\}$$

$$\frac{\partial J}{\partial s_{0c}}=\frac{1}{M}\sum_{k=1}^{M}\{-2g_kW_k'^TW_k'x_k+2g_k^2W_k'^TW_k's_{0c}+2g_k^2W_k'^TW_k's_{1k}\}=0$$

그러므로,

$$\sum_{k=1}^{M}(g_kW_k'^TW_k'x_k-g_k^2W_k'^TW_k's_{lk})=\sum_{k=1}^{M}g_k^2W_k'^TW_k's_{0c}$$

그 결과,

$$s_{0c}=\left\{\sum_{k=1}^{M}g_k^2W_k'^TW_k'\right\}^{-1}\cdot\left\{\sum_{k=1}^{M}g_kW_k'^TW_k'(x-g_ks_{lk})\right\}$$

여기서 0는 역행렬을 나타내고 $W_K'^T$는 W_K'의 이항행렬을 나타낸다.

이어서, 이득최적화를 고려한다.

이득의 코드워드 gc를 선택하는 k'번째 프레임에 관한 왜곡의 기대치는 수학식 32로 주어진다.

$$J_g=\frac{1}{M}\sum_{k=1}^{N}\|W_k'(x_k-g_c(s_{0k}+s_{1k}))\|^2$$
$$=\frac{1}{M}\sum_{k=1}^{M}\{x_k^TW_k'^TW_k'x_k-2g_cx_k^TW_k'^TW_k'(s_{0k}+s_{1k})$$
$$-g_c^2(s_{0k}^T+s_{1k}^T)W_k'^TW_k'(s_{0k}+s_{1k})\}$$

풀이하면

$$\frac{\partial J_g}{\partial g_c}=\frac{1}{M}\sum_{k=1}^{M}\{-2x_k^TW_k'^TW_k'(s_{0k}+s_{1k})$$
$$-2g_c(s_{0k}^T+s_{1k}^T)W_k'^TW_k'(s_{0k}+s_{1k})\}=0$$

우리는

$$\frac{\partial J_g}{\partial g_c} = \frac{1}{M}\sum_{k=1}^{M}\{-2x_k^T W_k'^T W_k'(s_{0k}+s_{1k}) -2g_c(s_{0k}^T+s_{1k}^T)W_k'^T W_k'(s_{0k}+s_{1k})\} = 0$$

와

$$g_c = \frac{\sum_{k=1}^{M} x_k^T W_k'^T W_k'(s_{0k}+s_{1k})}{\sum_{k=1}^{M}(s_{0k}^T+s_{1k}^T)W_k'^T W'(s_{0k}+s_{1k})}$$

를 얻는다.

상기 수학식 31과 수학식 32는 형상(s_{0i}, s_{1i}), 및 최적의 디코더출력인 $0 \le i \le 31$ 에 대한 이득(gi)에 대한 최적의 중심조건을 제공한다. 한편, s_{1i}도 s_{0i}와 동일한 방법으로 구해질 수 있다.

최적의 부호화조건 즉, 가장 근접하는 조건을 고려한다.

상기 왜곡측정, 즉, 식 $E = \| W'(X-gl(s_{1i}+s_{1j}))\|^2$를 최소화하는 s_{0i}과 s_{1i}을 구하기 위한 수학식 27은 입력(x)와 가중행렬(W')이 주어질 때마다 즉, 한 프레임씩 차례로 구해진다.

본래, E는 $gl(0 \le 1 \le 31)$, $s_{0i}(0 \le i \le 31)$ 및 $s_{0j}(0 \le i \le 31)$의 총합 즉, $32\times32\times32 = 32768$에 대해 사발통문(round robin)식으로 구해져서, E의 최소치를 제공할 한세트의 s_{0i}, s_{1i}를 구하게 된다. 그렇지만, 이는 방대한 계산이 요구되기 때문에, 본 실시예에 있어서는 형상과 이득을 연속적으로 탐색한다. 한편, s_{0i}, s_{1i}의 조합을 위해서도 사발통문식의 탐색을 사용한다. s_{0i}, s_{1i}에 대해서는 $32\times32 = 1024$개의 조합이 있다. 다음 설명에서는, $s_{1i} + s_{1j}$는 간결성을 위해 s_m으로 나타낸다.

상기 수학식 27은 $E = \| W'(x-glsm)\|^2$이 된다. 보다 간략하게 한다면, $x_w = W'x$이고 $s_w = W' s_m$일 때, 우리는 다음 수학식 33과 34를 얻는다.

$$E = \| x_w - g_l s_w \|^2$$

$$E = \|x_w\|^2 + \|s_w\|^2\left(g_l - \frac{x_w^T \cdot s_w}{\|s_w\|^2}\right)^2 - \frac{(x_w^T \cdot s_w)^2}{\|s_w\|^2}$$

따라서, gl을 충분히 정확하게 할 수 있으면, 탐색은 다음의 두가지 단계로 수행될 수 있다.

(1)

$$\frac{(x_w^T \cdot s_w)^2}{\|s_w\|^2}$$

를 최소화하게 될 s_w에 대한 탐색단계

(2)

$$\frac{x_w^T \cdot z_w}{\|z_w\|^2}$$

에 근접하는 g_1에 대한 탐색단계

상기 내용을 본래의 표기법을 이용해 다시 쓰면,

(1)'

$$\frac{\left(x^T W'^T W'(z_{0i} + z_{1j})\right)^2}{\|W'(z_{0i} + z_{1j})\|^2}$$

을 최소화할 s01 및 s_{1i}의 세트에 대한 탐색과,

(2)'

$$\frac{x^T W'^T W'(z_{0i} + z_{1j})}{\|W'(z_{0i} + z_{1j})\|^2}$$

에 가장 근접하는 g_1에 대한 탐색이 이루어진다.

상기 수학식 35는 최적의 부호화조건(가장근접한 조건)을 나타낸다.

수학식 31과 수학식 32의 조건(중심조건) 및 수학식 35의 조건을 이용하여, 코드북(CB0, CB1 및 CBg)을 소위 범용화된 로이드 알고리듬(GLA)으로 동시에 연습할 수 있다.

본 실시예에 있어서는, 입력 x의 노르마(norm)에 의해 나누어진 W'가 W'로 사용된다. 즉, W'/‖x‖는 수학식 31, 32 및 35에서 W'로 치환된다.

선택적으로, 벡터양자화기(116)에 의한 벡터양자화시에 청각가중용으로 사용된 가중(W')은 상기 수학식 26에 의해 정의된다. 그렇지만, 일시적인 마스킹(masking)을 고려하여, 그러한 가중(W')도 현재 가중(W')(여기서 지난 W'는 이미 고려되어 있다)을 구함으로써 구해질 수 있다.

시간(n)에서, 즉 n번째 프레임에서 구해진 상기 수학식 26에서 wh(1), wh(2), …, wh(L)의 값은 각각 whn(1), whn(2), …, whn(L)로 표시된다.

과거값을 고려하는 시간(n)에서의 무게를 An(i)(여기서 $1 \le i \le L$)로 정의하면,

$$An(i) = \lambda_{n-1} A(i) + (1-\lambda)\, whn(i), \ (whn(i) \le A_{n-1}(i))$$

$$= whn(i), \ (whn(i) > A_{n-1}(i))$$

여기서, λ는 예를들어 λ = 0.2로 설정될 수 있다. An(i)에서는, $1 \le i \le L$로, 그렇게 구해진, 대각원소로써 그러한 An(i)를 갖는 행렬이 상기 가중으로써 사용될 수 있다.

이와같은 식으로 가중벡터양자화에 의해 얻어진 형상지수값(s_{0i}, s_{1i})은 출력단자(520, 522)로 각각 출력되는 한편, 이득지수(gl)은 출력단자(521)로 출력된다. 또한, 양자화값(x_0')은 출력단자(504)로 출력되는 동시에 가산기(505)로 보내진다.

가산기(505)는 스펙트럼포락선벡터(x)에서 양자화값을 감산하여 양자화에러벡터(y)를 생성한다. 특히, 이러한 양자화에러벡터(y)는 벡터양자화부(511)로 보내져서 차원적으로 분할되게 되어 가중벡터양자화로 벡터양자화기(511_1~511_8)에 의해 양자화된다.

제 2벡터양자화부(510)는 제 1벡터양자화부(500)보다 큰수의 비트수를 사용한다. 결과적으로, 코드북의 메모리용량 및 코드북탐색에 대한 처리볼륨(복잡성)도 굉장히 증가된다. 따라서, 제 1벡터양자화부(500)의 차수와 동일한 44차수를 갖는 벡터양자화를 수행할 수 있게 된다. 따라서, 제 2벡터양자화부(510)내의 벡터양자화부(511)는 복수의 벡터양자화기들로 이루어지며, 그 입력양자화값은 복수의 저차원의 벡터들로 분할되어 가중벡터양자화를 수행하게 된다.

