

(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

| | | |
|---|-------------------------------------|--|
| (51) 。 Int. Cl. G10L 15/00 (2006.01) | (45) 공고일자 (11) 등록번호 (24) 등록일자 | 2006년07월21일 10-0543982 2006년01월10일 |
|---|-------------------------------------|--|

| | | | |
|-----------|-----------------|-----------|-----------------|
| (21) 출원번호 | 10-1997-0048321 | (65) 공개번호 | 10-1998-0024885 |
| (22) 출원일자 | 1997년09월23일 | (43) 공개일자 | 1998년07월06일 |

(30) 우선권주장 96-251614 1996년09월24일 일본(JP)

(73) 특허권자 소니 가부시끼 가이샤
일본국 도쿄도 시나가와구 기타시나가와 6쵸메 7반 35고

(72) 발명자 니시구찌 마사유키
일본국 도쿄도 시나가와구 기타시나가와 6쵸메 7방 35고 소니가부시끼
가이샤내

마쯔모토 준
일본국 도쿄도 시나가와구 기타시나가와 6쵸메 7방 35고 소니가부시끼
가이샤내

이이지마 가즈유키
일본국 도쿄도 시나가와구 기타시나가와 6쵸메 7방 35고 소니가부시끼
가이샤내

(74) 대리인 신관호

심사관 : 안준형

(54) 벡터양자화방법,음성부호화방법및장치

요약

벡터양자화를 위한 코드북탐색에 대한 처리량은 감소될 것이다. 사인파분석 부호기(114)의 스펙트럼 평가부(148)에서 벡터양자화용 벡터양자화기(116)에 조파들의 스펙트럼성분들의 엔벨로프를 나타내는 데이터를 보낼 때, 입력벡터와 코드북에 저장된 모든 코드벡터들 사이의 유사도가 좀더 작은 수의 코드벡터들을 예비선택하기 위해 근사화에 의해 구해진다. 이 복수개의 예비 선택된 코드벡터들로부터, 입력벡터에 대해 오차를 최소로 하는 코드벡터가 최종 선택된다. 이 방법에서, 좀더 작은 수의 후보 코드벡터들이 단순화된 처리를 포함하는 예비선택에 의해 예비 선택되고 결과적으로 고정밀도로 최종 선택된다.

대표도

도 1

명세서

도면의 간단한 설명

- 도 1은 본 발명에 따라서 부호화 방법을 수행하기 위한 음성신호 부호화장치(부호기)의 기본적인 구조를 나타내는 블록도이다.
- 도 2는 본 발명에 따라서 복호화 방법을 수행하기 위한 음성신호 복호화장치(복호기)의 기본적인 구조를 나타내는 블록도이다.
- 도 3은 도 1에 나타난 음성신호 부호기의 좀더 구체적인 구조를 나타내는 블록도이다.
- 도 4는 도 2에 나타난 음성신호 복호기의 좀더 상세한 구조를 나타내는 블록도이다.
- 도 5는 음성신호 부호화장치의 비트율을 나타내는 표이다.
- 도 6은 LSP양자화기의 좀더 상세한 구조를 나타내는 블록도이다.
- 도 7은 LSP양자화기의 기본적인 구조를 나타내는 블록도이다.
- 도 8은 벡터양자화기의 좀더 상세한 구조를 나타내는 블록도이다.
- 도 9는 벡터양자화기의 좀더 상세한 구조를 나타내는 블록도이다.
- 도 10은 가중(weighting)하기 위해 $W[i]$ 의 가중값의 구체적인 예를 도시하는 그래프이다.
- 도 11은 양자화값들과, 차원의 수와 비트의 수들 사이의 관계를 나타내는 표이다.
- 도 12는 가변차원 코드북 복구를 위한 벡터양자화기의 도식적 구조를 나타내는 블록회로도이다.
- 도 13은 가변차원 코드북 복구를 위한 벡터양자화기의 다른 도식적 구조를 나타내는 블록회로도이다.
- 도 14는 가변차원용 코드북과 고정차원용 코드북을 이용하는 벡터양자화기의 제 1도식적 구조를 나타내는 블록회로도이다.
- 도 15는 가변차원용 코드북과 고정차원용 코드북을 이용하는 벡터양자화기의 제 2도식적 구조를 나타내는 블록회로도이다.
- 도 16은 가변차원용 코드북과 고정차원용 코드북을 이용하는 벡터양자화기의 제 3도식적 구조를 나타내는 블록회로도이다.
- 도 17은 가변차원용 코드북과 고정차원용 코드북을 이용하는 벡터양자화기의 제 4도식적 구조를 나타내는 블록회로도이다.
- 도 18은 본 발명에 따라서 음성부호화장치의 CULP부호화부(제 2부호기)의 구체적인 구조를 나타내는 블록회로도이다.
- 도 19는 도 16에 나타난 구성에서 처리흐름을 나타내는 흐름도이다.
- 도 20은 다른 임계값들에서 클리핑 후의 잡음과 가우스잡음의 상태를 나타낸다.
- 도 21은 학습에 의해 형상코드북을 발생하는 시간(o)에서 처리흐름을 나타내는 흐름도이다.
- 도 22는 U/UV천이에 따라서 LSP스위칭의 상태를 나타내는 표이다.
- 도 23은 10차 LPC분석에 의해 얻어진 α 파라미터에 근거해서 10차 선형스펙트럼쌍(LSP)을 나타낸다.

- 도 24는 무성음(UV) 프레임에서 유성음(V) 프레임으로 변하는 이득의 상태를 도시한다.
- 도 25는 1프레임씩 합성된 스펙트럼성분이나 파형에 대한 보간동작을 도시한다.
- 도 26은 유성음(V) 프레임과 무성음(UV) 프레임 사이의 결합부에서 중첩을 도시한다.
- 도 27은 유성음의 합성시의 잡음부가처리를 도시한다.
- 도 28은 유성음의 합성 시에 부가된 잡음의 진폭계산의 예를 도시한다.
- 도 29는 포스트필터의 도시적 구조를 도시한다.
- 도 30은 필터계수를 갱신하는 주기와 포스트필터의 이득갱신주기를 도시한다.
- 도 31은 포스트필터의 이득과 필터계수의 프레임 경계부에서 합병하기 위한 처리를 도시한다.
- 도 32는 본 발명을 구체화하는 음성신호 부호화장치를 이용하는 휴대용 단말기의 송신 측의 구조를 나타내는 블록도이다.
- 도 33은 본 발명을 구체화하는 음성신호 복호화장치를 이용하는 휴대용 단말기의 수신 측의 구조를 나타내는 블록도이다.

* 도면의 주요부분에 대한 부호설명

110. 제 1부호화부 111. LPC역필터

113. LPC분석양자화부 114. 사인파분석 부호화부

115. V/UV판정부 116. 벡터양자화기

120. 제 2부호화부 121. 잡음코드북

122. 가중합성필터 123. 감산기

124. 거리계산회로 125. 청각가중필터

530. 코드북(코드벡터) 531. 형상코드북

532. 이득코드북 533. 이득회로

535. 선택회로 542. 가변/고정차원 변환회로

544. 고정/가변차원 변환회로 545. 선택회로

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 입력벡터가 코드벡터 중 최적 코드벡터의 인덱스를 출력하기 위한 코드북에 저장된 코드벡터들에 비교되는 벡터양자화 방법에 관한 것이다. 또한 본 발명은 입력음성신호가 블록이나 프레임과 같은 소정의 부호화 단위로 분할되고 벡터양자화를 포함하는 부호화 처리가 부호화 단위에 근거해서 수행되는 음성부호화 방법 및 장치에 관한 것이다.

지금까지 오디오나 비디오신호를 디지털화 하고 압축부호화 하기 위해서 여러 개의 입력데이터가 하나의 코드(인덱스)로서 나타내기 위한 벡터로 함께 묶여지는 벡터양자화가 알려져 있다.

상기 벡터양자화에서, 다양한 입력벡터들의 대표적인 패턴은 예를 들어 학습과 주어진 코드들이나 인덱스들에 의해 미리 결정되며, 다음에 코드북에 저장된다. 다음에 입력벡터가 가장 큰 유사도(similarity)나 상관(correlation)을 보이는 패턴의 코드를 출력하기 위한 패턴매칭(pattern matching)의 방법으로 대표적인 패턴(코드벡터)들에 비교된다. 이 유사도나 상관은 입력벡터와 대표적인 코드벡터들 사이의 왜곡량이나 오차에너지를 계산함으로써 구해지고 유사도나 상관이 높을 수록 왜곡이나 오차가 적어진다.

지금까지 시간영역이나 주파수영역에서 통계적 특성을 이용하는 다양한 부호화 방법과 신호압축에서 인간의 정신음향적(psychoacoustic) 특성을 이용한 다양한 부호화 방법이 알려져 있다. 이 부호화 방법은 대략 시간영역에서의 부호화, 주파수영역에서의 부호화, 그리고 합성에 의한 분석 부호화로 분류된다.

음성신호의 고효율 부호화의 예 중에는, 조파(harmonic) 부호화, 부대역 부호화(SBC), 선형예측 부호화(LPC), 이산 사인파 변환(DCT), 수정된 DCT(MDCT) 또는 고속 푸리에변환(FFT)과 같은 사인파분석 부호화가 있다.

음성신호의 고효율 부호화에서, 상술한 벡터양자화는 조파들의 스펙트럼성분과 같은 파라미터로 이용된다.

발명이 이루고자 하는 기술적 과제

반면에, 만약 코드북에 저장된 패턴들의 수, 즉 코드벡터들의 수가 크거나 또는 만약 벡터양자화기가 여러 개의 코드북으로 구성된 다단 구성이라면, 패턴매칭을 위한 코드벡터 탐색동작의 횟수가 증가되어서 처리량을 증가시킨다. 특히, 만약 여러 개의 코드북들이 함께 결합된다면, 코드북에서 코드벡터들의 수의 곱셈의 수의 유사도를 구하기 위한 처리가 수행될 필요가 있어서 코드북탐색 처리량을 획기적으로 증가시키게 된다.

그러므로 본 발명의 목적은 코드북탐색 처리량을 억제할 수 있는 벡터양자화 방법, 음성부호화 방법 및 음성부호화 장치를 제공하는 것이다.

발명의 구성 및 작용

상기 목적을 성취하기 위해서, 본 발명은 고정도의 유사도를 보이는 여러 개의 코드벡터들을 예비선택하기 위한 근사화로 벡터양자화될 입력벡터와 코드북에 저장된 모든 코드벡터들과의 사이의 유사도를 구하는 단계와 입력벡터에 대해 오차를 최소화하는 여러 개의 예비 선택된 코드벡터들 중 하나를 최종 선택하는 단계를 포함한다.

예비 선택한 후에 최종 선택을 수행함으로써, 더 작은 수의 후보 코드벡터들이 단순화된 처리를 따르는 예비선택에 의해 선택되고 고정도의 최종 선택이 수행되어서 코드북탐색의 처리량을 감소시키게 된다.

코드북은 최적 결함을 나타내는 여러 개의 코드벡터들이 선택될 수 있는 여러 개의 코드북으로 구성된다. 유사도는 입력벡터와 코드벡터의 내적일 수도 있고, 선택적으로는 각 코드벡터의 가중된 노름(norm)이나 노름에 의해 분할된다.

또한 본 발명은 입력음성신호나 입력음성신호의 단기예측 나머지(residuals)가 조파의 스펙트럼성분을 구하기 위해서 사인파분석에 의해 분석되고 입력벡터로서, 부호화 단위에 기초한 조파의 스펙트럼성분들에서 유래하는 파라미터들이 부호화를 위해 벡터양자화 되는 음성부호화 방법을 제공한다. 벡터양자화에서, 입력벡터와 코드북에 저장된 코드벡터들 사이의 유사도는 고정도의 유사도를 가지는 더 작은 여러 개의 코드벡터들을 예비선택하기 위해 근사화에 의해 구해지고 입력벡터에 대해 오차를 최소화하는 이 예비 선택된 코드벡터들 중 하나가 최종 선택된다.

유사도는 노름과 각 코드벡터의 가중된 노름에 의해 선택적으로 분할되는 코드벡터와 입력벡터 사이의 선택적으로 가중된 내적일 수도 있다. 노름을 가중하기 위해서, 저주파수대역으로 집중된 에너지와 높은 주파수 대역으로 감소된 에너지를 가지는 가중이 이용될 수도 있다. 따라서, 유사도는 가중된 코드벡터 노름에 의해 코드벡터의 가중된 내적을 분할함으로써 구해질 수 있다.

본 발명은 또한 음성부호화 방법을 수행하기 위한 음성부호화 장치에 대해 설명한다.

도면을 참조하여 본 발명의 구체적 실시예들을 상세하게 설명할 것이다.

도 1은 본 발명에 따라서 음성부호화 방법을 수행하기 위한 부호화장치(부호기)의 기본적인 구조를 나타낸다.

도 1의 음성신호 부호기에 기초한 기본적인 개념은 부호기가 조파부호화(harmonic encoding)와 같은 사인파분석을 달성하기 위해서 입력음성신호의 가령 선형예측부호화(LPC) 나머지와 같은 단기예측 나머지를 구하기 위한 제 1부호화부(110)와, 위상 재생성(reprodducibility)을 가지는 과형부호화로 입력음성신호를 부호화하기 위한 제 2부호화부(120)를 가지며, 제 1부호화부(110)와 제 2부호화부(120)는 각각 입력신호의 유성(V)음을 부호화하고 입력신호의 무성(UV)음을 부호화 하는데 이용된다는 것이다.

제 1부호화부(110)는 조파부호화나 다중대역여기(MBE) 부호화와 같은 사인파분석 부호화로 예를 들어 LPC나머지를 부호화 하는 구성을 이용한다. 제 2부호화부(120)는 페루프 탐색에 의한 최적 벡터의 페루프 탐색에 의한 벡터양자화를 이용하고 또한 예를 들어 합성에 의한 분석법을 이용하는 코드여기된 선형예측(CELP)을 수행하는 구성을 이용한다.

도 1에 나타낸 실시예에서, 입력단자(101)에 공급된 음성신호는 제 1부호화부(110)의 LPC역필터(111)와 LPC분석양자화부(113)에 전송된다. LPC분석양자화부(113)에 의해 얻어진 LPC계수 또는 소위 α 파라미터는 제 1부호화부(110)의 LPC역필터(111)에 전송된다. LPC역필터(111)에서는 입력음성신호의 선형예측 나머지(LPC나머지)가 취출된다. LPC분석합성부(113)에서는, 선형스펙트럼쌍(LSP)들의 양자화된 출력이 취출되고 나중에 설명할 출력단자(102)에 전송된다. LPC역필터(111)로부터의 LPC나머지는 사인파분석 부호화부(114)에 전송된다. 사인파분석 부호화부(114)는 V/UV판정부(115)에 의한 V/UV판정뿐만 아니라 스펙트럼 엔벌로프의 진폭의 피치검출과 계산을 수행한다. 사인파분석 부호화부(114)로부터의 스펙트럼 엔벌로프 진폭데이터는 벡터양자화부(116)에 전송된다. 스펙트럼 엔벌로프의 벡터양자화된 출력으로서 벡터양자화부(116)로부터의 코드북 인덱스가 스위치(117)를 거쳐 출력단자(103)에 출력되는 반면, 사인파분석 부호화부(114)의 출력은 스위치(118)를 거쳐 출력단자(104)에 전송된다. V/UV판정부(115)의 V/UV판정출력은 출력단자(105)에 전송되고, 제어신호로서 스위치(107, 108)에 전송된다. 만약 입력음성신호가 유성(V)음이면, 인덱스와 피치가 선택되고 각각 출력단자(103, 104)에서 취출된다.

본 실시예에서, 도 1의 제 2부호화부(120)는 코드여기된 선형예측 부호화(CELP부호화) 구성을 가지며, 잡음코드북(121)의 출력이 가중합성필터에 의해 합성되는 합성에 의한 분석법을 이용하는 페루프 탐색을 이용하여 시간영역 과형을 벡터양자화하고, 그 결과의 가중된 음성은 감산기(123)에 전송되고, 가중된 음성과 입력단자(101)에서 청각가중필터(125)를 통해 공급된 음성신호 사이의 오차가 취출되고, 그와 같이 구해진 오차는 거리계산회로(124)에 전송되어서 거리계산을 하고, 그리고 오차를 최소화하는 벡터가 잡음코드북(121)에 의해 탐색된다. 이 CELP부호화는 이미 설명한 것처럼 무성음을 부호화 하기 위해 이용된다. 코드북 인덱스는 잡음코드북(121)으로부터의 UV데이터로서, V/UV판정의 결과가 무성음(UV)일 때 온이 되는 스위치(127)를 거쳐 출력단자(107)에서 취출된다.

도 2는 도 1의 음성신호 부호기의 대응장치로서, 본 발명에 따라서 음성복호화 방법을 수행하기 위한 음성신호 복호기의 기본적인 구조를 나타내는 블록도이다.

도 2를 참조하면, 코드북 인덱스는 도 1의 출력단자(102)로부터의 선형스펙트럼쌍(LSP)들의 양자화출력으로서, 입력단자(202)에 공급된다. 도 1의 출력단자(103, 104, 105)들의 출력들, 즉 피치, V/UV판정출력, 그리고 인덱스데이터는 엔벌로프 양자화 출력데이터로서, 각각 입력단자(203 ~ 205)에 공급된다. 무성음데이터에 대한 데이터가 도 1의 출력단자(107)에서 공급되는 것과 같은 인덱스데이터는 입력단자(207)에 공급된다.

입력단자(203)의 엔벌로프 양자화 출력으로서의 인덱스는 유성음합성기(211)에 전송되는 LPC나머지의 스펙트럼 엔벌로프를 구하기 위한 역벡터양자화용 역벡터양자화부(212)에 전송된다. 유성음합성기(211)는 사인파합성으로 유성음부분의 선형예측 부호화(LPC) 나머지를 합성한다. 합성기(211)에는 또한 입력단자(204, 205)로부터 V/UV판정출력과 피치가 공급된다. 유성음합성부(211)로부터의 유성음의 LPC나머지는 LPC합성필터(214)에 전송된다. 입력단자(207)로부터의 UV데이터의 인덱스데이터는 무성음부분의 LPC나머지를 취출하기 위한 잡음코드북을 참조하는 무성음합성부(220)에 전송된다. 이러한 LPC나머지들은 또한 LPC합성필터(214)에 전송된다. LPC합성필터(214)에서, 유성음부분의 LPC나머지와 무성음부분의 LPC나머지는 LPC합성으로 처리된다. 선택적으로, 함께 합해진 유성음부분의 LPC나머지와 무성음부분의 LPC나머지는 LPC합성으로 처리되어도 좋다. 입력단자(202)로부터의 LSP인덱스데이터는 LPC의 α 파라미터가 취출되고 LPC합성필터(214)로 전송되는 LPC파라미터 재생부(213)에 전송된다. LPC합성필터(214)에 의해 합성된 음성신호는 출력단자(201)에서 취출된다.

이제 도 3을 참조하여, 도 1에 나타난 음성신호부호기의 좀더 상세한 구조를 설명할 것이다. 도 3에서, 도 1에 나타난 성분과 유사한 부분이나 성분은 동일한 부호로 표시된다.

도 3에 나타난 음성신호부호기에서, 입력단자(101)에 공급된 음성신호는 불필요한 범위의 신호를 제거하기 위한 고역통과필터(HPF)(109)에 의해 필터링되고 LPC분석양자화부(113)의 LPC(선형 예측 부호화)분석회로(132)와 LPC역필터(111)에 공급된다.

LPC분석양자화부(113)의 LPC분석회로(132)는 블록으로서 256샘플 정도를 입력신호파형의 길이로 해밍창(Hamming window)을 적용하고, 선형예측계수 즉 소위 α 파라미터를 자기상관법으로 구한다. 데이터출력 단위로서의 프레임(framing) 간격은 대략 160샘플로 설정된다. 만약 샘플링주파수(fs)가 8kHz라면, 예를 들어 1프레임 간격은 20msec 또는 160샘플이다.

LPC분석회로(132)로부터의 α 파라미터는 선형스펙트럼쌍(LSP) 파라미터로 변환하기 위한 α -LSP변환회로(133)에 전송된다. 이것은 직접형 필터계수에 의해 구해지는 것과 같은 α 파라미터를 예를 들어 10 즉 5쌍의 LSP파라미터로 변환한다. 이 변환은 예를 들어 뉴턴-랩슨(Newton-Rhapson)법으로 수행된다. α 파라미터가 LSP파라미터로 변환되는 이유는 LSP파라미터가 α 파라미터에 비해 보간특성이 우수하기 때문이다.

α -LSP변환회로(133)로부터의 LSP파라미터는 LSP양자화기(134)에 의해 매트릭스 또는 벡터양자화 된다. 벡터양자화에 프레임 대 프레임 차이를 취하는 것이 가능하거나 매트릭스양자화를 수행하기 위해서 여러 개의 프레임을 수집하는 것이 가능하다. 본 경우에는, 20msec마다 계산된 LSP파라미터의 각 20msec 길이, 즉 2개의 프레임은 함께 다루어지고 매트릭스양자화와 벡터양자화로 처리된다.

양자화기(134)의 양자화된 출력, 즉 LSP양자화의 인덱스데이터는 단자(102)에서 취출되는 반면, 양자화된 LSP벡터는 LSP보간회로(136)에 전송된다.

LSP보간회로(136)는 옥테튜플(octatuple) 비율을 제공하기 위해서 20msec 또는 40msec마다 양자화된 LSP벡터를 보간한다. 즉, LSP벡터는 2.5msec마다 갱신된다.

그 이유는 만약 나머지파형이 조파 부호화/복호화 방법에 의한 분석/합성으로 처리된다면, 합성파형의 엔벌로프가 극히 스무스한 파형을 타나내고, 그래서 만약 LPC계수가 20msec마다 급격히 변한다면, 외부 잡음이 생성되기 쉽기 때문이다. 즉, 만약 LPC계수가 2.5msec마다 서서히 변한다면, 상기 외부 잡음은 발생하지 않을 수도 있다.

2.5msec마다 생성된 보간된 LSP벡터를 이용하는 입력음성의 역필터링 때문에, LSP파라미터는 LSP- α 변환회로(137)에 의해 예를 들어 10차 직접형 필터의 필터계수인 α 파라미터로 변환된다. LSP- α 변환회로(137)의 출력은 2.5msec마다 갱신된 α 파라미터를 이용하는 스무스한 출력을 생성하기 위해서 역필터링을 수행하는 LPC역필터회로(111)에 전송된다. LPC필터(111)의 출력은 조파부호화회로와 같은 사인파분석 부호화부(114)의 가령 DCT회로와 같은 직교변환회로(145)에 전송된다.

LPC분석양자화부(113)의 LPC분석회로(132)로부터의 α 파라미터는 청각가중용 데이터가 구해지는 청각가중필터 계산회로에 전송된다. 이러한 가중데이터는 벡터양자화기(116)와 제 2부호화부(120)의 청각가중필터(125)와 청각가중합성필터(122)에 전송된다.

조파부호화회로의 사인파분석 부호화부(114)는 조파부호화의 방법으로 LPC역필터(111)의 출력을 분석한다. 즉, 피치검출, 각 조파의 진폭(Am)의 계산, 그리고 유성음(V)/무성음(UV)판정이 수행되고 진폭(Am)의 수나 각 조파의 엔벌로프는 피치에 따라 변하며 차원변환에 의해 일정하게 된다.

도 3에 나타난 사인파분석 부호화부(114)의 도식적 예에서, 평범한 조파부호화가 이용된다. 특히, 다중대역여기(MBE) 부호화에서, 모델링시 유성음부분과 무성음 부분이 동일한 시점에서 각 주파수영역이나 대역(동일한 블록이나 프레임)에 나타난다고 가정한다. 다른 조파부호화 기술에서는, 1블록 또는 1프레임의 음성이 유성음인지 무성음인지가 유일하게 판정된다. 다음의 설명에서, 만약 MBE부호화가 고려된다는 점에서 대역들의 전체가 UV라면, 주어진 프레임은 UV라고 판정된다.

위에서 설명한 것처럼 MBE를 위한 분석합성법 기술의 구체적 예는 본 출원의 출원인의 이름으로 제출된 일본 특허출원 제 4-91442에서 구해질 수도 있다.

도 3의 사인파분석 부호화부(114)의 개방루프 피치탐색부(141)와 영교차(zero-crossing) 카운터(142)에는 입력단자(101)로부터 입력음성신호가 공급되고 고역통과필터(HPF)(109)로부터 신호가 각각 공급된다. 사인파분석 부호화부(114)의 직교변환회로(145)에는 LPC역필터(111)로부터 LPC나머지나 선형예측 나머지가 공급된다. 개방루프 피치탐색부(141)는 개방루프 탐색으로 비교적 대략적인 피치탐색을 수행하기 위해서 입력신호의 LPC나머지를 취한다. 추출된 대략적인 피치데이터는 나중에 설명할 것으로 페루프 탐색에 의해 양호한 피치탐색부(141)에 전송된다.

개방루프 피치탐색부(141)에서는, 대략적인 피치데이터와 함께 LPC나머지의 자기상관의 최대값을 정규화 하여 얻어지는 정규화된 자기상관의 최대값이 V/UV판정부(115)에 전송되도록 대략적인 피치데이터와 함께 추출된다.

직교변환회로(145)는 시간축상의 LPC나머지를 주파수축상의 스펙트럼 진폭데이터로 변환하기 위해서 이산 푸리에변환(DFT)과 같은 직교변환을 수행한다. 직교변환회로(145)의 출력은 양호한 피치탐색부(146)에 전송되고 스펙트럼진폭이나 엔벌로프를 평가하도록 구성된 스펙트럼평가부(148)에 전송된다.

양호한 피치탐색부(146)에는 개방루프 피치탐색부(141)에 의해 추출된 비교적 대략적인 피치데이터가 공급되고 직교변환회로(145)에 의한 DFT로 얻어진 주파수영역 데이터가 공급된다. 양호한 피치탐색부(146)는 최종으로 최적 소수점을 가지는 양호한 피치데이터의 값에 도달하도록 대략적인 피치값 데이터를 중심으로 0.2에서 0.5까지의 비율로 소트 샘플마다 피치데이터를 스윙(swing)한다. 합성에 의한 분석법은 피치를 선택하기 위한 양호한 탐색기술로서 이용되어서, 전력스펙트럼이 원음의 전력스펙트럼에 제일 가까워질 것이다. 양호한 페루프 피치탐색부(146)로부터의 피치데이터는 스위치(118)를 거쳐 출력단자(104)에 전송된다.

스펙트럼평가부(148)에서, 각 조파의 진폭과 조파의 합으로서의 스펙트럼 엔벌로프는 LPC나머지의 직교변환출력으로서의 피치와 스펙트럼 진폭에 근거해서 평가되고, 양호한 피치탐색부(146), V/UV판정부(115), 그리고 청각가중 벡터양자화부(116)에 전송된다.

V/UV판정부(115)는 직교변환회로(145)의 출력, 양호한 피치탐색부(146)로부터의 최적 피치, 스펙트럼평가부(148)로부터의 스펙트럼 진폭데이터, 개방루프 피치탐색부(141)로부터의 정규화된 자기상관 $r(p)$ 의 최대값, 그리고 영교차 카운터(142)로부터의 영교차 카운트값에 근거해서 프레임의 V/UV를 판별한다. 또한, MBE에 대한 대역에 기초한 V/UV판정의 경계부분이 V/UV판정에 대한 조건으로서 이용될 수도 있다. V/UV판정부(115)의 판정출력은 출력단자(105)에서 추출된다.

스펙트럼평가부(148)의 출력부 또는 벡터양자화부(116)의 입력부에는 많은 데이터변환부(어떤 종류의 샘플링을 변환을 수행하는 부)가 제공된다. 데이터변환부의 수는 주파수축 상에서 분할된 대역의 수와 데이터의 수가 피치와 다른 것을 고려하여 엔벌로프의 진폭데이터($|A_m|$)를 일정한 값으로 설정하는데 이용된다. 즉, 만약 유효 대역이 3400kHz까지라면, 유효 대역은 피치에 따라서 8에서 63대역까지로 분할될 수 있다. 1대역씩 얻어지는 진폭데이터($|A_m|$)의 $mMX+1$ 의 수는 8에서 63까지의 범위에서 변화된다. 따라서 데이터수 변환부는 변수($mMX+1$)의 진폭데이터를 44데이터와 같은 데이터의 소정수(M)로 변환한다.

스펙트럼평가부(148)의 출력부나 벡터양자화부(116)의 입력부에 설치되는 데이터수 변환부로부터의 가령 44와 같은 소정수(M)의 진폭데이터나 엔벌로프데이터는 가중벡터양자화를 수행하는 방법으로 벡터양자화부(116)에 의해 가령 44데이터와 같은 데이터의 소정수에 관하여 함께 다루어진다. 이 가중은 청각가중필터 계산회로(139)의 입력에 의해 공급된다. 벡터양자화기(116)로부터의 엔벌로프의 인덱스는 출력단자(103)에서 스위치(117)에 의해 추출된다. 가중벡터양자화 전에, 데이터의 소정수로 이루어진 벡터에 대한 적당한 누설(leakage)계수를 이용하는 인터프레임(inter-frame) 차이를 취하는 것이 바람직하다.

제 2부호화부(120)를 설명한다. 제 2부호화부(120)는 소위 CELP부호화 구조를 가지고 특히 입력음성신호의 무성음부분을 부호화 하는데 이용된다. 입력음성신호의 무성음부분에 대한 CELP부호화구조에서, 잡음코드북 또는 소위 스토캐스틱(stochastic) 코드북(121)의 대표적인 출력값으로서 무성음의 LPC나머지에 해당하는 잡음출력은 이득제어회로(126)를 거쳐 청각가중합성필터(122)에 전송된다. 가중합성필터(122)는 LPC합성으로 입력잡음을 LPC합성하고 생성된 가중 무성음신호를 감산기(123)에 전송한다. 감산기(123)에는 고역통과필터(HPF)(109)를 거쳐 입력단자(101)에서 공급된 신호가 공급되고 청각가중필터(125)에 의해 청각가중된다. 감산기는 신호와 합성필터(122)로부터의 신호 사이의 오차 또는 차

이를 구한다. 반면에, 청각가중합성필터의 영입력 응답은 미리 청각가중필터(125)의 출력에서 감소된다. 이 오차는 거리를 계산하기 위한 거리계산회로(124)에 공급된다. 오차를 최소화할 대표적인 벡터값은 잡음코드북(121)에서 탐색된다. 위 부분은 합성에 의한 분석법으로 페루프 탐색을 이용하는 시간영역 파형의 벡터양자화를 요약해 놓는 것이다.

CELP부호화 구조를 이용하는 제 2부호기(120)로부터의 무성음(UV)부분에 대한 데이터로서, 잡음코드북(121)으로부터의 코드북의 형상인덱스와 이득회로(126)로부터의 코드북의 이득 인덱스가 추출된다. 잡음코드북(121)으로부터의 UV데이터인 형상인덱스는 스위치(127s)를 거쳐 출력단자(107s)에 전송되는 반면, 이득회로(126)의 UV데이터인 이득인덱스는 스위치(127g)를 거쳐 출력단자(107g)에 전송된다.

이 스위치(127s, 127g)들과 스위치(117, 118)들은 V/UV판정회로(115)의 V/UV판정결과에 따라서 온/오프된다. 특히, 스위치(117, 118)는 현재 송신된 프레임의 음성신호의 V/UV판정결과가 유성음(V)이면 온이 되는 반면, 스위치(127s, 127g)는 현재 송신된 프레임의 음성신호가 무성음(UV)일 때 온이 된다.

도 4는 도 2에 나타난 음성신호 복호기의 좀더 상세한 구조를 나타낸다. 도 4에서 동일한 부호가 도 2에 나타난 대응부분을 나타내는데 이용된다.

도 4에서, 도 1 및 도 3의 출력단자(102)에 해당하는 LSP들의 벡터양자화출력, 즉 코드북 인덱스는 입력단자(202)에 공급된다.

LSP인덱스는 LPC파라미터 재생부(213)에 대한 LSP의 역벡터양자화기(231)에 전송되어서 다음에 보간용 LSP보간회로(232, 233)에 공급되는 선형스펙트럼쌍(LSP)데이터로 역벡터양자화 된다. 그 결과의 보간된 데이터는 LSP- α 변환회로(234, 235)에 의해 LPC합성필터(214)에 전송되는 α 파라미터로 변환된다. LSP보간회로(232)와 LSP- α 변환회로(234)는 유성음용으로 설계된 반면, LSP보간회로(233)와 LSP- α 변환회로(235)는 무성음용으로 설계된다. LPC합성필터(214)는 유성음부분의 LPC합성필터(236)와 무성음부분의 LPC합성필터(237)로 구성된다. 즉, LPC계수보간은 유성음부분과 무성음부분에 대해 독립적으로 수행되는데, 유성음부분에서 무성음부분까지 일시적인 부분에 생성될 수도 있는 나쁜 영향이나 반대로 전체적으로 다른 특성의 LSP의 보간에 의한 나쁜 영향을 방지하기 위해서이다.

도 4의 입력단자(203)에 도 1 및 도 3의 부호기의 단자(103)의 출력에 해당하는 가중벡터양자화된 스펙트럼 엔벨로프(Am)에 해당하는 코드 인덱스데이터가 공급된다. 입력단자(204)에는 도 1 및 도 3의 단자(104)에서 피치데이터가 공급되고, 입력단자(205)에는 도 1 및 도 3의 단자(105)에서 V/UV판정데이터가 공급된다.

입력단자(203)로부터의 스펙트럼 엔벨로프(Am)의 벡터양자화된 인덱스데이터는 데이터수 변환과 반대의 변환이 수행되는 역벡터양자화용 역벡터양자화기(212)에 전송된다. 그 결과의 스펙트럼 엔벨로프데이터는 사인파합성회로(215)에 전송된다.