벡터양자화기(511_1~511_8)에서 사용된 양자화값(y_0~y_7)과, 차원수 및 비트수들간의 관계를 다음의 표 1에 나타내었다.

표 1.

양 자 화 값	차 원 수	비 트 수
y_0	4	10
y_1	4	10
y_2	4	10
y_3	4	10
y_4	4	9
y_5	8	8
y_6	8	8
y_7	8	7

벡터양자화기(511_1~511_8)에서 출력된 지수값(Id_{vq0}~Id_{vq7})은 출력단자(523_1~523_8)에서 출력된다. 이러한 지수데이터의 비트합은 72이다.

차원방향으로의 벡터양자화기(511_1~511_8)의 출력양자화값(y_0', y_7')을 접속함으로써 얻어진 값이 y'이면, 양자화값(y', x_0')은 가산기(513)에 의해 합산되어 양자화값(x_1')을 제공한다. 따라서, 양자화값(x_1')은 다음과 같이 표시된다.

$$\begin{aligned} x_1' &= x_0' + y' \\ &= x - y + y' \end{aligned}$$

즉, 최종적인 양자화에러벡터는 y' - y이다.

제 2벡터양자화기(510)로부터 양자화값(x_1')이 복호화되게 되면, 음성신호복호화장치는 제 1양자화부(500)로부터의 양자화값(x_1')에 있어서는 필요치 않다. 그렇지만, 제 1양자화부(500) 및 제 2양자화부(510)로부터의 지수데이터에 있어서는 필요하다.

벡터양자화부(511)에서의 학습방법 및 코드북탐색을 이하 설명한다.

학습방법에 대해서는, 양자화에러벡터(y)는 표 1에 나타낸 바와같이, 무게(W')를 사용하여, 8개의 저차수벡터(y_0, y_7)로 분할된다. 무게(W')가 다음의 수학식 36과 같이 대각원소로써 44포인트의 서브샘플처리된 값을 갖는 행렬인 경우,

$$W' = \begin{bmatrix} wh(1) & & & 0 \\ & wh(2) & & \\ & & \ddots & \\ 0 & & & wh(44) \end{bmatrix}$$

무게(W')는 다음의 8개의 행렬로 분할된다.

$$W_1' = \begin{bmatrix} wh(1) & & 0 \\ & \ddots & \\ 0 & & wh(4) \end{bmatrix}$$

$$W_2' = \begin{bmatrix} wh(5) & & 0 \\ & \ddots & \\ 0 & & wh(8) \end{bmatrix}$$

$$W_3' = \begin{bmatrix} wh(9) & & 0 \\ & \ddots & \\ 0 & & wh(12) \end{bmatrix}$$

$$W_4' = \begin{bmatrix} wh(13) & & 0 \\ & \ddots & \\ 0 & & wh(16) \end{bmatrix}$$

$$W_5' = \begin{bmatrix} wh(17) & & 0 \\ & \ddots & \\ 0 & & wh(20) \end{bmatrix}$$

$$W_6' = \begin{bmatrix} wh(21) & 0 & \\ & \ddots & \\ 0 & & wh(28) \end{bmatrix}$$

$$W_7' = \begin{bmatrix} wh(29) & & 0 \\ & \ddots & \\ 0 & & wh(36) \end{bmatrix}$$

$$W_8' = \begin{bmatrix} wh(37) & & 0 \\ & \ddots & \\ 0 & & wh(44) \end{bmatrix}$$

저차원들로 분할된 y와 W'를 Y_i와 W_i'(여기서 $1 \leq i \leq 8$)로 각각 명명한다.

왜곡측정(E)은 다음 수학식 37로 정의된다.

$$E = \| w_I' \ (y_i - s) \|^2$$

코드북벡터(s)는 y_i에 대한 양자화결과이다. 왜곡측정(E)을 최소화하는 그러한 코드북의 코드벡터를 탐색한다.

코드북학습에 있어서는, 또한 범용의 로이드 알고리듬(GLA)을 이용해 가중이 행해진다. 학습에 대한 최적의 중심조건을 먼저 설명한다. 최적의 양자화결과로서 코드벡터(s)를 선택하고 있는 M입력벡터(y)가 존재하고, 연습데이터가 (y_k)인 경우, 왜곡의 기대치(J)는 전체 프레임(k)에 대해 가중을 행하여 왜곡의 중심을 최소화하는 다음의 수학식 38에 의해 주어진다.

$$\begin{aligned} J &= \frac{1}{M}\sum_{k=1}^{M} \| W_k'(y_k - s) \|^2 \\ &= \frac{1}{M}\sum_{k=1}^{M} (y_k - s)^T W_k'^T W_k' (y_k - s) \\ &= \frac{1}{M}\sum_{k=1}^{M} y_k^T W_k'^T W_k' y_k - 2 y_k^T W_k'^T W_k' s \\ &\quad + s^T W_k'^T W_k' s \end{aligned} \qquad 0$$

$$\frac{\partial J}{\partial s} = \frac{1}{M}\sum_{k=1}^{M} (-2 y_k^T W_k'^T W_k' + 2 s^T W_k'^T W_k') = 0$$

풀이하면, 우리는

$$\sum_{k=1}^{M} y_k^T W_k'^T W_k' = \sum_{k=1}^{M} s^T W_k'^T W_k'$$

을 얻는다. 양측의 이항값을 취하여, 우리는

$$\sum_{k=1}^{M} W_k'^T W_k' y_k = \sum_{k=1}^{M} W_k'^T W_k' s$$

을 얻는다. 따라서, 수학식 39와 같이 된다.

$$s = \left(\sum_{k=1}^{M} W_k'^T W_k' \right)^{-1} \sum_{k=1}^{M} W_k'^T W_k' y_k$$

상기 수학식 39에서, s는 최적의 대표벡터이며, 최적의 중심조건을 나타낸다.

최적의 부호화조건에 대해서는, $\| W_i'(y_i - s) \|^2$의 값을 최소화하는 s에 대한 탐색으로 충분하다. 탐색동안의 W_i'는 학습동안의 W_i'와 똑같을 필요는 없으며, 가중행렬이 아니어도 좋다.

$$\begin{bmatrix} 1 & & 0 \\ & 1 & \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

2단의 벡터양자화부를 이용해 음성신호 엔코더내에 벡터양자화부(116)를 구성함으로써, 다양한 출력지수비트의 수를 표현할 수 있게 된다.

본 발명의 상술한 CELP엔코더구조를 이용하는 제 2부호화부(120)은 도 9에 나타낸 바와같이 다단의 벡터양자화 프로세서로 구성되어 있다. 도 9의 실시예에 있어서, 이러한 다단의 벡터양자화 프로세서는 2단의 부호화부(120_1, 120_2)로 구성되며, 여기서 송신비트속도가 예를들어 2kbps와 6kbps와의 사이로 절환될 수 있는 경우에는 6kbps의 송신비트속도를 수습하기 위한 구조를 나타내고 있다. 또한, 형상 및 이득지수출력은 23비트/5msec 및 15비트/ 5msec사이로 절환될 수 있다. 도 9의 구조에 있어서의 처리흐름을 도 10에 나타내었다.

도 9를 참고하면, 도 9의 LPC분석회로(302)는 도 3에 나타낸 LPC분석회로(132)에 상응하는 한편, LSP파라미터 양자화회로(303)는 도 3의 α→LSP변환회로(133)에서 LSP→α변환회로(137)까지의 구성에 대응하며, 청각가중필터(304)는 도 3의 청각가중필터계산회로(139)와 청각가중필터(125)에 상응한다. 따라서, 도 9에 있어서, 도 3의 제 1부호화부(113)의 LSP→α변환회로(137)의 출력과 동일한 출력이 단자(305)로 인가되는 한편, 도 3의 청각가중필터계산회로(139)의 출력과 동일한 출력이 단자(307)로 인가되며, 도 3의 청각가중필터(125)의 출력과 동일한 출력이 단자(306)로 인가된다. 그렇지만, 청각가중필터(125)로부터의 판별에 있어서, 도 9의 청각가중필터(304)는 LSP→α변환회로(137)의 출력을 이용하는 대신에, 입력음성데이터와 예양자화 α-파라미터를 이용하여, 청각가중신호 즉, 도 3의 청각가중필터(125)의 출력과 동일한 신호를 생성한다.

도 9에 나타낸 2단의 제 2부호화부(120_1, 120_2)에 있어서, 감산기(313, 323)는 도 3의 감산기(123)와 대응하는 한편, 거리계산회로(314, 324)는 도 3의 거리계산회로(124)와 대응한다. 또한, 이득회로(311, 321)는 도 3의 이득회로(126)에 대응하는 한편, 확률적 코드북(310, 320) 및 이득 코드북(315, 325)은 도 3의 잡음코드북(121)에 대응한다.

도 9의 구성에 있어서, 도 10의 단계(S1)에서 LPC분석회로(302)는 단자(301)로부터 인가된 입력음성데이터(x)를 상술한 바와같이 프레임으로 분할하여 α-파라미터를 구하기 위해 LPC분석을 수행하게 된다. LSP파라미터 양자화회로(303)는 LPC분석회로(302)로부터의 α-파라미터를 LSP파라미터를 변환하여 LSP파라미터를 양자화시킨다. 양자화된 LSP파라미터는 보간처리되어서 α-파라미터로 변환된다. LSP파라미터 양자화회로(303)는 양자화된 LSP파라미터에서 변환된 α-파라미터 즉, 양자화된 LSP파라미터에서 LPC분석필터함수(1/H(z))를 생성하며, 생성된 LPC분석필터함수(1/H(z))를 단자(305)를 통해서 제 1단의 제 2부호화부(120_1)의 청각가중합성필터(312)로 보낸다.

청각가중필터(304)는 도 3의 청각가중필터 계산회로(139)에 의해 생성된 데이터와 동일한 청각가중용 데이터를 LPC분석회로(302)로부터의 α-파라미터 즉, 예양자화 α-파라미터에서 구한다. 이러한 데이터는 단자(307)를 통해서 제 1단의 제 2부호화부(120_1)의 청각가중합성필터(312)로 인가된다. 청각가중필터(304)는 도 3의 청각가중필터(125)에 의해 출력된 신호와 동일한 신호인 청각가중신호를, 도 10의 단계(S2)에 나타낸 바와같이, 입력음성데이터와 예양자화 α-파라미터로부터 생성한다. 즉, 우선 LPC합성필터함수(W(z))를 예양자화 α-파라미터로부터 생성한다. 그렇게 생성된 필터함수(W(z))를 입력음성데이터(x)에 가해서, 청각가중신호로써 단자(306)를 통해서 제 1단의 제 2부호화부(120)의 감산기(303)에 인가되는 xw를 생성한다.

제 1단의 제 2부호화부(120_1)에 있어서, 9비트의 형상지수출력의 확률적 코드북(310)의 대표값출력은 이득회로(311)로 보내지며, 그후 확률적 코드북(310)으로부터의 대표값출력을 6비트의 이득지수출력의 이득 코드북(315)으로부터의 이득(척도)과 승산된다. 이득회로(311)에 의한 이득과 승산된 대표값출력은 $1/A(z) = (1/H(z))^{*}W(z)$를 갖는 청각가중합성필터(312)로 보내진다. 도 10에 단계(S3)로 나타낸 바와같이, 가중합성필터(312)는 1/A(z)의 제로입력응답출력을 감산기(313)로 보낸다. 감산기(313)는 청각가중합성필터(312)의 제로입력응답출력과 청각가중필터(304)로부터의 청각가중신호(xw)에 대해 감산을 행하며, 그 결과의 차 또는 에러는 기준벡터(r)로써 도출된다. 제 1단의 제 2부호화부(120)에서의 탐색동안에는, 도 10에 나타낸 단계(S4)에 나타낸 바와같이, 기준벡터(r)가 거리계산회로(314)로 보내지며, 여기서 거리가 계산되고 양자화에러에너지(E)를 최소화하는 형상벡터(s) 및 이득(g)이 탐색된다. 여기서, 1/A(z)는 제로상태에 있게 된다. 즉, 제로상태의 1/A(z)로 합성된 코드북내의 형상벡터(s)가 s_{syn}인 경우에는, 다음의 수학식 40을 최소화하는 형상벡터(s) 및 이득(g)이 탐색된다.

$$E = \sum_{n=0}^{N-1} \left(r(n) - g s_{syn}(n)\right)^2$$

비록 양자화에러에너지(E)를 최소화하는 s 및 g가 완전히 탐색될 수도 있지만, 계산량을 감소시키기 위해 다음의 방법을 사용할 수도 있다.

제 1방법은 E를 최소화하는 형상벡터(s)를 탐색하는 방법으로, 다음의 수학식 41에 정의되어 있다.

$$E_s = \frac{\sum_{n=0}^{N-1} r(n)s_{syn}(n)}{\sqrt{\sum_{n=0}^{N-1} s_{syn}(n)^2}}$$

제 1방법에 의해 얻어진 s로부터의, 이상적인 이득은 수학식 42에 나타낸 바와같다.

$$g_{ref} = \frac{\sum_{n=0}^{N-1} r(n)s_{syn}(n)}{\sum_{n=0}^{N-1} s_{syn}(n)^2}$$

따라서, 제 2방법으로써는, 다음 수학식 43을 최소화하는 g가 탐색된다.

$$Eg = (g_{ref} - g)^2$$

E는 g의 2차함수이기 때문에, Eg를 최소화하는 g는 E를 최소화한다.

제 1 및 제 2방법에 의해 얻어진 s 및 g로부터, 양자화에러벡터(e)는 다음 수학식 44에 의해 계산될 수 있다.

$$e = r - gs_{syn}$$

이것은 제 1단에서 제 2단의 제 2부호화부(120_2)의 기준만큼 양자화된다.

즉, 단자(305, 307)로 인가된 신호는 제 1단의 제 2부호화부(120_1)의 청각가중합성필터(312)로부터 제 2단의 제 2부호화부(120_2)의 청각가중합성필터(322)로 곧바로 인가된다. 제 1단의 제 2부호화부(120_1)에 의해 구해진 양자화에러벡터(e)는 제 2단의 제 2부호화부(120_2)의 감산기(323)로 인가된다.

도 10의 단계(S5)에서, 제 1단에서 수행된 처리와 동일한 처리가 제 2단의 제 2부호화부(120_2)에서 발생한다. 즉, 5비트의 형상지수출력의 확률적 코드북(320)으로부터의 대표값출력은 이득회로(321)로 보내지며, 여기서 코드북(320)의 대표값출력이 3비트의 이득지수출력의 이득코드북(325)으로부터의 이득과 승산된다. 가중합성필터(322)의 출력은 감산기(323)로 보내지고, 여기서 청각가중합성필터(322)의 출력과 제 1단의 양자화에러벡터(e)간의 차가 구해진다. 이러한 차는 거리계산을 위한 거리계산회로(324)로 보내져서 양자화에러벡터(E)를 최소화하는 형상벡터(s) 및 이득(g)을 탐색하게 된다.

제 1단의 제 2부호화부(120_1)의 확률적 코드북(310)의 형상지수출력과 이득코드북(315)의 이득지수출력과, 제 2단의 제 2부호화부(120_2)의 확률적 코드북(320)의 지수출력과 이득코드북(325)의 지수출력은 지수출력절환회로(330)로 보내진다. 제 2의 부호화부(120)로부터 23비트가 출력되면, 제 1단 및 제 2단의 제 2부호화부(120_1, 120_2)의 확률적 코드북(310, 320)과 이득코드북(315, 325)의 지수데이터가 합해져서 출력된다. 15비트가 출력되면, 제 1단의 제 2부호화부(120_1)의 확률적 코드북(310)과 이득코드북(315)의 지수데이터가 출력된다.

이어서, 필터상태는 단계(S6)에 나타낸 바와같이 업데이트된다.

본 실시예에 있어서, 제 2단의 제 2부호화부(120_2)의 지수비트수는 형상벡터에 대해서는 5만큼 작은 한편, 이득에 대해서는 3만큼 작다. 이 경우, 코드북에 적절한 형상과 이득이 존재하지 않으면, 양자화에러가 감소되는 대신 증가될 우려가 있다.

그러한 결점을 방지하기 위해 비록 0을 이득으로 부여할 수도 있지만, 이득용으로는 3개의 비트만이 있다. 그중 하나가 0으로 설정되면, 양자화성능을 굉장히 악화시킨다. 이러한 것을 고려하여, 보다 큰 비트수가 할당되어 있는 형상벡터용으로 모두 0인 벡터를 부여하고 있다. 상술한 탐색은 모든 제로벡터를 베제하고 수행되며, 모든 제로벡터는 양자화에러가 최종으로 증가되어 있는 경우에 선택된다.