만약 인터프레임 차이가 부호화 중에 스펙트럼의 벡터양자화전에 구해지면, 인터프레임 차이는 스펙트럼 엔벨로프데이터를 생성하기 위한 역벡터양자화 후에 복호화 된다.

사인파합성회로(215)에는 입력단자(204)로부터 피치가 공급되고 입력단자(205)로부터 V/UV판정데이터가 공급된다. 사인파합성회로(215)로부터는, 도 1 및 도 3에 나타난 LPC역필터(111)의 출력에 해당하는 LPC나머지 데이터가 추출되고 가산기(218)로 전송된다. 사인파합성의 구체적인 기술은 예를 들어 본 출원인에 의해 제안된 일본 특허 출원 제 4-91442와 제 6-19851에 설명되어 있다.

역벡터양자화기(212)의 엔벨로프데이터와 입력단자(204, 205)로부터의 피치 및 V/UV판정데이터는 유성음(V)부분에 대한 잡음가산용으로 구성된 잡음합성회로(216)에 전송된다. 잡음합성회로(216)의 출력은 가중 중첩가산회로(217)를 거쳐 가산기(218)에 전송된다. 특히, 만약 유성음의 LPC합성필터에 대한 입력으로서 여기(excitation)가 사인파형 합성에 의해 생성되고, 스템프트 필링(stuffed feeling)이 남성의 음성과 같은 낮은 피치음에서 생기고, 그리고 음질이 유성음과 무성음 사이에서 급격하게 변화된다면, 부자연스러운 청각을 유발시킨다는 것을 고려하면, 잡음은 LPC나머지 신호의 유성음부분에 가산된다. 유성음부분의 LPC합성필터 입력, 즉 여기에 관련하여, 상기 잡음은 피치, 스펙트럼 엔벨로프의 진폭, 프레임의 최대 진폭 또는 나머지신호 레벨과 같은 음성부호화 데이터에 관련한 파라미터를 없앤다.

가산기(218)의 합산 출력은 유성음에 대해 포스트필터(238v)에 의해 필터링되고 가산기(238)에 전송되는 시간파형 데이터를 형성하도록 LPC합성이 수행되는 LPC합성필터(214)의 유성음용 합성필터(236)에 전송된다.

도 3의 출력단자(107s, 107g)로부터의 UV데이터로서 형상인덱스와 이득인덱스는 도 4의 입력단자(207s, 207g)에 각각 공급되고, 다음에 무성음합성부(220)에 공급된다. 단자(207s)로부터의 형상인덱스는 무성음합성부(220)의 잡음코드북(221)에 전송되는 반면, 단자(207g)로부터의 이득인덱스는 이득회로(222)에 전송된다. 잡음코드북(221)에서 독출된 대표값 출력은 무성음의 LPC나머지에 해당하는 잡음신호 성분이다. 이것은 이득회로(222)에서 소정의 이득진폭이 되고 윈도우(windowing)회로(223)에 전송되어서 유성음부분과의 결합을 부드럽게 하기 위해서 윈도우(windowing)된다.

윈도우회로(223)의 출력은 LPC합성필터(214)의 무성음(UV)용 합성필터(237)에 전송된다. 합성필터(237)에 전송된 데이터는 LPC합성으로 처리되어서 무성음부분에 대한 시간과형 데이터가 된다. 무성음부분의 시간과형 데이터는 가산기(239)에 전송되기 전에 무성음부분용 포스트필터(238u)에 의해 필터링된다.

가산기(239)에서, 유성음용 포스트필터(238v)로부터의 시간과형 신호와 무성음용 포스트필터(238u)로부터의 시간과형 데이터는 서로 가산되고 그 결과의 합데이터는 출력단자(201)에서 취출된다.

위에서 설명한 음성신호부호기는 요구하는 음질에 따라서 다른 비트율(bit rate)의 데이터를 출력할 수 있다. 즉, 출력데이터는 다양한 비트율로 출력될 수 있다.

특히, 출력데이터의 비트율은 저비트율과 고비트율 사이에서 스위칭될 수 있다. 예를 들어, 만약 저비트율이 2kbps이고 고비트율이 6kbps이면, 출력데이터는 표 1에 타나낸 다음의 비트율을 갖는 비트율이 데이터이다.

출력단자(104)로부터의 피치데이터는 항상 유성음에 대해 8비트/20msec의 비트율로 출력되고, 출력단자(105)로부터의 V/UV판정출력은 항상 1비트/20msec이다. 출력단자(102)에서 출력된 LPS양자화용 인덱스는 32비트/40msec와 48비트/40msec 사이에서 스위칭된다. 다른 한편으로는, 출력단자(103)에 의해 출력된 유성음(V) 동안의 인덱스는 15비트/20msec와 87비트/20msec 사이에서 스위칭된다. 출력단자(107s, 107g)에서 출력된 무성음(UV)용 인덱스는 11비트/10msec와 23비트/5msec 사이에서 스위칭된다. 유성음(V)에 대한 출력데이터는 2kbps동안 40비트/20msec이고 6kbps동안 120비트/20msec이다. 다른 한편으로는, 무성음(UV)에 대한 출력데이터는 2kbps동안 39비트/20msec이고 6kbps동안 117비트/20msec이다.

LSP양자화용 인덱스, 유성음(V)용 인덱스, 그리고 무성음(UV)용 인덱스는 관련부분의 구성과 관련하여 이후에 설명할 것이다.

도 6과 도 7을 참조하여, LSP양자화기(134)의 매트릭스양자화(matrix quantization)와 벡터양자화를 상세하게 설명할 것이다.

LPC분석회로(132)로부터의 α 파라미터는 LSP파라미터로 변환하기 위한 α -LSP회로(133)에 전송된다. 만약 P차 LPC분석이 LPC분석회로(132)에서 수행되면, P개의 α 파라미터가 계산된다. 이러한 P개의 α 파라미터는 버퍼(610)에 유지되는 LSP파라미터로 변환된다.

버퍼(610)는 LSP파라미터의 2프레임을 출력한다. LSP파라미터의 2프레임은 제 1매트릭스양자화기(620₁)와 제 2매트릭스양자화기(620₂)로 구성된 매트릭스양자화기(620)에 의해 매트릭스양자화 된다. LSP파라미터의 2프레임은 제 1매트릭스양자화기(620₁)에서 매트릭스양자화 되고, 그 결과의 양자화 오차가 제 2매트릭스양자화기(620₂)에서 더 매트릭스양자화 된다. 매트릭스양자화는 시간축과 주파수축 양자의 상관(correlation)을 이용한다.

매트릭스양자화기(620₂)로부터의 2프레임에 대한 양자화 오차는 제 1벡터양자화기(640₁)와 제 2벡터양자화기(640₂)로 구성된 벡터양자화부(640)로 들어간다. 제 1벡터양자화기(640₁)는 2개의 벡터양자화부(650, 660)로 구성되며, 제 2벡터양자화기(640₂)는 2개의 벡터양자화부(670, 680)로 구성된다. 매트릭스양자화부(620)로부터의 양자화 오차는 제 1벡터양자화기(640₁)의 벡터양자화부(650, 660)에 의해 프레임마다 양자화 된다. 그 결과의 양자화 오차벡터는 제 2벡터양자화기(640₂)의 벡터양자화부(670, 680)에 의해 더 양자화 된다. 위에서 설명한 벡터양자화는 주파수축을 따라 상관을 이용한다.

위에서 설명한 것처럼 매트릭스양자화를 실행하는 매트릭스양자화부(620)는 적어도 제 1매트릭스양자화 단계를 수행하기 위한 제 매트릭스양자화기(620₁)와 제 1매트릭스양자화기에 의해 생성된 양자화 오차를 매트릭스양자화 하기 위한 제 2매트릭스양자화 단계를 수행하기 위한 제 2매트릭스양자화기(620₂)를 포함한다. 위에서 설명한 것과 같은 벡터양자화를 실행하는 벡터양자화부(640)는 적어도 제 1벡터양자화 단계를 수행하기 위한 제 1벡터양자화기(640₁)와 제 1벡터양자화에 의해 생성된 양자화 오차를 벡터양자화하는 제 2벡터양자화 단계를 수행하기 위한 제 2벡터양자화기(640₂)를 포함한다.

매트릭스양자화와 벡터양자화를 지금부터 상세하게 설명할 것이다.

버터(660)에 저장된 2프레임에 대한 LSP파라미터, 즉 10×2 매트릭스는 제 1매트릭스양자화기(620₁)에 전송된다. 제 1매트릭스양자화기(620₁)는 2프레임에 대한 LSP파라미터를 LSP파라미터 가산기(621)를 거쳐 최소값의 가중거리를 구하기 위한 가중거리 계산부(623)에 전송한다.

제 1매트릭스양자화기(620₁)에 의한 코드북 탐색동안 왜곡량(d_{MQ1})은 수학식 1로 주어진다.

[수학식 1]

$$d_{MQ1}(X_1, X'_1) = \sum_{t=0}^1 \sum_{i=1}^P w(t, i) (x_1(t, i) - x'_1(t, i))^2$$

여기서 X₁은 LSP파라미터이고 X'₁'은 양자화값이며, t와 i는 P차원의 수이다.

주파수축과 시간축으로의 가중제한이 고려되지 않는 가중(w)은 수학식 2로 주어진다.

[수학식 2]

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

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

수학식 2의 가중(w)은 또한 후단의 매트릭스양자화와 벡터양자화에 이용된다.

계산된 가중거리는 매트릭스양자화용 매트릭스양자화기(MQ₁)(622)에 전송된다. 이 매트릭스양자화에 의해 출력된 8비트 인덱스는 신호전환기(690)에 전송된다. 매트릭스양자화에 의한 양자화값은 가산기(621)에서 버터(610)로부터의 2프레임에 대한 LSP파라미터로부터 감산된다. 가중거리 계산부(623)는 2프레임마다 가중거리를 계산하여서 매트릭스양자화가 매트릭스양자화부(622)에서 수행된다. 또한, 가중거리를 최소화하는 양자화값이 선택된다. 가산기(621)의 출력은 제 2매트릭스양자화기(620₂)의 가산기(631)에 전송된다.

제 1매트릭스양자화기(620₁)와 마찬가지로, 제 2매트릭스양자화기(620₂)도 매트릭스양자화를 수행한다. 가산기(621)의 출력은 가산기(631)를 거쳐 최소 가중거리가 계산되는 가중거리 계산부(633)에 전송된다.

제 2매트릭스양자화기(620₂)에 의한 코드북 탐색동안 왜곡량(d_{MQ2})은 수학식 3으로 주어진다.

[수학식 3]

$$d_{MQ2}(X_2, X'_2) = \sum_{t=0}^1 \sum_{i=1}^P w(t, i) (x_2(t, i) - x'_2(t, i))^2$$

가중거리는 매트릭스양자화용 매트릭스양자화부(MQ₂)(632)에 전송된다. 매트릭스양자화에 의해 출력된 8비트 인덱스는 신호전환기(690)에 전송된다. 가중거리 계산부(633)는 계속해서 가산기(631)의 출력을 이용하는 가중거리를 계산한다. 가중거리를 최소화하는 양자화값이 선택된다. 가산기(631)의 출력은 1프레임씩 제 1벡터양자화기(640₁)의 가산기(651, 661)에 전송된다.

제 1벡터양자화기(640₁)는 1프레임씩 벡터양자화를 수행한다. 가산기(631)의 출력은 1프레임씩 가산기(651, 661)를 거쳐 최소 가중거리를 계산하기 위한 가중거리 계산부(653, 663)의 각각에 전송된다.

양자화 오차(X₂)와 양자화 오차(X'₂) 사이의 차이는 (10×2)의 매트릭스이다. 만약 차이가 $\mathbf{X}_2 - \mathbf{X}'_2 = [\mathbf{x}_{3-1}, \mathbf{x}_{3-2}]$ 과 같이 표현되면, 제 1벡터양자화기(640₁)의 벡터양자화부(652, 662)에 의한 코드북 탐색동안 왜곡량(d_{VQ1}, d_{VQ2})은 수학식 4와 수학식 5로 주어진다.

[수학식 4]

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

[수학식 5]

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

가중거리는 벡터양자화용 벡터양자화부(VQ₂)(662)와 벡터양자화부(VQ₁)(652)에 전송된다. 이 벡터양자화에 의해 출력된 각 8비트 인덱스는 신호전환기(690)에 전송된다. 양자화값은 가산기(651, 661)에 의해 입력 2프레임 양자화 오차벡터에서 감산된다. 가중거리 계산부(653, 663)는 계속해서 가중거리를 최소화하는 양자화값을 선택하기 위해서 가산기(651, 661)의 출력을 이용하는 가중거리를 계산한다. 가산기(651, 661)의 출력은 제 2벡터양자화기(640₂)의 가산기(671, 681)에 전송된다.

제 2벡터양자화기(640₂)의 벡터양자화기(672, 682)에 의한 코드북 탐색동안 왜곡량(d_{VQ3}, d_{VQ4})은

$$\mathbf{x}_{4-1} = \mathbf{x}_{3-1} - \mathbf{x}_{3-1}'$$

$$\mathbf{x}_{4-2} = \mathbf{x}_{3-2} - \mathbf{x}_{3-2}'$$

인 동안에 수학식 6과 수학식 7로 주어진다.

[수학식 6]

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

[수학식 7]

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

가중거리는 벡터양자화용 벡터양자화기(VQ₄)(682)와 벡터양자화기(VQ₃)(672)에 전송된다. 벡터양자화기로부터의 8비트 출력 인덱스데이터는 가산기(671, 681)에 의해 2프레임에 대한 입력 양자화 오차벡터에서 감산된다. 가중거리 계산부(673, 683)는 계속해서 가중거리를 최소화하는 양자화값을 선택하기 위해서 가산기(671, 681)의 출력을 이용하는 가중거리를 계산한다.

코드북 학습(learning)동안, 학습은 각 왜곡량에 근거해서 일반적인 로이드(Lloyd) 알고리즘에 의해 수행된다.

코드북 탐색동안 그리고 학습동안 왜곡량은 다른 값이 될 수도 있다.

매트릭스양자화부(622, 632)와 벡터양자화부(652, 662, 672, 682)로부터의 8비트 인덱스데이터는 신호전환기(690)에 의해 스위칭되고 출력단자(691)에 출력된다.

특히, 저비트율에 대해서는, 제 1매트릭스양자화 단계를 수행하는 제 1매트릭스양자화기(620₁), 제 2매트릭스양자화 단계를 수행하는 제 2매트릭스양자화기(620₂), 그리고 제 1벡터양자화 단계를 수행하는 제 1벡터양자화기(640₁)의 출력이 취출되는 반면, 고비트율에 대해서는, 저비트율에 대한 출력이 제 2벡터양자화 단계를 수행하는 제 2벡터양자화기(640₂)의 출력에 합산되고 그 결과의 합이 취출된다.

이것은 32비트/40msec의 인덱스와 48비트/40msec의 인덱스를 각각 2kbps와 6kbps동안 출력한다.

매트릭스양자화부(620)와 벡터양자화부(640)는 LPC계수를 나타내는 파라미터의 특성에 일치해서 주파수축 그리고/또는 시간축으로 제한된 가중을 수행한다.

LSP파라미터의 특성에 일치해서 주파수축으로 제한된 가중을 먼저 설명한다. 만약 차수가 P=10이라면, LSP파라미터(X(i))는 고, 중, 저 3개의 범위에 대해

$$L_1 = \{X(i) | 1 \leq i \leq 2\}$$

$$L_2 = \{X(i) | 3 \leq i \leq 6\}$$

$$L_3 = \{X(i) | 7 \leq i \leq 10\}$$

로 묶인다. 만약 그룹(L₁, L₂, L₃)의 가중이 각각 1/4, 1/2, 1/4라면, 주파수축으로만 제한된 가중은 수학식 8과 수학식 9와 수학식 10으로 주어진다.

[수학식 8]

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

[수학식 9]

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

[수학식 10]

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

각 LSP파라미터의 가중은 각 그룹에서만 수행되고 상기 가중은 각 그룹에 대한 가중에 의해 제한된다.

시간축으로 보았을 때, 각 프레임의 합계는 필수적으로 1이어서, 시간축방향으로의 제한은 프레임에 기초한다. 시간축 방향으로만 제한된 가중은 수학식 11로 주어진다.

[수학식 11]

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

여기서 $1 \leq i \leq 10$ 이고 $0 \leq t \leq 1$ 이다.

이 수학적 식 11에 의해서, 주파수축 방향으로 제한되지 않은 가중은 $t=0$ 과 $t=1$ 의 프레임을 가지는 2개의 프레임 사이에서 수행된다. 이 시간축 방향으로만 제한된 가중은 매트릭스양자화로 처리된 2개의 프레임 사이에서 수행된다.

학습동안, 총 수(T)를 가지는 학습데이터로서 이용된 프레임의 총계는 수학적 식 12에 따라서 가중된다.