이득은 제멋대로이다. 이것은 제 2단의 제 2부호화부(120_2)에서 양자화에러가 증가되는 것을 방지하는 것을 가능하게 해준다.

비록 2단의 구조가 상술되었지만, 그러한 단수는 2보다 클 수도 있다. 그러한 경우, 제 1단의 폐루프탐색에 의한 벡터양자화가 가까워지게 되면, N번째 단(여기서 2≤N)의 양자화는 기준입력으로써 (N-1)번째의 단의 양자화에러로 수행되게 되며, N번째 단의 양자화에러는 (N+ 1)번째 단에 대한 기준입력으로써 사용된다.

도 9와 도 10으로부터는 제 2부호화부에 대해 다단의 벡터양자화를 적용함으로써, 그 계산량을 동일수의 비트수를 갖는 직벡터양자화의 사용 또는 켤레코드북의 사용에 있어서의 계산량과 비교하여 보다 감소시킬수 있다. 특히, 분석/합성방법에 의한 폐루프탐색을 이용하여 시간축파형의 벡터양자화를 행하는 CELF부호화에 있어서, 보다 적은 탐색동작회수는 굉장히 중요하다. 또한, 그러한 비트수는 2단의 제 2부호화부(120_1, 120_2)의 지수출력을 모두 이용하는 것과 제 2단의 제 2부호화부(120_2)의 출력은 이용하지 않고 제 1의 제 2부호화부(120_1)의 출력만을 이용하는 것을 절환함으로써 손쉽게 절환할 수 있다. 제 1단 및 제 2단의 제 2부호화부(120_1, 120_2)의 지수출력이 합쳐져서 출력되는 경우에는, 디코더는 그러한 지수출력중 하나를 선택함으로써 그러한 구조를 손쉽게 수습할 수 있다. 즉, 2kbps에서 동작하는 디코더를 이용하여 예를 들어 6kbps로 부호화된 파라미터를 복호화함으로써 디코더의 구조를 손쉽게 수습할 수 있다. 또한, 제로벡터가 제 2단의 제 2부호화부(120_2)의 형상코드북에 포함되는 경우에는, 양자화에러가 0이 이득에 가해지는 경우보다 그 성능이 보다 덜한 악화로 증가되는 것을 방지할 수 있게 된다.

예를들어, 확률적 코드북의 코드벡터는 소위 가우스잡음을 클립핑(clipping)함으로써 생성될 수 있다. 특히, 코드북은 가우스잡음을 생성하고, 적당한 임계값으로 가우스잡음을 클립핑하고, 클립핑된 가우스잡음을 규격화함으로써, 생성될 수 있다.

그렇지만, 음성에는 다양한 형태가 있다. 예를들어, 가우스잡음은 "사, 시, 수, 세, 소"등의 잡음에 가까운 자음사운드의 음성을 수습할 수는 있는 한편, 그러한 가우스잡음은 "파, 피, 푸, 페, 포"등의 급격히 상승하는 자음에 대한 음성을 수습할 수는 없다. 본 발명에 따르면, 가우스잡음이 몇몇의 코드벡터에 가해지게 되는 한편, 코드벡터의 나머지부분은 학습에 의해 처리된다. 그 결과, 급격히 상승하는 자음사운드와 잡음에 가까운 자음사운드를 갖는 모든 자음을 수습할 수 있다. 예를들어, 임계값이 존재하게 되면, 6~7개의 큰 피크를 갖는 벡터가 얻어지는 반면, 임계값이 감소되면, 코드벡터는 가우스잡음에 근접하게 된다. 따라서, 임계값을 클립핑하는 데 있어서의 변동율을 증가시킴으로써, "파, 피, 푸, 페, 포"등의 급격한 상승부를 갖는 잡음 또는 "사, 시, 수, 세, 소"등의 잡음에 가까운 자음을 수습할 수 있게 되고, 따라서 명확성을 증가시키게 된다. 도 11은 가우스잡음 및 클립핑된 잡음의 외형을 각각 굵은 선 및 점선으로 나타내고 있다. 도 11a 도 11b는 1.0과 동등한 클립핑 임계값 즉, 보다 큰 임계값을 갖는 잡음과, 0.4와 동등한 클립핑 임계값 즉, 보다 작은 임계값을 갖는 잡음을 나타내고 있다. 도 11a 및 도 11b로부터는 임계값이 크게 선택될 경우에는 6~7개의 큰 피크를 갖는 벡터가 얻어지는 한편, 임계값이 작은 값으로 선택되는 경우에는 잡음은 스스로 가우스잡음에 접근한다.

이를 실현하기 위해서, 초기의 코드북은 가우스잡음을 클립핑함으로써 준비되며, 적당한 수의 비학습코드벡터가 설정된다. 비학습코드벡터는 "사, 시, 수, 세, 소"등의 잡음에 가까운 자음을 수습하기 위해 증가하는 변동값의 차수로 선택된다. 학습에 의해 구해진 벡터는 학습용 LBG알고리듬을 사용한다. 가장 근접한 이웃조건에서 부호화는 고정된 코드벡터와 학습으로 얻어진 코드벡터를 모두 사용한다.

중심조건에서는, 학습용으로 설정된 코드벡터만이 업데이트된다. 따라서, 학습용으로 설정된 코드벡터는 "파, 피, 푸, 페, 포"등의 급격히 상승하는 자음들을 수습할 수 있다.

최적의 이득은 통상의 학습에 의해 이러한 코드벡터에 대해 학습될 수 있다.

도 12는 가우스잡음을 클립핑하여 코드북을 구성하기 위한 처리흐름을 나타낸다.

도 12에 있어서, 학습의 회수(n)는 초기화를 위한 단계(S10)에서 n = 0으로 설정한다. 에러(D_0) = ∞임에 따라, 학습의 최대회수(n_{max})가 설정되며, 학습종료조건을 설정하는 임계값(∈)이 설정된다.

다음단계(S11)에서는, 가우스잡음을 클립핑하여 초기의 코드북을 생성한다. 단계(S12)에서는 코드벡터의 일부를 비학습코드벡터로 고정한다.

다음 단계(S13)에서는, 상기 코드북을 읊조려서 부호화를 행한다. 단계(S14)에서는, 에러를 계산한다. 단계(S15)에서는, $D_{n-1} - D_n / D_n < \in$, 혹은 $n = n_{max}$인지를 판별한다. 그 결과가 "예"이면, 처리를 종료한다. 그 결과가 "아니오"이면, 처리를 단계(S16)로 전환한다.

단계(S16)에서는, 부호화에 사용되지 않은 코드벡터를 처리한다. 다음 단계(S17)에서는, 코드북을 업데이트한다. 단계(S18)에서는, 학습의 회수를 단계(S13)로 되돌아가기 전에 증분시킨다.

상술한 신호부호화 및 신호복호화장치를 예를들어 도 14에 나타낸 휴대용 통신단말기나 휴대용 전화기에 채용되는 음성코드북으로써 사용할 수도 있다.

도 13은 도 1 및 도 3에 나타낸 바와 같이 구성된 음성부호화부(160)을 적용하고 있는 휴대용 단말기의 송신측을 나타낸다. 마이크(161)에서 수집된 음성신호는 증폭기(162)에 의해 증폭되어 아날로그/디지털(A/D)변환기(163)에 의해 디지털신호롤 변환되어서, 도 1 및 도 3에 나타낸 바와 같이 구성된 음성부호화부(160)으로 보내진다. A/D변환기(163)로부터의

디지털신호는 입력단자(101)로 인가된다. 음성부호화부(160)는 도 1 및 도 3과 결부하여 설명한 바와같이 부호화를 수행한다. 도 1 및 도 2의 출력단자의 출력신호는 음성부호화부(160)의 출력신호로써 송신채널부호화부(164)로 보내져서, 그 후, 공급된 신호에 따라서 채널부호화를 수행하게 된다.

송신채널부호화부(164)의 출력신호는 변조를 위해 변조회로(165)로 보내지고, 그런다음 디지털/아날로그(D/A)변환기(166)와 RF증폭기(167)를 통해서 안테나(168)로 인가된다.

도 14는 도 4에 나타낸 바와같이 구성된 음성복호화부(260)를 사용하는 휴대용 단말기의 수신측을 나타낸다. 도 14의 안테나(261)에 의해 수신된 음성신호는 RF증폭기(262)에서 증폭되어서, 아날로그/디지털(A/D)변환기(263)를 통해 변조회로(264)로 인가되며, 그로부터 변조된 신호는 송신채널복호화부(265)로 인가된다. 복호화부(265)의 출력신호는 도 2 및 도 4에 나타낸 바와같이 구성된 음성복호화부(260)로 인가된다. 음성복호화부(260)는 도 2 및 도 4와 결부시켜 설명한 바와같이 신호를 복호화한다. 도 2 및 도 4의 출력단자(201)에서의 출력신호는 음성복호화부(260)의 신호로써 디지털/아날로그(D/A)변환기(266)로 보내진다. D/A변환기(266)로부터의 아날로그음성신호는 스피커(268)로 보내진다.