[수학적 식 12]

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

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

주파수축 방향과 시간축 방향으로 제한된 가중을 설명한다. 차수가 $P=10$ 이면, LSP파라미터($x(i,t)$)는 고, 중, 저 3개의 범위에 대해

$$L_1 = \{x(i, t) | 1 \leq i \leq 2, 0 \leq t \leq 1\}$$

$$L_2 = \{x(i, t) | 3 \leq i \leq 6, 0 \leq t \leq 1\}$$

$$L_3 = \{x(i, t) | 7 \leq i \leq 10, 0 \leq t \leq 1\}$$

으로 묶인다. 만약 그룹(L_1, L_2, L_3)에 대한 가중이 $1/4, 1/2, 1/4$ 라면, 주파수축으로만 제한된 가중은 수학적 식 13, 수학적 식 14, 그리고 수학적 식 15로 주어질 것이다.

[수학적 식 13]

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

[수학적 식 14]

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

[수학적 식 15]

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

수학적 식 13과 수학적 식 14와 수학적 식 15에 의해서, 가중제한이 주파수축 방향으로 3프레임마다 수행되고 시간축 방향으로 매트릭스양자화로 처리된 2개의 프레임에 걸쳐 수행된다. 이것은 코드북 탐색동안 그리고 학습동안에 효과적이다.

학습동안, 가중은 전체 데이터의 프레임의 총계이다. LSP파라미터($x(i,t)$)는 고, 중, 저 3개 범위에 대해서

$$L_1 = \{x(i, t) | 1 \leq i \leq 2, 0 \leq t \leq T\}$$

$$L_2 = \{x(i, t) | 3 \leq i \leq 6, 0 \leq t \leq T\}$$

$$L_3 = \{x(i, t) | 7 \leq i \leq 10, 0 \leq t \leq T\}$$

으로 묶인다. 만약 그룹(L_1, L_2, L_3)에 대한 가중이 각각 $1/4, 1/2, 1/4$ 라면, 주파수축과 시간축 방향으로 제한된 그룹(L_1, L_2, L_3)에 대한 가중은 수학적 식 16과 수학적 식 17과 수학적 식 18로 주어진다.

[수학식 16]

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

[수학식 17]

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

[수학식 18]

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

수학식 16과 수학식 17과 수학식 18에 의해서, 가중은 주파수축 방향으로 3개의 범위에 대해서 수행될 수 있고 시간축 방향으로 프레임의 총계에 걸쳐 수행될 수 있다.

또한, 매트릭스양자화부(620)와 벡터양자화부(640)는 LSP파라미터의 변화의 크기에 따라서 가중을 수행한다. 음성프레임의 총계 중에서 소수(minority) 프레임을 나타내는 V에서 UV까지 또는 UV에서 V까지의 일시적인 영역에서, LSP파라미터는 주로 자음과 모음 사이의 주파수 응답의 차이에 의해서 변화된다. 그러므로, 수학식 19로 나타난 가중에 일시적인 영역을 강조하는 가중을 수행하기 위해서 가중(W'(i,t))이 곱해진다.

[수학식 19]

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

다음의 수학식 20은 수학식 19대신에 이용될 수도 있다.

[수학식 20]

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

따라서 LSP양자화부(134)는 2단 매트릭스양자화와 2단 벡터양자화를 실행하여서 출력인텍스의 비트수를 변하게 한다.

벡터양자화부(116)의 기본적인 구조를 도 8에 나타내며, 도 8에 나타난 벡터양자화부(116)의 좀더 상세한 구조를 도 9에 나타낸다. 벡터양자화부(116)에서 엔벨로프(Am)에 대한 가중벡터 양자화의 도식적 구조를 지금부터 설명한다.

먼저, 도 3에 나타난 음성신호 부호화장치에서, 스펙트럼 평가부(148)의 출력측 또는 벡터양자화부(116)의 입력측에 스펙트럼 엔벨로프의 진폭데이터의 일정한 수를 제공하기 위한 데이터수 변환에 대한 도식적 구조를 설명한다.

다양한 방법이 상기 데이터수 변환을 위해 고안될 수도 있다. 본 실시예에서, 1블록의 마지막 데이터에서 그 블록의 처음 데이터까지의 값들을 보간하는 더미데이터나 또는 1블록에서 마지막 데이터나 처음 데이터를 반복하는 데이터와 같은 소정의 데이터가 주파수축상의 유효대역에 1블록의 진폭데이터에 첨부하고, 데이터의 수를 N_F까지 증가시키고 대역폭 제한형의 예를 들어 8배 오버샘플링인 0s오버샘플링에 의해 예를 들어 8배인 0s배의 개수인 진폭데이터를 구한다. ((mMx+1)×0s) 진폭데이터는 2048과 같은 더 큰 N_M 수로 확장하기 위해 선형보간된다. 이 N_M 데이터는 44데이터와 같은 위에서 설명한 데이터의 상기 소정수(M)로 변환하기 위해 부샘플링(sub-sampling)된다. 실제로, 최종적으로 요구되는 M데이터를 표현하는데 필요한 데이터만이 위에서 설명한 N_M 데이터의 모든 것을 구하지 않고도 오버샘플링과 선형보간에 의해 계산된다.

도 8의 가중벡터양자화를 수행하기 위한 벡터양자화부(116)는 적어도 제 1벡터양자화 단계를 수행하기 위한 제 1벡터양자화부(500)와 제 1벡터양자화부(500)에 의한 제 1벡터양자화 동안 생성된 양자화 오차벡터를 양자화 하는 제 2벡터양자화 단계를 수행하기 위한 제 2벡터양자화부(510)를 포함한다. 이 제 1벡터양자화부(500)는 소위 제 1단 벡터양자화부인 반면, 제 2벡터양자화부(510)는 소위 제 2단 벡터양자화부이다.

스펙트럼 평가부(148)의 출력벡터(\underline{x}), 즉 소정수(M)를 가지는 엔벌로프는 제 1벡터양자화부(500)의 입력단자(501)에 들어간다. 이 출력벡터(\underline{x})는 벡터양자화부(502)에 의해 가중벡터양자화로 양자화 된다. 따라서 벡터양자화부(502)에 의해 출력된 형상인덱스는 출력단자(503)에서 출력되는 반면, 양자화값(\underline{x}_0')은 출력단자(504)에서 출력되고 가산기(505, 513)에 전송된다. 가산기(505)는 양자화값(\underline{x}_0')을 소스(source)벡터(\underline{x})에서 감산하여서 다중차수 양자화 오차벡터(\underline{y})를 제공한다.

양자화 오차벡터(\underline{y})는 제 2벡터양자화부(510)의 벡터양자화부(511)에 전송된다. 이 제 2벡터양자화부(511)는 여러 개의 벡터양자화기 또는 도 8에서 2개의 벡터양자화기(511₁, 511₂)로 구성된다. 양자화 오차벡터(\underline{y})는 2개의 벡터양자화기(511₁, 511₂)에서 가중벡터양자화에 의해 양자화되도록 차원으로 분할된다. 이 벡터양자화기(511₁, 511₂)에 의해 출력된 형상인덱스는 출력단자(512₁, 512₂)에 출력되는 반면, 양자화값($\underline{y}_1', \underline{y}_2'$)은 차원방향으로 연결되고 가산기(513)에 전송된다. 가산기(513)는 양자화값($\underline{y}_1', \underline{y}_2'$)을 양자화값(\underline{x}_0')에 가산하여서 출력단자(514)에서 출력되는 양자화값(\underline{x}_1')을 생성한다.

따라서, 저비트율에 대해서, 제 1벡터양자화부(500)에 의한 제 1벡터양자화 단계의 출력이 추출되는 반면, 고비트율에 대해서 제 1벡터양자화 단계의 출력과 제 2양자화부(510)에 의한 제 2양자화 단계의 출력이 출력된다.

특히, 벡터양자화부(116)에 있는 제 1벡터양자화부(500)의 벡터양자화기(502)는 도 9에 나타난 것처럼 44차원의 2단 구조와 같은 L차수이다.

즉, 이득(g_i)이 곱해진 32의 코드북 크기를 갖는 44차원의 벡터양자화 코드북의 출력벡터의 합은 44차원의 스펙트럼 엔벌로프 벡터(\underline{x})의 양자화값(\underline{x}_0')으로 이용된다.

따라서, 도 8에 나타난 것처럼, 2개의 코드북이 CB0과 CB1인 반면, 출력벡터는 \underline{s}_{1i} 와 \underline{s}_{1j} 이며 여기서 $0 \leq i$ 이고 $j \leq 31$ 이다. 다른 한편으로는, 이득코드북(CB_g)의 출력은 g_1 이며, 여기서 $0 \leq 1 \leq 31$ 이고 g_1 은 스칼라이다. 최종 출력(\underline{x}_0')은 $g_1(\underline{s}_{1i} + \underline{s}_{1j})$ 이다.

상기 LPC나머지의 MBE분석에 의해 구해지고 소정의 차원으로 변환된 스펙트럼 엔벌로프(A_m)는 \underline{x} 이다. 어떻게 \underline{x} 가 효과적으로 양자화 되는지가 결정적이다.

양자화 오차에너지(E)는

[수학식 21]

$$E = \|W\{H\underline{x} - Hg_1(\underline{s}_{0i} + \underline{s}_{1j})\}\|^2$$

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

으로 정의되며, H는 LPC합성필터의 주파수축상의 특성을 나타내고 W는 주파수축상의 청각가중에 대한 특성을 나타내기 위한 가중용 매트릭스를 나타낸다.

만약 현재 프레임의 LPC분석의 결과에 의한 α 파라미터가 $\alpha_i(1 \leq i \leq P)$ 로 표시된다면, 예를 들어 점들에 해당하는 44차원인 L차원의 값들이 수학식 22의 주파수응답으로부터 샘플링된다.

[수학식 22]

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

계산을 위해서, 예를 들어 256점 데이터를 제공하기 위해서 1, $a_1, a_2, \dots, a_p, 0, 0, \dots, 0$ 의 열을 제공하도록 1, a_1, a_2, \dots, a_p ,의 열 다음에 0들이 채워진다. 다음에, 256점 FFT함으로서, $(r_e^2 + im^2)^{1/2}$ 가 0에서 π 까지의 범위와 관련된 점들에 대해 계산되고 그 결과의 역수들이 구해진다. 이러한 역수들은 44점과 같은 L점으로 부샘플링되고, 매트릭스가 이러한 L 점을 대각선 성분으로 가지면서 형성된다.

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

청각가중매트릭스(W)는 수학식 23으로 주어진다.

[수학식 23]

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

여기서 a_i 는 LPC분석의 결과이고, λ_a 와 λ_b 는 $\lambda_a=0.4$ 와 $\lambda_b=0.9$ 와 같은 상수이다.

매트릭스(W)는 상기 수학식 23의 주파수응답으로부터 계산될 수도 있다. 예를 들어, FFT가 $0 \leq i \leq 128$ 인 곳에서 0에서 π 까지의 범위에 대해 $(re^2[i] + Im_2[i])^{1/2}$ 를 구하기 위해서 1, $a_1\lambda_b, a_2\lambda_b^2, \dots, a_p\lambda_b^p, 0, 0, \dots, 0$ 의 256점 데이터에 대해 실행된다. 분모의 주파수응답은 $0 \leq i \leq 128$ 인 곳에서 $(re'^2[i] + im'^2[i])^{1/2}$ 을 구하기 위해서 128점에서 1, $a_1\lambda_b, a_2\lambda_b^2, \dots, a_p\lambda_b^p, 0, 0, \dots, 0$ 에 대해 0에서 π 까지의 영역에 대해 256점 FFT에 의해 구해진다. 수학식 23의 주파수응답은

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

에 의해 구해지며, $0 \leq i \leq 128$ 이다.

이것은 예를 들어 다음의 방법으로 44차원의 벡터의 각 관련된 점에 대해 구해진다. 좀더 자세하게는, 선형보간이 이용되어야 한다. 하지만, 다음의 예에서, 가장 가까운 점이 대신 이용된다.

즉,

$$\omega[i] = \omega_0[\text{nint}(128i/L)], 1 \leq i \leq L \text{이다.}$$

수학식에서 $\text{nint}(X)$ 는 X에 가장 가까운 값을 돌려주는 함수이다.

H에 대해서, $h(1), h(2), \dots, h(L)$ 가 유사한 방법으로 구해진다. 즉, 다음과 같다.

[수학식 24]

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

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

다른 예로서, $H(z)W(z)$ 가 먼저 구해지고 다음에 주파수응답이 FFT의 횟수를 감소시키기 위해서 구해진다. 즉,
[수학식 25]

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

수학식 25의 분모는

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

으로 확장된다. 예를 들어 256점 데이터는 $1, \beta_1, \beta_2, \dots, \beta_{2P}, 0, 0, \dots, 0$ 의 열을 이용하여 생성된다. 다음에, 256점 FFT가 진폭이

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

인 진폭의 주파수응답과 함께 실행되며, $0 \leq i \leq 128$ 이다. 이것으로부터,

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

이며, $0 \leq i \leq 128$ 이다. 이것은 L차원 벡터의 각 해당 점들에 대해 구해진다. 만약 FFT의 점들의 수가 작다면, 선형보간이 이용되어야 한다. 하지만, 여기서는 가장 가까운 값이

$$wh[i] = wh_0\left[\text{rint}\left(\frac{128}{L} \cdot i\right)\right]$$

으로 구해지며, $1 \leq i \leq L$ 이다. 만약 대각선 성분으로서 이것들을 가지는 매트릭스가 W'라면, 다음과 같이 된다.

[수학식 26]

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

수학식 26은 상기 수학식 24와 동일한 매트릭스이다.

선택적으로, $|H(\exp(j\omega))W(\exp(j\omega))|$ 은 $\omega = i\pi$ 에 대해서 수학식 25로부터 직접 계산될 수도 있으며, $1 \leq i \leq 128$ 이며, $wh[i]$ 에 이용된다.

선택적으로, 수학식 25의 임펄스(impulse)응답의 가령 40점과 같은 적당한 길이는 이용된 진폭의 주파수응답을 구하기 위해서 구해지고 FFT될 수도 있다.

$$E = \|W_k'(\underline{x} - \underline{g}_k(\underline{s}_{0c} + \underline{s}_{ik}))\|^2$$

이 매트릭스를 이용하여 수학식 21을 다시 쓰면, 즉, 가중합성의 주파수특성이 필터링되면,

[수학식 27]

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

이 구해진다.

형상코드북과 이득코드북의 학습을 위한 방법을 설명한다.

왜곡의 기대값(expected value)은 코드벡터(s_{0c})가 CBO에 대해 선택되는 모든 프레임(k)에 대해 최소화된다. 만약 프레임이 M개 있다면,

[수학식 28]

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

이 최소화된다면 만족된다. 수학식 28에서, W_k' , x_k , g_k 그리고 s_{1k} 는 각각 k번째 프레임에 대한 가중, k번째 프레임에 대한 출력, k번째 프레임의 이득, 그리고 k번째 프레임에 대한 코드북(CB1)의 출력을 나타낸다.

수학식 28을 최소화하기 위해서,

[수학식 29]

$$\begin{aligned} J &= \frac{1}{M} \sum_{k=1}^M \{ (x_k^T - g_k(s_{0c}^T + s_{1k}^T)) W_k'^T W_k' (x_k - g_k(s_{0c} + s_{1k})) \} \\ &= \frac{1}{M} \sum_{k=1}^M \{ x_k^T W_k'^T W_k' x_k - 2g_k(s_{0c}^T + s_{1k}^T) W_k'^T W_k' x_k \\ &\quad + g_k^2(s_{0c}^T + s_{1k}^T) W_k'^T W_k' (s_{0c} + s_{1k}) \} \\ &= \frac{1}{M} \sum_{k=1}^M \{ x_k^T W_k'^T W_k' x_k - 2g_k(s_{0c}^T + s_{1k}^T) W_k'^T W_k' x_k \\ &\quad + g_k^2 s_{0c}^T W_k'^T W_k' s_{0c} + 2g_k^2 s_{0c}^T W_k'^T W_k' s_{1k} + g_k^2 s_{1k}^T W_k'^T W_k' s_{1k} \} \end{aligned}$$

[수학식 30]

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

이 된다.

따라서,

$$\sum_{k=1}^M (g_k W_k'^T W_k' x_k - g_k^2 W_k'^T W_k' s_{1k}) = \sum_{k=1}^M g_k^2 W_k'^T W_k' s_{0c} \text{ 에서}$$

[수학식 31]

$$s_{0c} = \left\{ \sum_{k=1}^M g_k^2 W_k'^T W_k' \right\}^{-1} \cdot \left\{ \sum_{k=1}^M g_k W_k'^T W_k' (x_k - g_k s_{1k}) \right\}$$

이며, $\{ \}^{-1}$ 은 역매트릭스를 나타내고 $W_k'^T$ 는 W_k' 의 전치매트릭스를 나타낸다.

다음에 이득최적화를 고려한다.

이득의 코드워드(\$g_c\$)를 선택하는 \$k\$번째 프레임을 고려한 왜곡의 기대값은 다음과 같이 주어진다.

$$J_g = \frac{1}{M} \sum_{k=1}^M \|W_k'(x_k - g_c(s_{0k} + s_{1k}))\|^2$$

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

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

을 풀면

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

[수학식 32]

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

을 구한다.

수학식 31과 수학식 32는 \$0 \le i \le 31\$, \$0 \le j \le 31\$, 그리고 \$0 \le l \le 31\$인 동안 형상(\$s_{0i}\$, \$s_{1i}\$)과 이득(\$g_l\$)에 대한 최적 중심(centroid) 조건을 제공한다. 즉, 최적의 부호기 출력이다. 반면에, \$s_{1i}\$은 \$s_{0i}\$에 대해 동일한 방법으로 구해질 수도 있다.

다음에 최적의 부호화 조건, 즉 가장 가까운 인접(neighbor)조건이 고려된다.

식 \$E = \|W'(x - g_l(s_{0i} + s_{1i}))\|^2\$을 최소화하는 \$s_{0i}\$와 \$s_{1i}\$인 왜곡량을 구하기 위한 상기 수학식 27이 입력(\$x\$)과 가중매트릭스(\$W\$)이 주어질 때마다, 즉 프레임 단위마다 구해진다.

본래, \$E\$는 \$E\$의 최소값을 제공할 \$(s_{0i}, s_{1i})\$의 세트를 구하기 위해서, \$g_l\$ (\$0 \le l \le 31\$), \$s_{0i}\$ (\$0 \le i \le 31\$), 그리고 \$s_{1i}\$ (\$0 \le j \le 31\$)의 모든 조합, 즉 \$32 \times 32 \times 32 = 32768\$동안 라운드로빈(round robin)형으로 구해진다. 하지만, 이것은 많은 계산을 요하기 때문에, 형상과 이득은 차례로 본 실시예에서 탐색된다. 반면에, 라운드로빈 탐색은 \$s_{0i}\$과 \$s_{1i}\$의 조합에 대해 이용된다. \$s_{0i}\$과 \$s_{1i}\$에 대해 \$32 \times 21 = 1024\$ 조합이 있다.

다음의 설명에서, \$s_{0i} + s_{1i}\$는 간단히 \$s_m\$으로 표시된다.

상기 수학식 27은 \$E = \|W'(x - g_l s_m)\|^2\$ 된다. 만약 더 간단히 \$x_w = W' x\$이고 \$s_w = W' s_m\$라면, 다음의 식을 구할 수 있다.

[수학식 33]

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

[수학식 34]

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

그러므로 만약 \$g_l\$이 충분히 정확히 될 수 있다면, 탐색은 다음의 2단계로 수행될 수 있다.

(1) 다음 식을 최대화 하는 \mathbf{s}_w 를 탐색하고,

$$\frac{(\mathbf{x}_w^T \cdot \mathbf{s}_w)^2}{\|\mathbf{s}_w\|^2}$$

(2) 다음 식에 가장 가까운 g_1 을 탐색한다.

$$\frac{\mathbf{x}_w^T \cdot \mathbf{s}_w}{\|\mathbf{s}_w\|^2}$$

만약 상기 식을 원래의 표기를 이용하여 다시 쓰면.

(1)' 다음 식을 최대화 하는 \mathbf{s}_{0i} 와 \mathbf{s}_{1i} 에 대해 탐색이 되고,

$$\frac{(\mathbf{x}^T \mathbf{W}'^T \mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1i}))^2}{\|\mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1i})\|^2}$$

(2)' 다음의 식에 가장 가까운 g_1 에 대해 탐색이 이루어진다.

[수학식 35]

$$\frac{(\mathbf{x}^T \mathbf{W}'^T \mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1i}))^2}{\|\mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1i})\|^2}$$

상기 수학식 35는 최적 부호화 조건(가장 가까운 인접조건)을 나타낸다.

벡터양자화를 위해 코드북 탐색을 실행하는 경우의 처리량을 지금부터 고려한다.

K의 \mathbf{s}_{0i} 와 \mathbf{s}_{1i} 의 차원으로, 그리고 L_0 과 L_1 의 코드북(CB0, CB1)의 크기로, 즉

$$0 \leq i < L_0, 0 \leq j < L_1 \text{으로}$$

분자가 각각 1인 제곱과 곱의 합과 가산에 대한 처리량으로, 그리고 분모가 각각 1인 곱과 곱의 합의 처리량으로, 수학식 35의 (1)'의 처리량은 대략 다음과 같아서,

$$\text{분자: } L_0 \cdot L_1 \cdot (K \cdot (1+1) + 1)$$

$$\text{분모: } L_0 \cdot L_1 \cdot K \cdot (1+1)$$

$$\text{크기비교: } L_0 \cdot L_1$$

$L_0 \cdot L_1 \cdot (4K + 2)$ 의 합을 제공한다. 만약 $L_0=L_1=32$ 이고 $K=44$ 라면, 처리량은 182272의 차수이다.

따라서, 수학식 35의 (i)'의 처리가 모두 실행되지 않지만, 각 벡터($\mathbf{s}_{0i}, \mathbf{s}_{1i}$)의 P개가 미리 선택된다. 음의 이득 엔트리(entry)가 가정(또는 허용)되기 때문에, 수학식 35의 (1)'은 수학식 35의 (2)'의 분자의 값이 항상 양의 값되도록 탐색된다. 즉, 수학식 35의 (1)'은 $\mathbf{x}^t \mathbf{W}'^t \mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{1i})$ 의 극성을 포함하여 최대로 된다.

예비선택(pre-selection) 방법의 도식적 예로서, 다음과 같은 방법이 서술될 수도 있다.

(순서 1) 위 차수부터 계수하는 \mathbf{s}_{0i} 의 P_0 개를 선택해서, $\mathbf{x}^t \mathbf{W}'^t \mathbf{W}' \mathbf{s}_{0i}$ 을 최대화 하고,

(순서 2) 위 차수부터 계수하는 \mathbf{s}_{ii} 의 P_1 개를 선택해서, $\mathbf{x}^t \mathbf{W}'^t \mathbf{W}' \mathbf{s}_{ii}$ 을 최대로 하고,

(순서 3) \mathbf{s}_{0i} 의 P_0 개 \mathbf{s}_{ii} 의 P_1 개의 모든 조합에 대해 수학식 35의 (1)'의 식을 평가한다.

이것은, 수학식 35의 식 (1)'의 제곱근인 다음 식의 평가에서

[수학식 36]

$$\frac{\mathbf{x}^t \mathbf{W}'^t \mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{ij})}{\|\mathbf{W}' (\mathbf{s}_{0i} + \mathbf{s}_{ij})\|}$$

의 평가가

$\mathbf{s}_{0i} + \mathbf{s}_{ii}$ 의 가중형인 분모가 i 또는 j 에 관계없이 항상 일정하다는 가정 하에서 유효하다. 실제로, 수학식 36의 분모의 크기는 일정하지 않다. 이것을 고려한 예비선택 방법을 다음에 설명할 것이다.

여기서, 수학식 36의 분모가 일정하다는 가정 하에서 처리량을 감소시키는 효과를 설명한다. $L_0 \cdot K$ 의 처리량이 (순서 1)의 탐색 동안 요구되기 때문에,

$$\begin{aligned} & (L_0 - 1) + (L_0 - 2) + \dots + (L_0 - P_0) \\ & = P \cdot L_0 - P_0(1 + P_0)/2 \end{aligned}$$

의 처리량은 크기 비교를 위해 요구되는 동안, 처리량의 합은 $L_0(K + P_0) - P_0(1 + P_0)/2$ 이다. (순서 2)도 또한 유사한 처리량이 필요하다. 이것들을 함께 합할 때, 예비선택을 위한 처리량은 다음과 같다.

$$L_0(K + P_0) + L_1(K + P_1) - P_0(1 + P_0)/2 - P_1(1 + P_1)/2$$

순서 3의 최종 선택의 처리로 되돌아가면,

$$\text{분자: } P_0 \cdot P_1 \cdot (1 + K + 1)$$

$$\text{분모: } P_0 \cdot P_1 \cdot K \cdot (1 + 1)$$

$$\text{크기 비교: } P_0 \cdot P_1$$

수학식 35의 (1)'의 처리를 고려하면, $P_0 \cdot P_1(3K + 3)$ 의 총계를 준다.

예를 들어, 만약 $P_0 = P_1 = 6$, $L_0 = L_1 = 32$, 그리고 $K = 44$ 라면, 최종 선택에 대한 처리량과 예비선택에 대한 처리량은 각각 4860과 3158이어서, 8081의 차수의 총계를 준다.

만약 예비선택의 개수가 $P_0 = P_1 = 10$ 과 같이 10으로 증가되면, 최종 선택에 대한 처리량은 13500인 반면, 예비선택에 대한 처리량은 3346이 되어서, 16846의 차수의 총계를 준다.

만약 예비선택 벡터의 개수가 각 코드북에 대해 10으로 설정되면, 182272의 생략되지 않은 계산에 대한 처리량에 비교해 볼 때 처리량은

$$16846/182272$$

이며, 이전 양의 약 1/10이다.

반면에, 수학식 (35)의 식 (1)'의 분모의 크기는 일정하지 않지만 선택된 코드벡터에 따라서 변화된다. 이 형(norm)의 대략적인 크기를 고려한 예비선택 방법을 지금부터 설명한다.

수학식 35의 식 (1)'의 제곱근인 수학식 36의 최대값을 구하기 위해서는,

[수학식 37]

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}(\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|} \leq \frac{\mathbf{x}'\mathbf{W}'\mathbf{W}(\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}'(\mathbf{s}_{0i} + \mathbf{s}_{1j})\|}$$

이기 때문에, 수학식 37의 좌측을 최대로 하는 것으로 만족된다. 따라서, 이 좌측은

[수학식 38]

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}(\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|} = \frac{\mathbf{x}'\mathbf{W}'\mathbf{W}\mathbf{s}_{0i}}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|} + \frac{\mathbf{x}'\mathbf{W}'\mathbf{W}\mathbf{s}_{1j}}{\|\mathbf{W}'\mathbf{s}_{0i}\| + \|\mathbf{W}'\mathbf{s}_{1j}\|}$$

으로 확장되고, 다음에 제 1 및 제 2항이 최대로 된다.

수학식 38의 제 1항의 분자가 단지 \mathbf{s}_{0i} 의 함수이기 때문에, 제 1항은 \mathbf{s}_{0i} 에 대해 최대로 된다. 다른 한편으로는, 수학식 38의 제 2항이 단지 \mathbf{s}_{1j} 의 함수이기 때문에, 제 2항이 \mathbf{s}_{1j} 에 대해 최대로 된다. 즉, 다음과 같은 방법이 구체화되며,

[수학식 39]

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}\mathbf{s}_{0i}}{\|\mathbf{W}'\mathbf{s}_{0i}\|}$$

[수학식 40]

$$\frac{\mathbf{x}'\mathbf{W}'\mathbf{W}\mathbf{s}_{1j}}{\|\mathbf{W}'\mathbf{s}_{1j}\|}$$

다음의 순서를 포함한다.

(순서 1): 수학식 39를 최대로 하는 벡터들의 상위 순서에서부터 \mathbf{s}_{0i} 의 Q0개를 선택.

(순서 2): 수학식 40을 최대로 하는 벡터들의 상위 순서에서부터 \mathbf{s}_{1j} 의 Q1개를 선택,

(순서 3): \mathbf{s}_{0i} 의 선택된 Q0개와 \mathbf{s}_{1j} 의 선택된 Q1개의 모든 조합에 대해 수학식 35의 식 (1)'을 평가.

반면에, $W'=WH/\|\mathbf{x}\|$ 이며, W와 H 둘 다는 입력벡터(\mathbf{x})의 함수이며, W는 원래 입력벡터(\mathbf{x})의 함수이다.

그러므로, 수학식 39와 수학식 40의 분모를 계산하기 위해서, W는 원래 입력벡터(\mathbf{x})마다 계산되어야 한다. 하지만, 예비 선택을 위해 과도하게 처리량을 소비하는 것은 바람직하지 않다. 그러므로, 이러한 분모들은 W'의 실제의 또는 대표값을 이용하여 \mathbf{s}_{0i} 와 \mathbf{s}_{1j} 에 대해 미리 계산되고, \mathbf{s}_{0i} 와 \mathbf{s}_{1j} 의 값들에 따라서 표에 저장된다. 반면에, 실제의 탐색처리에서 나누기는 처리시의 부하를 의미하기 때문에, 수학식 41과 수학식 42의 값은

[수학식 41]

$$\frac{1}{\|\mathbf{W}^*\mathbf{s}_{0i}\|} \quad 0 \leq i < L_0$$

[수학식 42]

$$\frac{1}{\|\mathbf{W}^*\mathbf{s}_{1j}\|} \quad 0 \leq j < L_1$$

이 저장된다. 상기 수학식에서, \mathbf{W}^* 은 다음의 수학식 43으로 주어진다.

[수학식 43]

$$W = \frac{1}{N} \sum_{k=1}^N W_k$$

여기서, W_k '은 V/UV 가 다음과 같이 유성으로 구해졌던 프레임의 W '이다.

[수학식 44]

$$W = \frac{WH}{\|x\|}$$

도 10은 W^* 가 다음의 수학식 45에 의해 기술되는 경우에 $W[0] \sim W[43]$ 의 각각의 구체적인 예를 나타낸다.

[수학식 45]

$$W^* = \begin{bmatrix} W[0] & & & & 0 \\ & W[1] & & & \\ & & W[2] & & \\ & & & \ddots & \\ 0 & & & & W[43] \end{bmatrix}$$

수학식 39와 수학식 40의 분자에 대해서, W '는 입력벡터(\underline{x})마다 구해지고 이용된다. 그 이유는 어떠한 비율로도 s_{0i} 및 s_{1i} 와 \underline{x} 와의 내적이 계산될 필요가 있기 때문에, $\underline{x}'W'W$ '이 한 번 계산되면 처리량이 단지 약간 증가하기 때문이다.

예비선택방법에 요구되는 처리량의 대략적인 판단에 대해서, $L_0(K+1)$ 의 처리량은 순서 1의 탐색에 대해 요구되는 반면,

$$Q_0 \cdot L_0 - Q_0(1+Q_0)/2$$

의 처리량은 크기 비교에 대해서 요구된다. 상기 순서 2는 또한 유사한 처리에 필요하다. 이러한 처리량을 함께 합할 때, 예비선택에 대한 처리량은 다음과 같다.

$$L_0(K+Q_0+1) + L_1(K+Q_1+1) - Q_0(1+Q_0)/2 - Q_1(1+Q_1)/2$$

순서 3의 최종 선택의 처리에 대해서,

$$\text{분자: } Q_0 \cdot Q_1 \cdot (1+K+1)$$

$$\text{분모: } Q_0 \cdot Q_1 \cdot K \cdot (1+1)$$

$$\text{크기비교: } Q_0 \cdot Q_1$$

$$\text{총 } Q_0 \cdot Q_1(3K+3)\text{이다.}$$

예를 들어, 만약 $Q_0=Q_1=6$, $L_0=L_1=32$, 그리고 $K=44$ 라면, 최종 선택의 처리량과 예비선택의 처리량은 각각 4860과 3222이며, 총 8082(크기의 8차)이다. 만약 예비선택에 대한 벡터의 수가 $Q_0=Q_1=10$ 과 같이 10으로 증가된다면, 최종 선택의 처리량과 예비선택의 처리량은 각각 13500과 3410이 되며, 총 16910(크기의 8차)이다.

이 계산된 결과들은 $P_0=P_1=6$ 에 대해 대략 8081 또는 정규화가 있을 때 $P_0=P_1=10$ 에 대해 대략 16846의 처리량과 같은 크기의 차수와 같다. 예를 들어, 만약 각 코드북에 대한 벡터의 수가 10으로 설정되면, 처리량은

$$19610/182272$$

로 감소하며, 182272는 빠뜨림 없는 처리량이다. 따라서 처리량은 원래의 처리량의 1/10보다 적지 않게 감소된다.

상기 설명한 예비선택을 기준으로 행하지 않고 분석 및 합성된 음성을 이용하여 예비선택이 이루어지는 경우에 SNR(S/N비)와 20msec 세그먼트에 대하여 세그먼트 SNR에 대하여 구체적인 예를 나타내면, P0=P1=6이고 정규화하지 않고서 예비선택에 대하여 동일수의 벡터로 SNR이 14.8dB이고, 세그먼트SNR이 17.5dB에 비하여 정규화하고 가중이 없는 경우에 SNR이 16.8dB이고, 세그먼트 SNR이 18.7dB인 반면, 정규화하고 가중이 있는 경우에 SNR이 17.8dB이고, 세그먼트 SNR이 19.6dB이다. 즉, 정규화가 없을 때의 동작 대신에 가중과 정규화가 있을 때의 동작을 이하에 2대 3으로 SNR과 세그먼트SNR이 향상된다.

수학식 31과 수학식 32의 조건(중심조건(Centroid Condition))과, 수학식 35의 조건을 이용하면 소위 일반화 로이드 알고리즘(Generalized Lloyd Algorithm : GLA)을 사용하여 코드북(CB0, CB1, CBg)을 동시에 트레이닝(train)시킬 수 있다.