발명의 효과

본 발명을 따르는 음성부호화방법 및 장치에 의하면, 적어도 입력음성신호의 유성음부에 대하여 단기간예측 오차(residual)가 구해지고, 단기간예측 오차(residual)를 기초로 하여 정현분석 부호화파라미터가 구해진다. 이러한 정현분석 부호화파라미터는 청각가중벡터양자화에 의해 양자화된다. 입력음성신호의 유성음부는 위상재생도를 갖는 파형부호화로 부호화된다. 청각가중벡터양자화에 있어서는, 제 1벡터양자화가 우선 행해지며, 제 1벡터양자화시에 생성된 양자화에러벡터는 제 2벡터양자화에 의해 양자화된다. 이러한 식으로, 출력 부호화데이 터의 비트수를 데이터송신채널의 용량에 따라 용이하게 절환시킬 수 있어, 복수의 데이터비트속도를 수습할 수 있다. 또한, 그러한 부호화데이터열을 생성하여 엔코더와 디코더간의 비트속도가 다른 경우라도 디코더측에 대하여 용이하게 수습할 수 있다.

(57) 청구의 범위

청구항 1.

시간축상의 입력음성신호를 단위로써 블록으로 분할하여 그 결과의 신호를 부호화하는 음성부호화방법에 있어서,

적어도 입력음성신호의 유성음부에 대해서 단기간예측 오차(residual)를 구하는 단계와,

그렇게 구해진 단기간예측 오차(residual)를 기초로 하여 정현분석 부호화파라미터를 구하는 단계와,

정현분석 부호화파라미터상에 청각가중벡터양자화를 수행하는 단계와,

파형부호화를 이용하여 입력음성신호의 무성음부를 부호화하는 단계를 포함하는 것을 특징으로 하는 음성부호화방법.

청구항 2.

제 1항에 있어서,

입력음성신호가 유성음인지 무성음인지를 판정하고, 그 판정결과를 기초로, 유성음으로 구해지는 입력음성신호부분은 상기 정현분석부호화로 처리하고 무성음으로 구해지는 입력음성신호부분은 분석/합성방법을 이용하는 페루프 최적벡터탐색으로 벡터양자화하는 것을 특징으로 하는 음성부호화방법.

청구항 3.

제 1항에 있어서,

상기 청각가중벡터양자화가 수행된 정현분석파라미터로써 스펙트럼포락선을 나타내는 데이터를 사용하는 것을 특징으로 하는 음성부호화방법.

청구항 4.

시간축상의 입력음성신호를 단위로써 블록으로 분할하여 그 결과의 신호를 부호화하는 음성부호화장치에 있어서,

적어도 입력음성신호의 단기간예측 오차(residual)를 구하는 수단과,

그렇게 구해진 단기간예측 오차(residual)를 기초로 하여 정현분석 부호화파라미터를 구하는 수단과,

정현분석 부호화파라미터상에 청각가중벡터양자화를 수행하는 수단과,

파형부호화를 이용하여 입력음성신호의 무성음부를 부호화하는 수단을 포함하여 구성된 것을 특징으로 하는 음성부호화장치.

청구항 5.

시간축상의 입력음성신호를 단위로써 블록으로 분할하여 그 결과의 신호를 부호화하는 음성부호화방법에 있어서,

적어도 입력음성신호의 유성음부에 대해서 단기간예측 오차(residual)를 구하는 단계와,

그렇게 구해진 단기간예측 오차(residual)를 기초로 하여 정현분석 부호화파라미터를 구하는 단계와,

정현분석 부호화파라미터상에 청각가중벡터양자화를 수행하는 단계를 포함하는 것을 특징으로 하는 음성부호화방법.

청구항 6.

제 5항에 있어서,

제 1벡터양자화단계와,

상기 제 1벡터양자화시에 생성된 양자화에러벡터를 양자화하는 제 2벡터양자화단계를 적어도 포함하는 것을 특징으로 하는 음성부호화방법.

청구항 7.

제 6항에 있어서,

제 1벡터양자화단계의 출력을 저속의 비트속도로 도출하고, 상기 제 1벡터양자화단계의 출력과 상기 제 2벡터양자화단계의 출력들을 고속의 비트속도로 도출하는 것을 특징으로 하는 음성부호화방법.

청구항 8.

시간축상의 입력음성신호를 단위로써 블록으로 분할하여 그 결과의 신호를 부호화하는 음성부호화장치에 있어서,

입력음성신호의 단기간예측 오차(residual)를 구하는 수단과,

그렇게 구해진 단기간예측 오차(residual)로부터 정현분석 부호화파라미터를 구하는 수단과,

정현분석 부호화파라미터상에 청각가중된 다단의 벡터양자화를 수행하는 수단을 포함하여 구성된 것을 특징으로 하는 음성부호화장치.

청구항 9.

입력음성신호를 증폭하기 위한 증폭수단과,

상기 변조된 신호를 A/D변환하기 위한 A/D변환수단과,

상기 A/D변환수단의 음성출력을 부호화하기 위한 음성부호화수단과,

결과의 부호화신호를 채널부호화하기 위한 송신경로부호화수단과,

상기 송신경로부호화수단의 출력을 변조하기 위한 변조수단과,

결과의 변조신호를 D/A변환하기 위한 D/A변환수단과,

상기 D/A변환수단으로부터의 신호를 증폭하여 결과의 증폭신호를 안테나로 인가하기 위한 증폭수단을 포함하는 것으로,

상기 음성부호화수단은

상기 입력음성신호의 단기간예측 오차(residual)를 구하는 수단과,

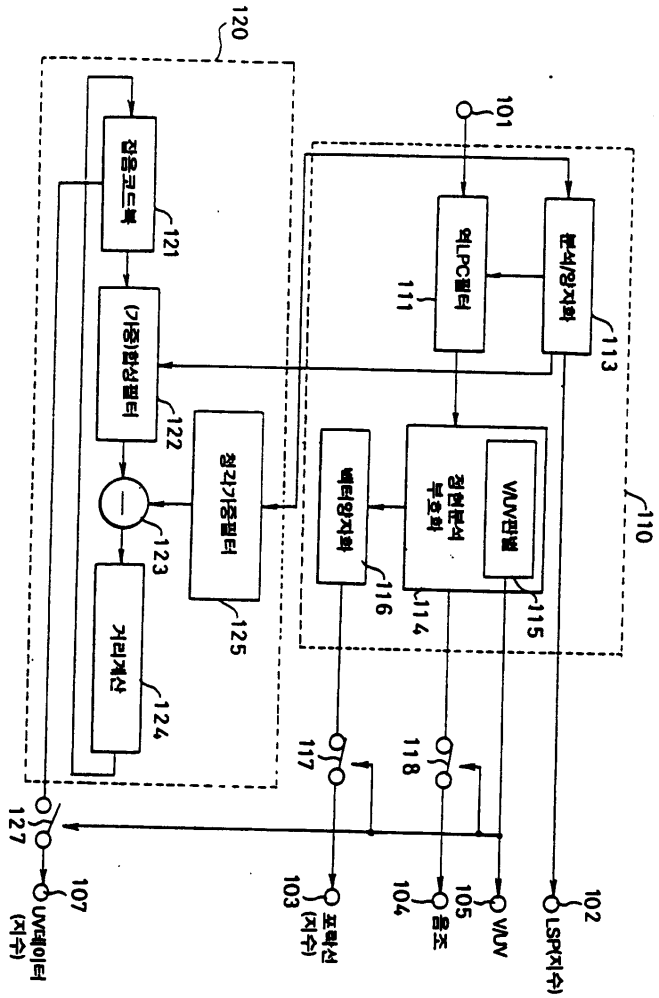
그렇게 구해진 단기간예측 오차(residual)로부터 정현분석 부호화파라미터를 구하는 수단과,

상기 정현분석 부호화파라미터상에 청각가중벡터양자화를 수행하는 수단과,

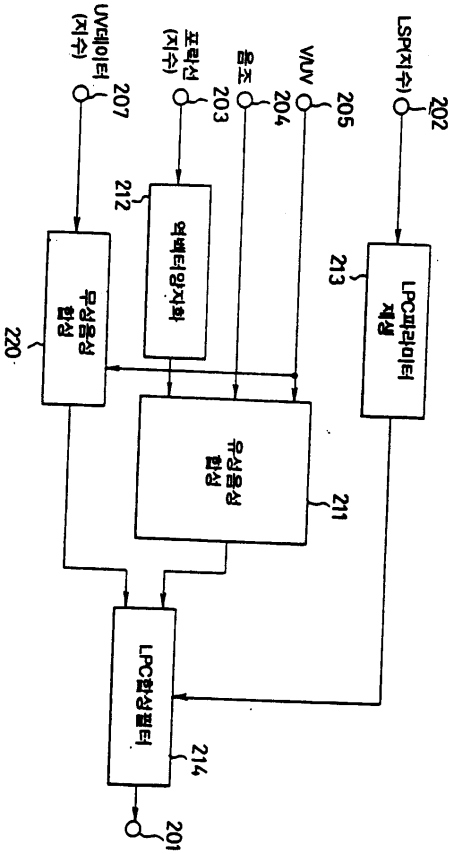
파형부호화로 상기 입력음성신호를 부호화하는 수단을 더 포함하여 구성된 것을 특징으로 하는 휴대용 무선단말장치.