본 실시예에서 W'로서 입력(x)의 형으로 분할한 W'를 W'로 사용하고 있다. 즉, 수학식 31과 수학식 32와 수학식 35에서 W'에 W'/||x||를 대입하고 있다.

선택적으로, 벡터양자화기(116)에 의해 벡터양자화 할 때 청각가중에 이용되는 가중(W')은 상기 수학식 26으로 정의된다. 그렇지만, 과거의 W'를 고려하여 현재의 W'를 구함으로써 일시적인 마스킹(masking)을 고려한 W'를 구할 수 있다.

상기 수학식 26에서 wh(1), wh(2), ..., wh(L)의 값은 시간(n)에서 즉 n번째 프레임에서 산출된 것으로 각각 whn(1), whn(2), ..., whn(L)로 나타낸다.

시간(n)에서 과거의 값을 고려한 가중이 An(i), 1≤i≤L로 정의하면,

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

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

여기에서, λ는 예를 들면 λ=0.2로 설정된다. 이와 같이 하여 구한 An(i), 1≤i≤L에 대하여 An(i)를 대각선 요소로서 가지는 매트릭스를 상기 가중으로서 이용해도 좋다.

이와 같이 가중 벡터양자화에 의해 얻어진 형상인덱스(s_{0i}, s_{1i})는 출력단자(520, 522)에서 각각 출력되고, 이득인덱스(g₁)는 출력단자(521)에서 출력된다. 또한, 양자화값(x₀')은 출력단자(504)에서 출력되면서 가산기(505)에 보내진다.

가산기(505)는 스펙트럼 엔벨로프 벡터(x)로부터 양자화값(x₀')을 감산하고, 양자화 오차벡터(y)가 생성된다. 특히, 이 양자화 오차벡터(y)는 벡터양자화부(511)에 보내지고, 차원분할되고, 벡터양자화기(511₁ ~ 511₈)에서 가중의 벡터양자화로 양자화 된다. 제 2양자화부(510)는 제 1벡터양자화부(500)보다 큰 비트수를 사용한다. 따라서, 코드북의 기억용량과 코드북탐색의 처리량(복합성)은 현저하게 증가한다. 따라서, 제 1벡터양자화부(500)와 같은 44차원으로 벡터양자화를 실행하는 것이 불가능하게 된다. 그러므로, 제 2벡터양자화부(510)에서 벡터양자화부(511)는 여러 개의 벡터양자화기로 구성되고 양자화된 입력값은 다수의 벡터양자화를 실행하기 위하여 다수의 저차원벡터로 차원분할된다.

벡터양자화기(511₁ ~ 511₈)에서 사용되는 양자화값(y₀ ~ y₁), 차원수, 비트수의 관계를 다음 표 2에서 나타낸다.

벡터양자화기(511₁ ~ 511₈)에서 출력된 인덱스값(Id_{vq0} ~ Id_{vq7})은 출력단자(523₁ ~ 523₈)에서 출력된다. 이들 인덱스 데이터의 합계는 72이다.

벡터양자화기(511₁ ~ 511₈)의 양자화된 출력값(y₀'~y₁')을 차원방향으로 연결하여 얻어진 값이 y'이면, 양자화된 값(y')과 값(x₀')이 가산기(513)에 의해 합해져서 양자화된 값(x₁')을 산출한다. 그러므로, 양자화된 출력값(x₁')은

$$x_1' = x_0' + y'$$

$$= x - y + y'$$

으로 표시된다.

즉, 최종 양자화 오차벡터는 $\underline{y}' - \underline{y}$ 이다.

제 2양자화기(510)로부터 양자화된 값(\underline{x}_1')이 복호화 되면, 음성신호 복호화 장치는 제 1양자화기(500)로부터 양자화된 값(\underline{x}_1')을 필요로 하지 않는다. 그러므로, 제 1양자화기(500)와 제 2양자화기(510)로부터 인덱스테이터를 필요로 한다.

벡터양자화부(511)에서 학습법과 코드북탐색을 이하 설명한다.

학습법에 대하여 양자화 오차벡터(\underline{y})는 표 2에 나타난 것같이 8개의 저차원벡터($\underline{y}_0 \sim \underline{y}_7$)로 분할된다. 가중(W')이 대각선 요소로서 44포인트 부샘플값을 가지는 매트릭스라면,

[수학식 46]

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

가중(W')은 다음의 8개의 매트릭스로서 분할된다.

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

$$W_1' = \begin{bmatrix} wh(5) & & 0 \\ & \ddots & \\ 0 & & wh(8) \end{bmatrix}$$

$$W_1' = \begin{bmatrix} wh(9) & & 0 \\ & \ddots & \\ 0 & & wh(12) \end{bmatrix}$$

$$W_1' = \begin{bmatrix} wh(13) & & 0 \\ & \ddots & \\ 0 & & wh(16) \end{bmatrix}$$

$$W_1' = \begin{bmatrix} wh(17) & & 0 \\ & \ddots & \\ 0 & & wh(20) \end{bmatrix}$$

$$W_1' = \begin{bmatrix} wh(21) & & 0 \\ & \ddots & \\ 0 & & wh(28) \end{bmatrix}$$

$$W_1' = \begin{bmatrix} wh(29) & & 0 \\ & \ddots & \\ 0 & & wh(36) \end{bmatrix}$$

$$W_1' = \begin{bmatrix} wh(37) & & 0 \\ & \ddots & \\ 0 & & wh(44) \end{bmatrix}$$

이렇게 저차원으로 분할된 \underline{y} 와 W' 는 각각 $Y_i, W_i', 1 \leq i \leq 8$ 로 한다.

왜곡량(E)은 다음과 같이 정의된다.

[수학식 47]

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

코드북벡터(s)는 y_i 의 양자화의 결과이다. 왜곡량(E)을 최소화하는 코드북의 이러한 코드벡터가 탐색된다.

코드북학습에서, 또한 가중치 일반화 로이드 알고리즘(GLA)을 사용하여 실행된다. 학습의 최적의 중심조건이 먼저 설명된다. 코드벡터(s)를 최적 양자화결과로서 선택한 M 입력벡터(y)가 있으면, 트레이닝데이터는 (y_k)이고, 왜곡의 기대값(J)은 모든 프레임(k)에 대하여 가중치에 왜곡의 중앙을 최소화하는 수학적식 48에 의해 주어진다.

[수학적식 48]

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

을 풀면

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

을 얻는다.

양측을 전환한 값을 취하면

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

얻는다.

그러므로,

[수학적식 49]

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

이다.

상기 수학적식 49에서, s 는 최적의 대표벡터이고 최적 중심조건을 나타낸다.

최적 부호화조건에 대하여, $\| W_i'(y_i - s) \|^2$ 의 값을 최소화하는 s 를 탐색하기에 충분하다. 탐색시의 W_i' 는 반드시 학습시의 W_i' 와 동일할 필요는 없고, 가중치의 매트릭스로 할 수 있다.

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

음성신호 부호화기 내의 벡터양자화부(116)를 2단의 벡터양자화부로 구성함으로써 출력하는 인덱스의 비트수를 가변으로 할 수 있다.

그런데, 스펙트럼 엔벌로프평가부(148)에 있어서 얻어지는 조파들의 데이터수는 피치에 따라서 변화하고 예를 들면 유효 주파수대역이 3400kHz이면, 데이터수는 8개로부터 63개 정도까지의 범위가 된다. 이들 데이터로 이루어진 서로 블록화된 벡터(\mathbf{v})는 가변차원벡터이다. 상술의 구체적인 예에서는 벡터양자화의 전에 일정한 데이터수, 44차원의 입력벡터(\mathbf{x})와 같이 차원변환하고 있다. 이 가변/고정차원변환은 상술한 데이터수 변환을 의미하고, 특히 오버샘플링 및 선형 보간을 이용하여 실현할 수 있다.

이와 같은 고정차원으로 변환한 벡터(\mathbf{x})에 대하여 오차처리를 행하여 오차를 최소화하는 코드북탐색을 행하면, 반드시 원래의 가변차원벡터(\mathbf{v})에 대한 오차를 최소화하는 코드벡터가 선택되는 것은 아니다.

그래서, 본 실시예에서는 고정차원의 코드벡터를 선택하는데 있어서 복수의 코드벡터를 일시적으로 선택하고, 이들의 일시 선택된 복수의 코드벡터로부터 최종적인 최적 가변차원 코드벡터를 선택한다. 한편, 고정차원 일시선택을 행하지 않고 가변차원에서의 선택처리가 실시될 수 있다.

도 12는 원래의 가변차원 최적 벡터선택을 위한 구성의 일 예를 나타내고 있다.

입력단자(541)에는 스펙트럼 엔벌로프평가부(148)에 의해 얻어지는 스펙트럼 엔벌로프의 가변개수의 데이터, 즉 가변차원벡터(\mathbf{v})가 입력되고 있다. 가변차원의 입력벡터(\mathbf{v})는 상술한 데이터수 변환회로로서 가변/고정차원 변환회로(542)에 의해 (44개의 데이터로 구성되는 고정차원(44차원)과 같은) 고정차원벡터(\mathbf{x})로 변환되고, 단자(501)에 보내지고 있다. 고정차원 입력벡터(\mathbf{x})와 고정차원의 코드북(530)에서 읽어낸 고정차원 코드벡터는 고정차원 선택회로(535)에 보내지고, 이들 사이의 가중치의 오차 혹은 왜곡을 최소로 감소시키는 코드벡터를 코드북(530)에서 선택하는 선택동작 혹은 코드북탐색이 행해진다.

도 12의 실시예에 있어서, 고정차원의 코드북(530)에서 얻어진 고정 2차원의 코드벡터가 원래의 차원과 같은 가변차원의 고정/가변차원 변환회로(544)에 의해 변환된다. 변환된 차원코드벡터가 가변차원의 변환회로(544)에 보내지고, 코드벡터와 상기 입력벡터(\mathbf{v})와의 사이의 가중치 왜곡의 계산을 행하고, 그 왜곡을 최소로 줄이는 코드벡터를 코드북(530)에서 선택하기 위하여 선택처리 혹은 코드북탐색을 행하고 있다.

즉, 고정차원의 선택회로(535)는 일시선택으로서 가중왜곡을 최소화하는 후보의 코드벡터로서 여러 가지 코드벡터를 선택하고 있고, 이들 후보 코드벡터에 대하여 가변차원의 선택회로(545)에서 가중치 왜곡계산을 행하여, 왜곡을 최소로 하는 코드벡터를 최종적으로 선택한다.

일시선택 및 최종선택을 이용하는 벡터양자화의 적용범위를 간단히 설명한다. 벡터양자화는 조파부호화, LPC나머지의 조파부호화, 본건 출원인이 일본 특허 공개 제4-91422에 설명한 MBE(다중대역여기)부호화, LPC나머지의 MBE부호화에 있어서의 조파의 스펙트럼에서 특정 성분에 대하여 차원변환을 이용하여 가변차원의 조파의 가중벡터양자화에 적용할 수 있을 뿐 아니라, 고정차원 코드북을 이용하여 가변차원 입력벡터의 벡터양자화에도 적용할 수 있다.

일시선택에 대하여는 다단의 양자화기 구성의 일부를 선택한다든지 형상코드북과 이득코드북으로 이루어지는 코드북의 경우에 형상코드북만을 일시 선택하여 탐색하고 가변차원의 왜곡계산에 의해 이득을 결정하는 것이 가능하다. 또한, 일시선택에 대하여 상술한 예비선택이 사용될 수 있다. 특히, 고정차원의 벡터(\mathbf{x})와 이 코드북에 저장된 모든 코드벡터사이의 유사도를 근사계산(가중왜곡의 근사계산)에 의해 구하고 유사도가 높은 복수의 코드벡터를 선택한다. 이 경우, 고정차원에서의 일시선택을 상기 예비선택으로 하고, 예비선택된 후보의 코드벡터에 대하여 가변차원에서의 가중왜곡을 최소화하는 최종 선택을 행하여도 좋고, 또 일시선택의 공정에서 상기 예비선택뿐만 아니라 고정밀도의 왜곡계산에 의한 교환을 다시 행한 후에 최종 선택에 돌아가도록 하여도 좋다.

일시선택 및 최종선택을 이용한 벡터양자화의 구체적인 예에 대하여 도면을 참조하면서 상세히 설명한다.

도 12에 있어서는 코드북(530)은 형상코드북(531)과 이득코드북(532)으로 이루어진다. 형상코드북(531)은 2개의 코드북(CB0, CB1)으로 구성된다. 이 형상코드북(CB0, CB1)의 출력 코드벡터는 $\mathbf{s}_0, \mathbf{s}_1$ 로서 표시되는 반면, 이득코드북(532)에 의해 결정되는 이득회로(533)의 이득이 g 로 표시된다. 입력단자(541)로부터 가변차원 입력벡터(\mathbf{v})는 가변/고정차원 변환

회로(542)에 의해 차원변환(여기에서 D1로 표시)처리되고, 코드북(530)에서 읽어낸 고정차원 코드벡터로부터 벡터(\underline{x})의 차이가 구해지고 가중회로(537)에 의해 가중된 오차 최소화회로(538)에 공급되도록 선택회로(535)의 감산기(536)로 고정차원의 벡터(\underline{x})로서 단자(501)를 통하여 공급된다. 이 가중회로(537)는 가중(W')을 사용한다. 코드북(530)으로부터 고정차원 코드벡터는 가변/고정차원 변환회로(542)에 의해 차원변환(여기에서 D2로 표시)처리되고, 가변차원 입력벡터(\underline{y})의 차이가 구해지고 가중회로(537)에 의해 가중된 가변차원 선택회로(545)의 선택기(546)로 공급되어 오차 최소화회로(548)에 공급되도록 한다. 이 가중회로(537)는 가중(W_v)을 사용한다.

오차 최소화회로(538, 548)의 오차는 상기 왜곡 혹은 왜곡량을 의미한다. 오차 또는 왜곡이 작게 된다는 것은 유사도 혹은 상관성을 증가시키는 것과 동일하다.

수학식 50에 의해 표시되는 왜곡량(E₁)을 최소화하는 $\underline{s}_0, \underline{s}_1, g$ 에 대하여 고정차원 일시선택 탐색을 실행하고, 수학식 27을 참조하여 설명한다.

[수학식 50]

$$(E_1) = \|W'(\underline{x} - g(\underline{s}_0 + \underline{s}_1))\|^2$$

가중회로(537)에서 가중치(W)는

[수학식 51]

$$W' = WH / \|\underline{x}\|$$

에 의해 주어진다.

여기에서 H는 LPC합성필터의 주파수응답특성을 대각선 요소로 가지는 매트릭스를 나타내고, 또한 W는 청각가중필터의 주파수응답특성을 대각선 요소로 가지는 매트릭스를 나타낸다.

먼저, 수학식 50의 왜곡량(E₁)을 최소화하는 $\underline{s}_0, \underline{s}_1, g$ 가 탐색된다. L세트의 $\underline{s}_0, \underline{s}_1, g$ 는 왜곡량(E₁)을 작게 하는 순서로 상위 순서 측에서 시작하여 고정차원에서 일시선택을 거쳐서 취해진다. 그러면, 최종선택이 최적 코드벡터로서 최소화하는 L세트의 $\underline{s}_0, \underline{s}_1, g$ 에서 실행된다.

[수학식 52]

$$(E_2) = \|W_v(\underline{y} - D_2g(\underline{s}_0 + \underline{s}_1))\|^2$$

수학식 50에 대한 탐색과 학습이 수학식 27과 다음 식을 참조하여 설명된다.

수학식 52에 의거하여 코드북학습에 대한 중심조건을 이하 설명한다.

코드북(530)에 있어서 형상코드북(531)의 하나로서 코드북(CB0)에 대하여 모든 프레임(k)에 대한 왜곡의 기대값은 코드 벡터(\underline{s}_0)를 선택하는 것으로부터 최소화된다. 이와 같은 프레임이 M개이면, 최소화하는 것을 만족시킨다.

[수학식 53]

$$J = \frac{1}{M} \sum_{k=1}^M \|W_{vk}(\underline{y}_k - g_k D_{2k}(\underline{s}_{0c} + \underline{s}_{1k}))\|^2$$

수학식 53을 최소화하기 위하여 수학식 54가 해석되어

[수학식 54]

$$\frac{\partial J}{\partial \underline{s}_{0c}} = \frac{1}{M} \sum_{k=1}^M \{-2g_k D_{2k}^T W_{vk}^T W_{vk} \underline{y}_k + 2g_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0c} + 2g_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{1k}\} = 0$$

수학식 55가 주어진다.

[수학식 55]

$$s_{0c} = \left\{ \sum_{k=1}^M g_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \right\}^{-1} \times \left\{ \sum_{k=1}^M g_k D_{2k}^T W_{vk}^T W_{vk} (v_k - g_k D_{2k} s_{1k}) \right\}$$

이 수학식 55에서 0^{-1} 은 역매트릭스를 표시하고 W_{vk}^T 는 W_{vk} 의 전치매트릭스를 표시한다. 이 수학식 55는 형상벡터(s_0)의 최적 중심조건을 나타낸다.

코드북(530)에서 다른 형상코드북(531)의 코드북(CB1)에 대하여 코드벡터(s_1)의 선택이 상기와 동일한 방법으로 실행되므로 간략하게 나타내기 위해 설명을 생략한다.

그러면, 코드북(530)에서 이득코드북(532)으로부터 이득(g)의 중심조건을 고찰한다.

코드워드(g_c)를 선택하기 위하여 k번째 프레임에 대한 왜곡의 기대치가 수학식 56에 의해 주어진다.

[수학식 56]

$$J_g = \frac{1}{M} \sum_{k=1}^M \| W_{vk} (v_k - g_c D_{2k} (s_{0k} + s_{1k})) \|^2$$

수학식 56을 최소화하기 위하여 다음 수학식 57이 해석되고,

[수학식 57]

$$\begin{aligned} \frac{\partial J_g}{\partial g_c} &= \frac{1}{M} \sum_{k=1}^M \{ -2v_k^T W_{vk}^T W_{vk} D_{2k} (s_{0k} + s_{1k}) + 2g_c (D_{2k} (s_{0k} + s_{1k}))^T W_{vk}^T W_{vk} D_{2k} (s_{0k} + s_{1k}) \} \\ &= 0 \end{aligned}$$

[수학식 58]

$$g_c = \frac{\sum_{k=1}^M v_k^T W_{vk}^T W_{vk} D_{2k} (s_{0k} + s_{1k})}{\sum_{k=1}^M (D_{2k} (s_{0k} + s_{1k}))^T W_{vk}^T W_{vk} D_{2k} (s_{0k} + s_{1k})}$$

이 주어진다.

이 수학식 58은 이득에 대한 중심조건을 나타낸다.

다음에, 수학식 52에 의거하여 가장 가까운 인접조건이 고려된다.

수학식 3에 의해 탐색된 s_0, s_1, g 세트의 수가 고정차원의 일시선택에 의해 제한되기 때문에, 수학식 52는 최적 코드벡터로서 왜곡(E)을 최소화하는 s_0, s_1, g 의 세트를 선택하기 위하여 s_0, s_1, g 세트(L)에 대하여 직접 계산된다.

일시선택에 대하여 L이 매우 크거나 s_0, s_1, g 가 일시선택을 하지 않고 직접 가변차원에서 선택되면, 유효하다고 받아들인 형상과 이득에 대한 순차 탐색방법을 설명한다.

수학식 52의 s_0, s_1, g 에 인덱스(i, j, l)가 가해지고 이러한 형태로 수학식 3b가 다시 쓰여지면,

[수학식 59]

$$(E2) = \|W_v(\mathbf{y} - D_2\mathbf{g}_1(\mathbf{s}_{0i} + \mathbf{s}_{1j}))\|^2$$

을 얻는다.

수학식 59를 최소화하는 $g, \mathbf{s}_0, \mathbf{s}_1$ 가 라운드로빈 방식으로 탐색되지만, $0 \leq l < 32, 0 \leq i < 32, 0 \leq j < 32$ 이면, 상기 수학식 59는 $32^3 = 32768$ 패턴에 대하여 계산될 필요가 있고, 그래서 팽대한 처리량으로 된다. 형상과 이득을 순차적으로 탐색하는 방법을 이하 설명한다.

형상코드벡터($\mathbf{s}_{0i}, \mathbf{s}_{1j}$)를 결정한 뒤에 이득(g_1)이 결정된다. $\mathbf{s}_{0i} + \mathbf{s}_{1j} = \mathbf{s}_m$ 으로 설정하면, 수학식 59는

[수학식 60]

$$(E_2) = \|W_v(\mathbf{y} - D_2\mathbf{g}_1\mathbf{s}_m)\|^2$$

로 표시될 수 있다.

$\mathbf{y}_w = W_v\mathbf{y}, \mathbf{s}_w = W_v D_2 \mathbf{s}_m$ 으로 설정하면, 수학식 11은

[수학식 61]

$$E_2 = \| \mathbf{y}_w - g_1 \mathbf{s}_w \|^2 = \| \mathbf{y}_w \|^2 + \| \mathbf{s}_w \|^2 \left(g_1 - \frac{\mathbf{y}_w^T \mathbf{s}_w}{\| \mathbf{s}_w \|^2} \right)^2 - \left(\frac{\mathbf{y}_w^T \mathbf{s}_w}{\| \mathbf{s}_w \|^2} \right)^2$$

이 된다.

그러므로, (g_1)가 충분한 정밀도가 되면,

[수학식 62]

$$\frac{(\mathbf{y}_w^T \mathbf{s}_w)^2}{\| \mathbf{s}_w \|^2}$$

을 최대화하는 \mathbf{s}_w 와

[수학식 63]

$$\frac{\mathbf{y}_w^T \mathbf{s}_w}{\| \mathbf{s}_w \|^2}$$

에 가장 근접한 g_1 이 탐색된다.

원래의 변수를 대입하여 수학식 62와 수학식 63을 다시 쓰면, 다음 수학식 64와 수학식 65를 얻는다.

[수학식 64]

$$\frac{(\mathbf{y}^T W_v^T W_v D_2 (\mathbf{s}_{0i} + \mathbf{s}_{1j}))^2}{\| W_v D_2 (\mathbf{s}_{0i} + \mathbf{s}_{1j}) \|^2}$$

을 최대화하는 $\mathbf{s}_{0i}, \mathbf{s}_{1j}$ 의 세트와

[수학식 65]

$$\frac{\mathbf{v}^T \mathbf{W}_v^T \mathbf{W}_v \mathbf{D}_2 (\mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|\mathbf{W}_v \mathbf{D}_2 (\mathbf{s}_{0i} + \mathbf{s}_{1j})\|^2}$$

에 가장 근접한 g_1 이 탐색된다.

수학식 55와 수학식 58의 형상과 이득에 대한 중심조건과 수학식 64와 수학식 65의 최적부호화조건(가장 근접한 인접조건)을 사용하여, 코드북(CB0, CB1, CBg)이 일반화 로이드 알고리즘에 의해 동시에 학습될 수 있다.

상술한 바와 같이 수학식 27이하 특히, 수학식 31과 수학식 32와 수학식 35를 이용하는 방법과 비교하여, 상기 수학식 55와 수학식 58과 수학식 64와 수학식 65를 이용하는 학습법은 원래의 입력벡터(\mathbf{v})를 가변차원벡터로 변환한 후의 왜곡을 최소화하는데 우수하다.

그러나, 수학식 55와 수학식 58에 의한 처리가 특히 수학식 55에서는 복잡하기 때문에, 수학식 27을 최적화 하는 데에서 들어가는 중심조건이 수학식 64와 수학식 65의 가장 인접한 인접조건만을 사용하는데 즉, (b1)에서 사용될 수 있다.

코드북의 학습동안 상기 수학식 27 이하의 설명에 서술한 방법으로 행하고, 탐색동안에만 상기 수학식 64와 수학식 65를 이용하는 방법도 권해진다. 고정차원에서의 일시선택을 상기 수학식 27 이하의 설명에 서술한 방법으로 행하고, 선택된 복수개(L개)의 조에 대하여만 수학식 52를 직접 평가하여 탐색을 행하도록 하여도 좋다.

어떠한 경우에도, 상기 수학식 52의 왜곡평가에 의한 탐색을 일시 선택 후에, 혹은 라운드로빈방식으로 사용함으로써 최종적으로 보다 왜곡이 적은 코드벡터탐색 혹은 학습을 행하는 것이 가능하다.

원래의 입력벡터(\mathbf{v})와 동일한 가변차원에서 왜곡계산을 행하는 것이 바람직한 이유에 대하여 간단히 서술한다.

고정차원에서의 왜곡의 최소화가 가변차원에서와 일치하면, 가변차원에서의 왜곡의 최소화는 불필요하다. 그러나, 고정/가변차원 변환회로(544)에서의 차원변환(D2)이 직교매트릭스들은 아니기 때문에, 이들의 왜곡최소화는 일치하지 않는다. 그래서, 고정차원에서 왜곡을 최소화하여도 이러한 최소화가 반드시 가변차원에서 최적으로 최소화하는 것은 아니고, 결과적으로 가변차원의 벡터의 벡터가 최적화 하면 가변차원에서 왜곡을 최적화 하는 것이 필요하게 된다.

도 13은 코드북을 형상코드북과 이득코드북으로 나눌 수 있을 때의 이득이 가변차원에서의 이득으로 하고, 가변차원에서 왜곡이 최적화 되는 예를 나타내고 있다.

특히, 형상코드북(531)에서 읽어낸 고정차원의 코드벡터를 고정/가변차원 변환회로(544)에 보내져서 가변차원의 벡터로 변환한 후, 이득제어회로(533)에 보내진다. 선택회로(545)는 이득회로(533)에서 가변차원의 코드벡터와 입력벡터(\mathbf{v})에 의거하여, 고정/가변차원 변환된 코드벡터에 대한 이득회로(533)에서의 최적이득을 선택하기에 충분하다. 혹은, 이득회로(533)로의 입력벡터와 입력벡터(\mathbf{v})와의 내적에 의거하여 최적이득을 선택할 수 있다. 다른 구성 및 동작은 상기 도 12의 예와 동일하다.

형상코드북(531)에 대하여는 선택회로(535)에 있어서의 고정차원에서 선택하는 동안 유일한 코드벡터가 선택될 수 있고, 가변차원에서의 선택은 이득만으로 구성될 수 있다.

고정/가변차원 변환회로(544)에서 변환한 코드벡터에 이득을 곱하여서, 도 10에 나타난 것과 같은 이득을 곱한 코드벡터를 고정/가변차원 변환하는 것을 고려한 고정/가변차원변환에 의한 영향으로 최적의 이득을 선택할 수 있다.

고정차원에서의 일시선택과 가변차원에서의 최종선택을 결합하여서 벡터양자화의 보다 구체적인 예를 이하 설명한다.

아래의 구체적인 예에서, 제 1코드북에서 읽어낸 고정차원의 제 1코드벡터는 입력벡터의 가변차원으로 변환하고, 제 2코드북에서 읽어낸 고정차원의 제 2코드벡터는 상기 서술한 것같이 고정/가변차원변환에 의해 처리된 가변차원의 제 1코드벡터에 합해진다.

가산의 결과, 최종적인 합계코드벡터를 형상과 입력벡터에서 오차를 최소화하는 최적코드벡터가 적어도 제 2코드북에서 선택된다.

도 14의 예에서, 제 1코드북(CB0)에서 읽어낸 고정차원의 제 1코드벡터(\underline{s}_0)는 고정/가변차원 변환회로(544)에 보내져서 단자(541)에서 입력벡터(\underline{y})에서와 동일한 가변차원으로 변환된다. 제 2코드북(CB1)에서 읽어낸 고정차원의 제 2코드벡터는 가산기(549)로 보내져서 고정/가변차원 변환회로(544)로부터 가변차원의 코드벡터에 합해지게 된다. 가산기(549)의 결과로서 얻어진 코드벡터합계는 선택회로(545)로 보내지고, 가산기(549)로부터 합계벡터 또는 입력벡터(\underline{y})로부터 오차를 최소화하는 최적의 코드벡터가 선택된다. 제 2코드북(CB1)의 코드벡터는 입력벡터의 조파의 저부로부터 코드북(CB1)의 범위에 적용된다. 이득(g)의 이득회로(533)는 제 1코드북(CB1)과 고정/가변차원 변환회로(544) 사이에만 설치된다. 다른 구조는 도 12와 유사하기 때문에, 간단히 하기 위해서 동일 부호와 대응하는 설명을 생략한다.

그래서, 코드북(CB1)에서의 고정차원에 남아 있는 코드벡터와 코드북(CB0)으로부터 읽어내어 가변차원으로 변환된 코드벡터를 가산함으로써 서로 합해져서 코드북(CB1)으로부터 고정차원의 코드벡터에 의해 고정/가변 차원변환에 의해 생성된 왜곡을 감소한다.

도 14의 선택회로(545)에 의해 계산된 왜곡(E_3)은

[수학식 66]

$$E_3 = \| \mathbf{W}_v(\underline{y} - (D_2 \mathbf{g} \underline{s}_0 + \underline{s}_1)) \|^2$$

에 의해 주어진다.

도 15의 예에서, 이득회로(533)는 가산기(549)의 출력 측에 배열된다. 그래서, 코드북(CB0)에서 읽어내어 고정/가변차원 변환회로(544)에 의해 변환된 코드벡터와 제 2코드북(CB1)으로부터 읽어낸 코드벡터에 이득(g)이 곱해진다. CB0에서의 코드벡터와 곱해지는 이득이 보정분(양자화오차의 양자화)에 대한 코드북(CB1)에서의 코드벡터와 곱해지는 이득과 매우 유사하게 나타나므로 공통이득이 사용된다. 도 15의 선택회로(545)에 의해 계산된 왜곡(E_4)은

[수학식 67]

$$E_4 = \| \mathbf{W}_v(\underline{y} - \mathbf{g}(D_2 \mathbf{g} \underline{s}_0 + \underline{s}_1)) \|^2$$

로 주어진다.

이 예의 다른 구성은 도 14의 예에서와 동일하므로, 간략히 하기 위해 설명은 생략한다.

도 16의 예에서, 이득(g)을 가지는 이득회로(535A)가 도 14의 예에서 제 1코드북(CB0)의 출력 측에 설치될 뿐만 아니라, 이득(g)을 가지는 이득회로(533B)가 제 2코드북(CB1)의 출력 측에 설치된다. 도 16의 선택회로(545)에 의해 계산된 왜곡은 수학식 67에 나타낸 왜곡(E_4)과 동일하다. 도 16의 예의 다른 구성은 도 14의 예에서와 동일하므로, 간략히 하기 위해서 대응부분의 설명은 생략한다.

도 17은 도 14의 제 1코드북이 2개의 형상코드북(CB0, CB1)으로 구성되는 예를 나타낸다. 이들 형상코드북으로부터의 코드벡터($\underline{s}_0, \underline{s}_1$)는 서로 합해져서 그 결과의 합계는 고정/가변차원 변환회로(544)에 보내지기 전에 이득회로(533)에 의해 이득(g)이 곱해진다. 고정/가변차원 변환회로(544)로부터의 가변차원 코드벡터와 제 2코드북(CB2)에서의 코드벡터(\underline{s}_2)는 선택회로(545)에 보내지기 전에 서로 가산기(549)에 의해 합쳐진다.

도 17의 선택회로(545)에 의해 계산된 왜곡(E_5)은

[수학식 68]

$$E_5 = \| \mathbf{W}_v(\underline{y} - (\mathbf{g}D_2(\underline{s}_0 + \underline{s}_1) + \underline{s}_2)) \|^2$$

로 주어진다.

도 16의 예의 다른 구성은 도 14의 예에서와 동일하므로, 간단히 하기 위해서 대응하는 설명은 생략한다.

일 예로서, 제 1탐색방법은

[수학식 69]

$$E_4' = \|W'(\mathbf{x} - \mathbf{g}_1 \mathbf{s}_{0i})\|^2$$

을 최소화하는 탐색($\mathbf{s}_{0i}, \mathbf{g}_1$)을 포함한다. 그리고, 탐색(\mathbf{s}_{0i})은

[수학식 70]

$$E_4 = \|W_v(\mathbf{y} - \mathbf{g}_1(D_2 \mathbf{s}_{0i} + \mathbf{s}_{1j}))\|^2$$

을 최소화시킨다.

다른 일 예로서,

[수학식 71]

$$\frac{(\mathbf{s}_{0i}^T W'^T W' \mathbf{x})^2}{\|W' \mathbf{s}_{0i}\|^2}$$

을 최대화시키는 이러한 \mathbf{s}_{0i} 가 탐색되고,

[수학식 72]

$$\frac{(\mathbf{y}^T W_v^T (D_2 \mathbf{s}_{0i} + \mathbf{s}_{1j}))^2}{\|W_v(D_2 \mathbf{s}_{0i} + \mathbf{s}_{1j})\|^2}$$

을 최대화시키는 이러한 \mathbf{s}_{1j} 가 탐색되고,

[수학식 73]

$$\frac{\mathbf{y}^T W_v^T W_v (D_2 \mathbf{s}_{0i} + \mathbf{s}_{1j})}{\|W_v^T (D_2 \mathbf{s}_{0i} + \mathbf{s}_{1j})\|^2}$$

에 근접한 이득(\mathbf{g}_1)이 탐색된다.

제 3탐색방법으로서,

[수학식 74]

$$E_4' = \|W(\mathbf{x} - \mathbf{g}_1 \mathbf{s}_{0i})\|^2$$

을 최소화하는 이러한 $\mathbf{s}_{0i}, \mathbf{g}_1$ 이 탐색된다.

[수학식 75]

$$\frac{(\mathbf{y}^T W_v^T W_v (D_2 \mathbf{s}_{0i} + \mathbf{s}_{1j}))^2}{\|W_v(D_2 \mathbf{s}_{0i} + \mathbf{s}_{1j})\|^2}$$

을 최대화하는 이러한 s_{1j} 가 탐색되고,

[수학식 76]

$$\frac{\underline{v}^T W_v^T W_v (D_2 s_{0i} + s_{1j})}{\| W_v^T (D_2 s_{0i} + s_{1j}) \|^2}$$

에 근접한 이득(g1)이 최종적으로 선택된다.