도면

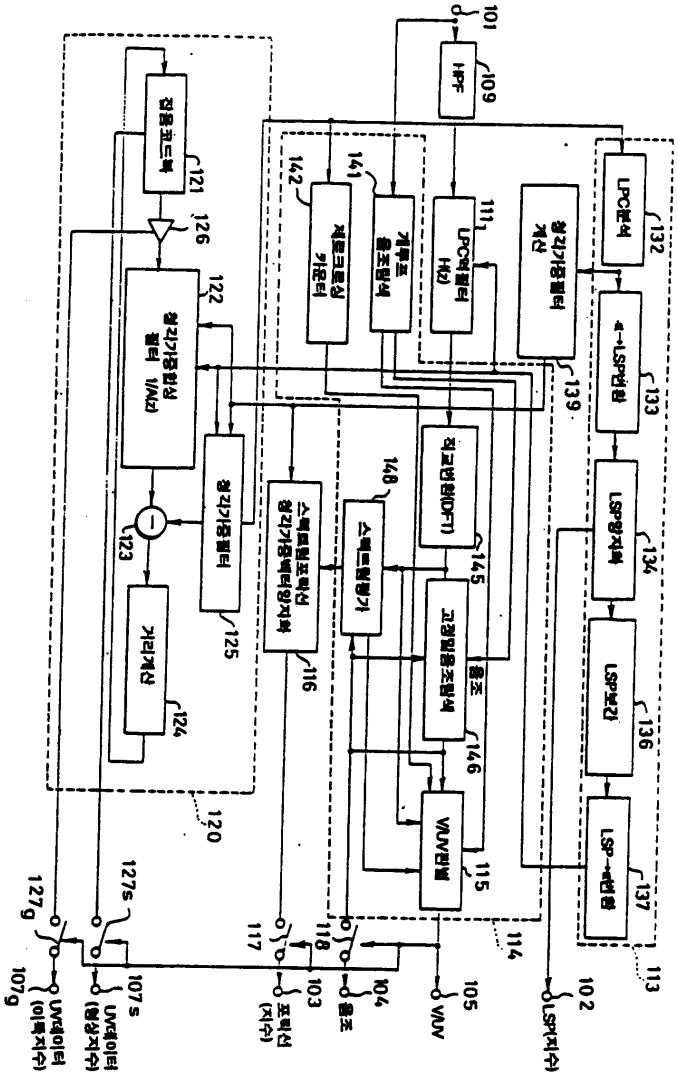
도면1



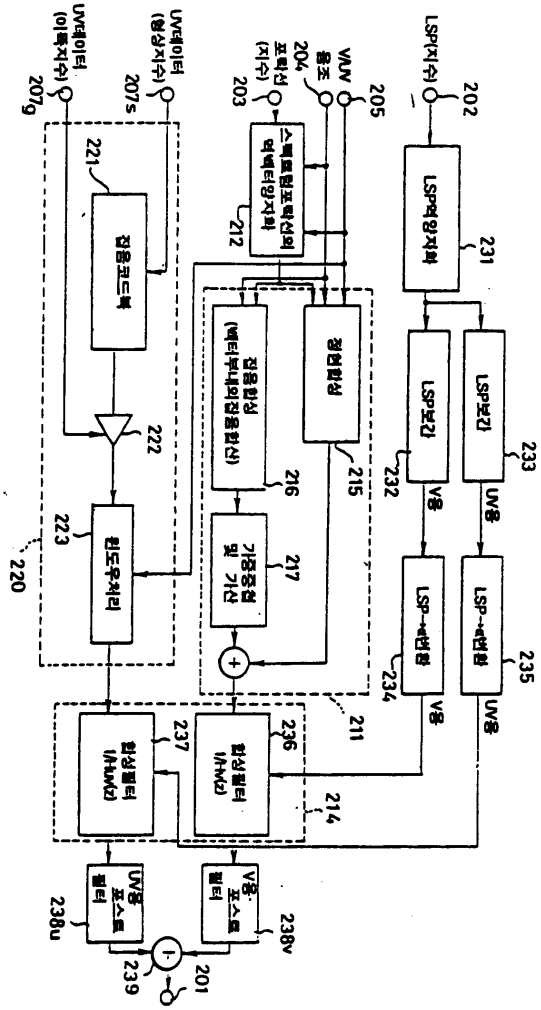
도면2


LSP(지수)
202
VUV
205
음조
204
포락선
(지수)
203
UV데이터
(지수)
207
LPC파라미터
재생
213
역벡터양자화
212
무성음성
합성
220
유성음성
합성
211
LPC합성필터
214
201

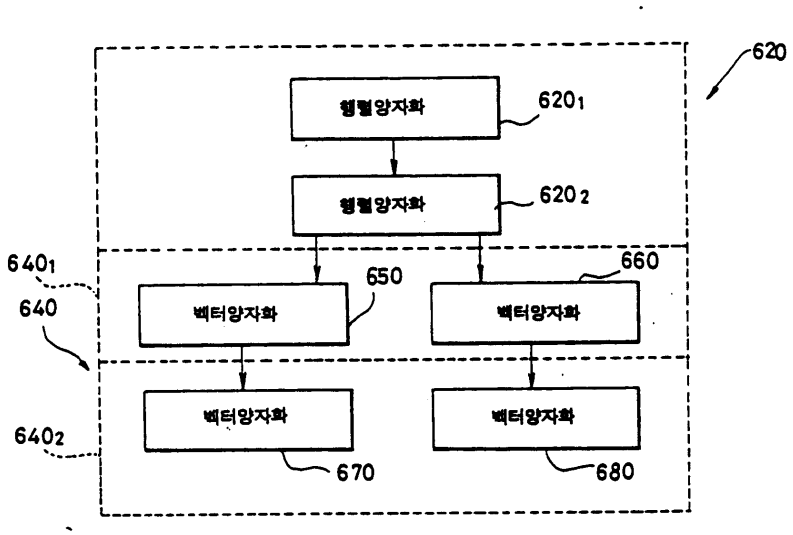
도면3


101
HPF
109
LPC분석
132
LSP변환
133
LSP양자화
134
LSP보간
136
LSP→α변환
137
113
LPC역필터 H(z)
111
LSP→α변환
139
제로크로싱 카운터
142
개루프 음조탐색
141
직교변환(DFT)
145
스펙트럼평가
148
스펙트럼포락선 청각가중벡터양자화
116
고정밀음조탐색
146
VUV판별
115
114
102
LSP(지수)
105
VUV
104
음조
103
포락선 (지수)
117
118
127s
127g
107s
UV데이터 (청상지수)
107g
UV데이터 (이득지수)
120
경음코드북
121
126
122
청각가중합성 필터 1/A(z)
123
청각가중필터
125
거리계산
124

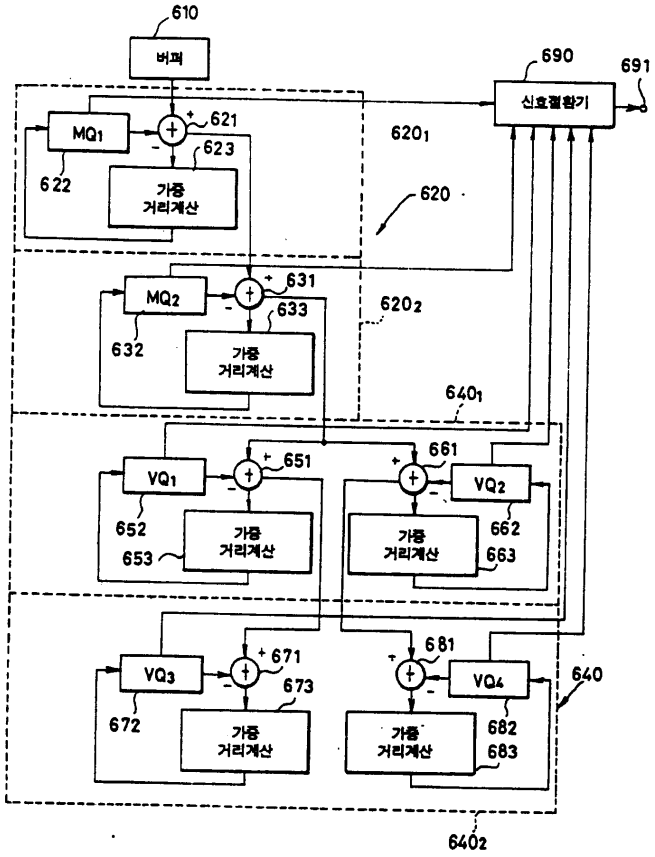
도면4


LSP(지수)
202
LSP역양자화
231
LSP보간
232
LSP보간
233
V용
UV용
LSP→α변환
234
LSP→α변환
235
V용
UV용
VUV
음조
204
205
포락선
(지수)
203
스펙트럼포락선의
역벡터양자화
212
정현합성
215
잡음합성
(벡터부내의잡음합산)
216
가중중첩
및 가산
217
211
UV데이터
(형상지수)
207s
UV데이터
(이득지수)
207g
잡음코드북
221
222
윈도우처리
223
220
합성필터
(HUV(z))
237
합성필터
(HV(z))
236
214
UV용
포스트
필터
238u
V용
포스트
필터
238v
239
201

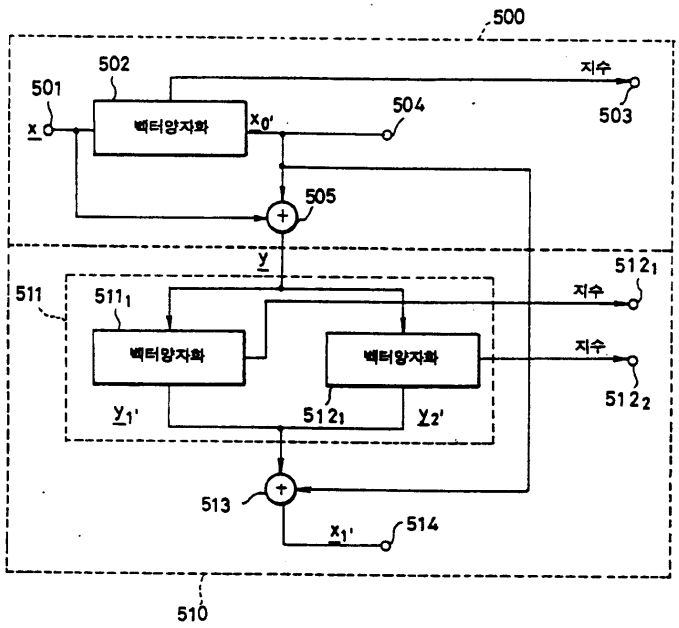
도면5


620
행렬양자화
620₁
행렬양자화
620₂
640₁
640
650
660
벡터양자화
벡터양자화
벡터양자화
벡터양자화
640₂
670
680

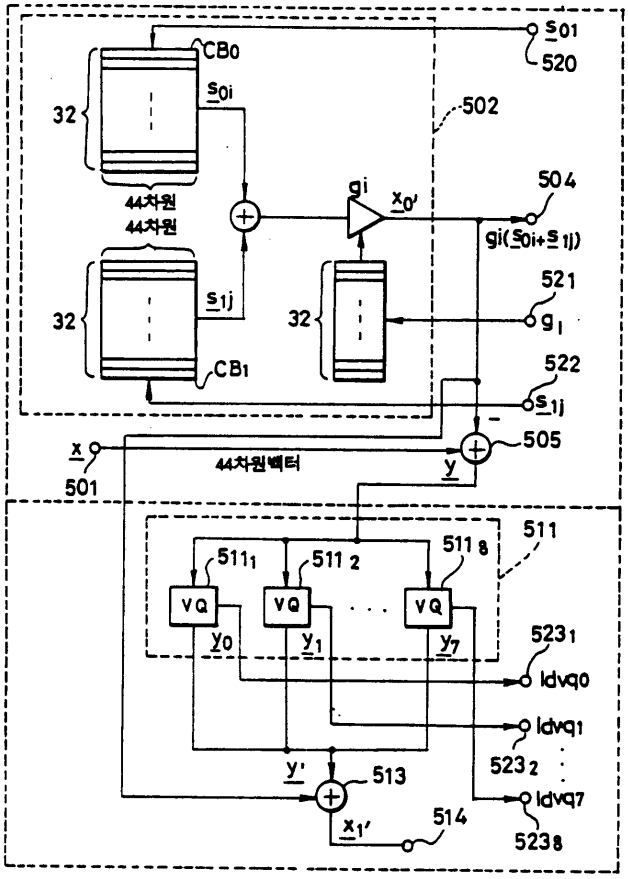
도면6


610
버퍼
690
691
신호결합기
MQ1
621
623
622
가중
거리계산
620₁
620
MQ2
631
633
632
가중
거리계산
620₂
640₁
VQ1
651
652
653
가중
거리계산
661
VQ2
662
663
가중
거리계산
VQ3
671
673
672
가중
거리계산
681
VQ4
682
683
가중
거리계산
640
640₂

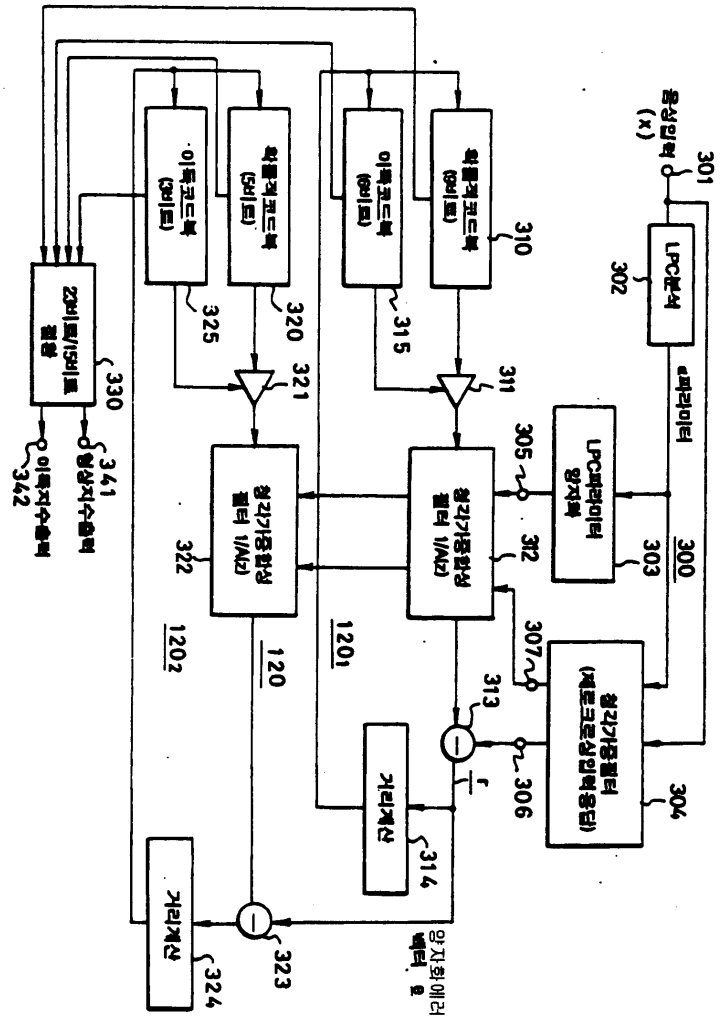
도면7


500
502
501
x
벡터양자화
x0'
504
지수
503
505
y
512₁
511
511₁
지수
벡터양자화
벡터양자화
지수
512₂
y1'
512₁
y2'
513
x1'
514
510

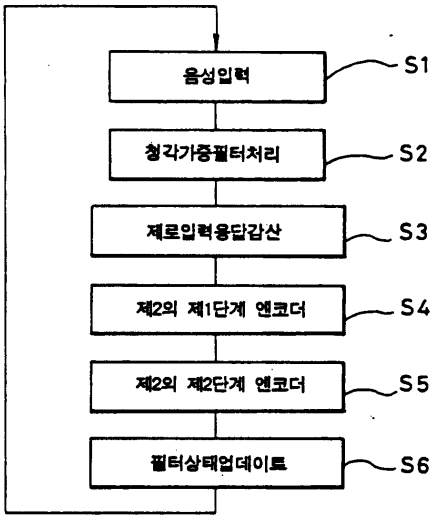
도면8


CB0
s0i
32
44차원
44차원
s01
520
502
gi
x0'
504
gi(s0i+s1j)
s1j
32
32
CB1
521
g1
522
s1j
x
501
44차원벡터
y
505
511
511 1
511 2
511 8
VQ
VQ
VQ
y0
y1
y7
523 1
ldvq0
ldvq1
523 2
ldvq7
523 8
y'
513
x1'
514

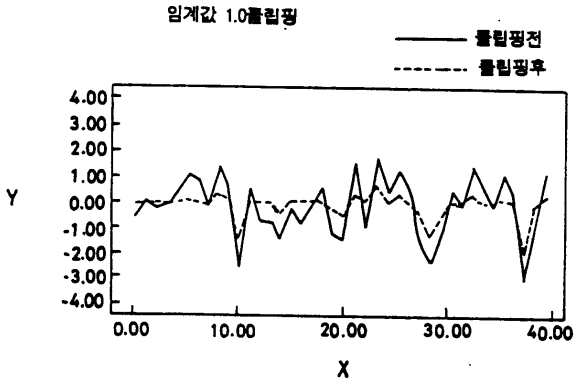
도면9


음성입력
(x)
301
LPC분석
302
α파라미터
LPC파라미터
양자화
303
300
305
312
청각가중합성
필터 1/A(z)
307
313
306
청각가중필터
(제로크로상입력응답)
304
r
거리계산
314
양자화에러
벡터 e
310
311
315
320
321
325
이득코드북
(3비트)
이득코드북
(5비트)
이득코드북
(6비트)
확률계코드북
(9비트)
23비트/15비트
결합
330
341
청상지수출력
342
이득지수출력
청각가중합성
필터 1/A(z)
322
120
120₁
120₂
거리계산
323
324

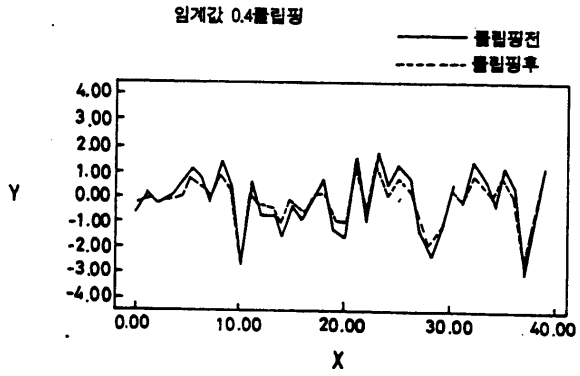
도면10


음성입력
S1
청각가중필터처리
S2
제로입력응답감산
S3
제2의 제1단계 엔코더
S4
제2의 제2단계 엔코더
S5
필터상태업데이트
S6

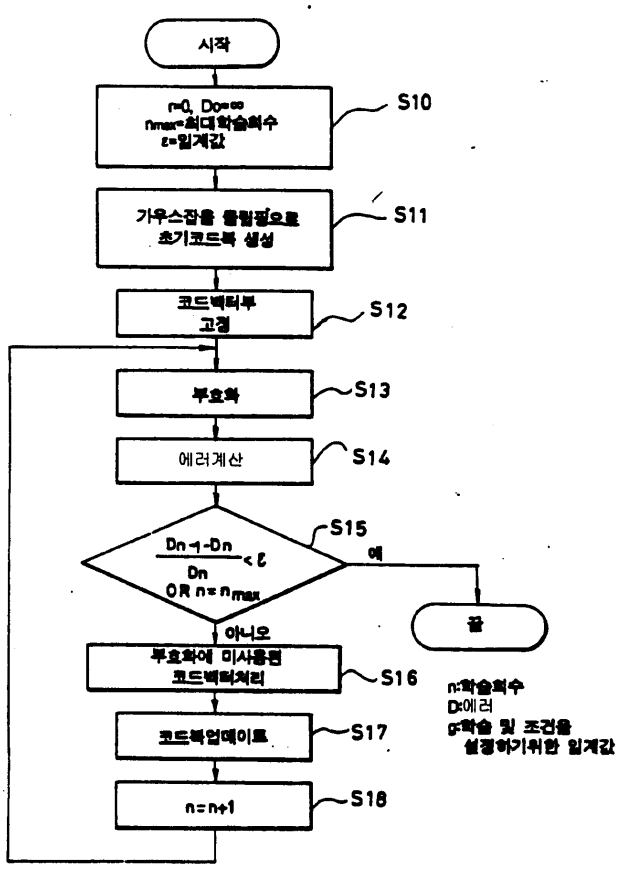
도면11a


임계값 1.0클립핑
클립핑전
클립핑후
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
Y
0.00
10.00
20.00
30.00
40.00
X

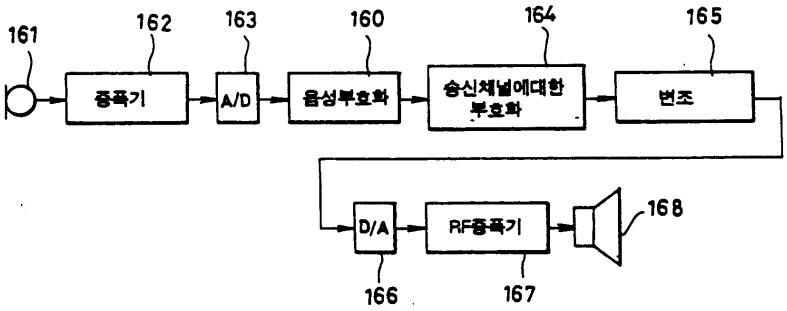
도면11b


임계값 0.4클립핑
클립핑전
클립핑후
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
Y
0.00
10.00
20.00
30.00
40.00
X

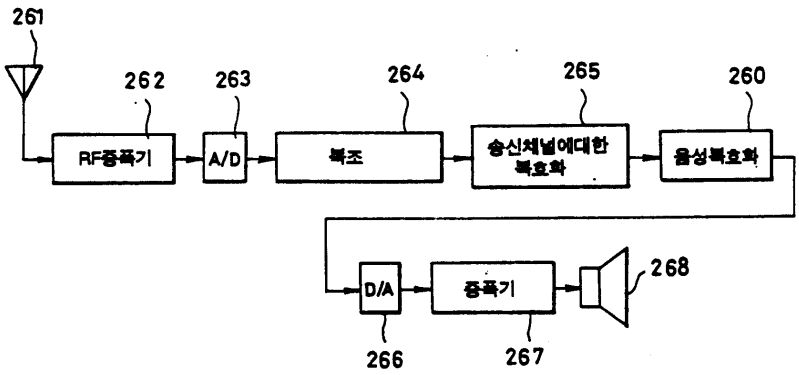
도면12


시작
n=0, Do=∞
nmax=최대학습회수
ε=임계값
S10
가우스잡음 클럽핑으로
초기코드북 생성
S11
코드벡터부
고정
S12
부호화
S13
에러계산
S14
Dn-1-Dn / Dn < ε OR n=nmax
S15
예
끝
아니오
부호화에 미사용된
코드벡터처리
S16
코드북업데이트
S17
n=n+1
S18
n:학습회수
D:에러
ε:학습 및 조건을
설정하기위한 임계값

도면13


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

도면14



도면15

	2kbps	6kbps
U/V판별출력	1비트/20msec	1비트/20msec
LSP양자화지수	32비트/40msec	48비트/40msec
유성음성용(V)	음조데이터 8비트/20msec	음조데이터 8비트/20msec
	스펙트럼 포락선 벡터 양자화 지수15비트/20msec	스펙트럼 포락선 벡터 양자화 지수87비트/20msec
	형상(제1단계용),5+5 비트/20msec 이득, 5비트/20msec	형상(제1단계용),5+5 비트/20msec 이득, 5비트/20msec 지수(제2단계용),72 비트/20msec
무성음성용(UV)	시간영역 벡터양자화 지수11비트/10msec	시간영역 벡터양자화 지수23비트/5msec
	형상(제1단계용),7비트 /10msec 이득, 4비트/10msec	제1단계용 형상,9비트 /5msec 이득,6비트/5msec 제2단계용 형상,5비트 /5msec 이득,3비트/5msec
유성음성용 무성음성용	40비트/20msec 39비트/20msec	120비트/20msec 117비트/20msec