다음에, 제 1탐색법의 수학식 69의 중심조건을 설명한다. 코드벡터(s_{0i})의 중심(s_{0c})에서

[수학식 77]

$$J = \frac{1}{M} \sum_{k=1}^M \| W'_k (x_k - g_k s_{0c}) \|^2$$

이 최소화된다. 이 최소화를 위해서,

[수학식 78]

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

가 해석되어,

[수학식 79]

$$s_{0c} = \left\{ \sum_{k=1}^M g_k^2 W'_k{}^T W'_k \right\}^{-1} \sum_{k=1}^M g_k W'_k{}^T W'_k x_k$$

을 준다. 유사하게, 이득(g)의 중심(g_c)에 대하여,

[수학식 80]

$$J = \frac{1}{M} \sum_{k=1}^M \| W'_k (x_k - g_c s_{0k}) \|^2$$

과,

[수학식 81]

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

이 상기 수학식 69로부터 해석되어

[수학식 82]

$$g_c = \frac{\sum_{k=1}^M s_{0k} W'_k{}^T W'_k x_k}{\sum_{k=1}^M s_{0k} W'_k{}^T W'_k s_{0k}}$$

을 준다.

한편, 제 1탐색법의 수학식 70의 중심조건으로서,

[수학식 83]

$$J = \frac{1}{M} \sum_{k=1}^M \| W_{vk} (\underline{v}_k - \underline{g}_k (D_{2k} \underline{s}_{0k} + \underline{s}_{1c})) \|^2$$

와

[수학식 84]

$$\frac{\partial J}{\partial \underline{s}_{1c}} = \frac{1}{M} \sum_{k=1}^M \{ -2 \underline{g}_k W_{vk}^T W_{vk} \underline{v}_k + 2 \underline{g}_k^2 W_{vk}^T W_{vk} D_{2k} \underline{s}_{0k} + 2 \underline{g}_k^2 W_{vk}^T W_{vk} \underline{s}_{1c} \} = 0$$

이 벡터(\underline{s}_{1j})의 중심(\underline{s}_{1c})에 대하여 해석되어,

[수학식 85]

$$\underline{s}_{1c} = \left\{ \sum_{k=1}^M \underline{g}_k^2 W_{vk}^T W_{vk} \right\}^{-1} \sum_{k=1}^M \underline{g}_k W_{vk}^T W_{vk} (\underline{v}_k - \underline{g}_k D_{2k} \underline{s}_{0k})$$

을 준다.

수학식 70으로부터 벡터(\underline{s}_{0i})의 중심(\underline{s}_{0c})이 구해져서,

[수학식 86]

$$J = \frac{1}{M} \sum_{k=1}^M \| W_{vk} (\underline{v}_k - \underline{g}_k (D_{2k} \underline{s}_{0c} + \underline{s}_{1k})) \|^2$$

[수학식 87]

$$\frac{\partial J}{\partial \underline{s}_{0c}} = \frac{1}{M} \sum_{k=1}^M \{ -2 \underline{g}_k D_{2k}^T W_{vk}^T W_{vk} \underline{v}_k + 2 \underline{g}_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0c} + 2 \underline{g}_k^2 D_{2k}^T W_{vk}^T W_{vk} \underline{s}_{1k} \} = 0$$

[수학식 88]

$$\underline{s}_{0c} = \left\{ \sum_{k=1}^M \underline{g}_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \right\}^{-1} \sum_{k=1}^M \underline{g}_k D_{2k}^T W_{vk}^T W_{vk} (\underline{v}_k - \underline{g}_k \underline{s}_{1k})$$

가 주어진다.

유사하게, 이득(g)의 중심(g_c)이

[수학식 89]

$$g_c = \frac{\sum_{k=1}^M (\underline{v}_k^T W_{vk}^T W_{vk} (D_{2k} \underline{s}_{0k} + \underline{s}_{1k}))}{\sum_{k=1}^M (D_{2k} \underline{s}_{0k} + \underline{s}_{1k})^T W_{vk}^T W_{vk} (D_{2k} \underline{s}_{0k} + \underline{s}_{1k})}$$

에 의해 구해진다.

상기 수학식 69에 의해 코드벡터(\underline{s}_{0i})의 중심을 계산하는 방법과 이득(g)의 중심(g_c)을 계산하는 방법을 수학식 82에 의해 나타내고 있다. 수학식 70에 의해 중심을 계산하는 방법으로서, 벡터(\underline{s}_{1j})의 중심(\underline{s}_{1c}), 벡터(\underline{s}_{0i})의 중심(\underline{s}_{0c}), 이득(g)의 중심(g_c)이 각각 수학식 85와 수학식 88과 수학식 89에 의해 표시되어 있다.

실제의 GLA에 의한 코드북의 학습에는 수학식 79와 수학식 85와 수학식 89를 사용하여 \underline{s}_0 , \underline{s}_1 , g 를 동시에 학습하는 방법이 있다. 상기 수학식 71과 수학식 72와 수학식 73은 탐색방법(가장 근접한 인접조건)에 대하여 사용될 수 있다. 또한, 수학식 79와 수학식 82와 수학식 85와 수학식 88과 수학식 85와 수학식 89에 의해 나타낸 중심조건의 여러 가지 조합이 선택적으로 사용될 수 있다.

도 14에 대응하는 수학식 66의 왜곡척도에 대한 탐색방법을 설명한다. 이 경우에서,

[수학식 90]

$$E_3' = \| W'(\underline{x} - g_1 \underline{s}_{0i}) \|^2$$

을 최소화하는 탐색(\underline{s}_{0i} , g_1)을 만족시키고, 따라서

[수학식 91]

$$E_3 = \| W_v(\underline{x} - g_1(D_2 \underline{s}_{0i} + \underline{s}_{1j})) \|^2$$

을 최소화하는 탐색(\underline{s}_{1j})을 만족시킨다.

상기 수학식 90에서, 모든 세트의 (g_1 , \underline{s}_{0i})를 획득하는 것은 현실적이지 못하므로,

[수학식 92]

$$\frac{(\underline{x}^T W'^T W' \underline{s}_{0i})^2}{\| W' \underline{s}_{0i} \|^2}$$

을 최대화하는 벡터(\underline{s}_{0i})의 상위 L수와,

[수학식 93]

$$\frac{\underline{x}^T W'^T W' \underline{s}_{0i}}{\| W' \underline{s}_{0i} \|^2}$$

에 근접한 이득의 L수가

[수학식 94]

$$E_3 = \| W_v(\underline{y} - (D_2 g \underline{s}_{0i} + \underline{s}_{1j})) \|^2$$

을 최소화하는 \underline{s}_{1j} 와 상기 수학식 92를 연합하여 탐색된다.

다음에, 중심조건이 수학식 90과 수학식 91로부터 주어진다. 이러한 경우에, 사용되는 식에 따라서 처리가 변화한다.

먼저, 수학식 90이 사용되면, 코드벡터(\underline{s}_{0i})의 중심이 \underline{s}_{0c} 이므로,

[수학식 95]

$$J = \frac{1}{M} \sum_{k=1}^M \| W'_k (x_k - g_k s_{0c}) \|^2$$

이 최소화되어,

[수학식 96]

$$s_{0c} = \left\{ \sum_{k=1}^M g_k^2 W_k'^T W_k' \right\}^{-1} \sum_{k=1}^M \{ g_k W_k'^T W_k' x_k \}$$

을 얻는다. 유사하게, 중심(g_c)에 대하여는 다음 수학식

[수학식 97]

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

이 수학식 92의 경우에서처럼, 상기 수학식 90에 의해 얻어진다.

벡터(s_{1j})의 중심(s_{1c})이 수학식 91을 사용하여 구해지면,

[수학식 98]

$$J = \frac{1}{M} \sum_{k=1}^M \| W_{vk} (v_k - D_{2k} g_k s_{0k} - s_{1c}) \|^2$$

과

[수학식 99]

$$\frac{\partial J}{\partial s_{1c}} = \frac{1}{M} \sum_{k=1}^M \{ 2 W_{vk}^T W_{vk} s_{1c} - 2 W_{vk}^T W_{vk} v_k + 2 g_k W_{vk}^T W_{vk} D_{2k} s_{0k} \} = 0$$

이 해석되어

[수학식 100]

$$s_{1c} = \left\{ \sum_{k=1}^M W_{vk}^T W_{vk} \right\}^{-1} \sum_{k=1}^M W_{vk}^T W_{vk} (v_k - g_k D_{2k} s_{0k})$$

이 주어진다.

유사하게, 벡터(s_{0i})의 중심(s_{0c})과 이득(g)의 중심(g_c)이 수학식 91에서 구해진다.

[수학식 101]

$$s_{0c} = \left\{ \sum_{k=1}^M g_k^2 D_{2k}^T W_{vk}^T W_{vk} D_{2k} \right\}^{-1} \sum_{k=1}^M g_k D_{2k}^T W_{vk}^T W_{vk} (v_k - g_k s_{1k})$$

[수학식 102]

$$J = \frac{1}{M} \sum_{k=1}^M \{ \underline{u}_k^T W_{vk}^T W_{vk} \underline{u}_k + g_c^2 (D_{2k} \underline{s}_{0k})^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0k} \\ + \underline{s}_{1k}^T W_{vk}^T W_{vk} \underline{s}_{1k} - 2g_c (D_{2k} \underline{s}_{0k})^T W_{vk}^T W_{vk} \underline{u}_k \\ - 2 \underline{s}_{1k}^T W_{vk}^T W_{vk} \underline{u}_k + 2 \underline{s}_{1k}^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0k} g_c \}$$

[수학식 103]

$$\frac{\partial J}{\partial g_c} = \frac{1}{M} \sum_{k=1}^M \{ 2g_c (D_{2k} \underline{s}_{0k})^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0k} - 2 (D_{2k} \underline{s}_{0k})^T W_{vk}^T W_{vk} \underline{u}_k + 2 \underline{s}_{1k}^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0k} \}$$

$$= 0$$

[수학식 104]

$$g_c = \frac{\sum_{k=1}^M \{ (D_{2k} \underline{s}_{0k})^T W_{vk}^T W_{vk} (\underline{u}_k - \underline{s}_{1k}) \}}{\sum_{k=1}^M \{ (D_{2k} \underline{s}_{0k})^T W_{vk}^T W_{vk} D_{2k} \underline{s}_{0k} \}}$$

한편, 상기 수학식 96과 수학식 97과 수학식 100을 이용하거나 상기 수학식 100, 수학식 101, 수학식 104를 이용하여 GLA에 의한 코드북학습을 실행한다.

본 발명의 CELP부호화 구조를 사용하는 제 2부호화기(120)는 단단 벡터 양자화 처리부(도 18의 본 실시예에서 2단 부호화부(120₁ ~ 120₂))를 가진다.

도 18은 전송비트율을 예를 들면 상기 2kbps와 6kbps로 전환가능한 경우에 있어서, 6kbps의 전송비트율에 대응한 구성을 나타내고 있고, 또한 형상 및 이득 인덱스 출력을 23비트/5msec와 15비트/5msec로 전환되도록 하고 있다. 도 18의 구성에 있어서의 처리의 흐름은 도 19에 나타낸 것과 같다.

도 18을 참조하여, 도 18의 제 1부호화부(300)는 도 3의 제 1부호화부(113)와 같고, 도 18의 LPC분석회로(302)는 도 3에 나타낸 LPC분석회로(132)에 대응하면서, LSP 파라미터 양자화회로(303)는 도 3의 α→LSP변환회로(133)에서 LSP→α 변환회로(137)까지의 구성에 대응하고, 도 18의 청각가중필터(304)는 도 3의 상기 청각가중필터 산출회로(139) 및 청각가중필터(125)와 대응하고 있다. 그러므로, 도 18에 있어서, 단자(305)에 상기 도 3의 제 1부호화부(113)의 LSP→α변환회로(137)에서의 출력과 동일한 것이 공급되고, 또 단자(307)에는 상기 도 3의 청각가중필터 산출회로(139)에서의 출력과 동일한 것이 공급되고, 또 단자(306)에는 상기 도 3의 청각가중필터(125)에서의 출력과 동일한 것이 공급된다. 그러나, 청각가중필터(125)로부터의 왜곡에서, 도 18의 청각가중필터(304)는 상기 도 3의 청각가중필터(125)와 같고 상기 LSP→α변환회로(137)의 출력을 이용하는 대신에 입력음성 데이터와 양자화 전의 α파라미터를 사용하여 청각가중된 신호를 생성하고 있다.

도 18에 나타내는 2단 제 2부호화부(120₁ ~ 120₂)에 있어서, 감산기(313 및 323)는 도 3의 감산기(123)와 대응하고, 거리계산회로(314 및 324)는 도 3의 거리계산회로(124)와 대응한다. 또한, 이득회로(311 및 321)는 도 3의 이득회로(126)와 대응하는 한편, 스테리케스틱 코드북(310, 320) 및 이득코드북(315, 325)은 도 3의 잡음코드북(121)과 대응하고 있다.

도 18의 구성에 있어서, 도 19의 단계(S1)에 나타낸 것같이, LPC분석회로(302)는 단자(301)에서 공급된 입력음성데이터(\underline{x})를 상술한 바와 같이 프레임으로 분할하여 LPC분석을 행하고 α파라미터를 구한다. LSP파라미터 양자화회로(303)는 LPC분석회로(302)에서의 α파라미터를 LSP파라미터로 변환하여 LSP파라미터를 양자화 한다. 양자화된 LSP데이터를 보간한 후, α파라미터로 변환한다. LSP파라미터 양자화회로(303)는 양자화된 LSP파라미터를 변환한 α파라미터에서 LPC 합성필터 함수(1/H(z))를 생성하고, 생성된 LPC합성필터 함수(1/H(z))를 단자(305)를 통하여 제 1단의 제 2부호화부(120₁)의 청각가중합성필터(312)에 보낸다.

청각가중필터(304)에서는 LPC분석회로(302)에서 α파라미터(즉 양자화 전의 α파라미터)에서 상기 도 3의 청각가중필터 산출회로(139)에 의해 산출된 것과 동일한 청각가중을 위한 데이터를 구한다. 이들 가중데이터가 단자(307)를 통하여, 제 1단의 제 2부호화부(120₁)의 청각가중합성필터(312)에 공급된다. 청각가중필터(304)는 도 19의 단계(S2)에 나타낸 것같이, 입력음성데이터와 양자화 전의 α파라미터에서 도 3의 청각가중필터(125)에 의한 출력과 동일신호의 청각가중된 신호

를 생성한다. 즉, 먼저 양자화 전의 α 파라미터에서 청각가중 필터함수($W(z)$)가 생성되고, 이렇게 생성된 필터함수($W(z)$)는 입력음성데이터(\underline{x})에 적용되어 단자(306)를 통하여 제 1단의 제 2부호화부(120₁)의 감산기(313)에 청각가중한 신호로서 보낸 \underline{x}_w 를 생성한다.

제 1단의 제 2부호화부(120₁)에서는 9비트 형상인덱스출력의 스테레오 코덱(310)에서의 대표값 출력이 이득회로(311)에 보내지고, 스테레오 코덱(310)에서의 대표값 출력이 6비트 이득인덱스 출력의 이득코덱(315)에서의 이득(스칼라값)을 곱한다. 이득회로(311)에서 이득이 곱해진 대표값 출력이 $1/A(z)=(1/H(z))*W(z)$ 의 청각가중의 합성필터(312)에 보내진다. 가중의 합성필터(312)에서는 도 19의 단계(S3)와 같이 $1/A(z)$ 의 제로입력 응답출력이 감산기(313)에 보내진다. 감산기(313)에서는 상기 청각가중 합성필터(312)에서의 제로입력 응답출력과, 상기 청각가중필터(304)에서의 상기 청각가중한 신호(\underline{x}_w)를 이용한 감산이 행해지고, 이 차분 혹은 오차가 참조벡터(\underline{r})로서 취해진다. 제 1단의 제 2부호화부(120₁)에서 참조벡터(\underline{r})는 거리가 계산되는 거리계산회로(314)에 보내지고 형상벡터(\underline{s})와 양자화 오차에너지를 최소화하는 이득(g)이 도 19에서 단계(S4)에 나타낸 것과 같이 탐색된다. 여기에서, $1/A(z)$ 는 제로상태에 있다. 즉, 제로상태에서 $1/A(z)$ 로 합성된 코덱에서 형상벡터(\underline{s})가 \underline{s}_{syn} 이면, 수학식 105를 최소화하는 형상벡터(\underline{s})와 이득이 탐색된다.

[수학식 105]

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

양자화 오차에너지를 최소화하는 \underline{s} 와 g 가 충분히 탐색되면, 계산의 양을 줄이기 위하여 다음의 방법이 사용될 수 있다.

제 1방법은 다음 수학식 106에 의해 정의된 E_s 를 최소화시키는 형상벡터(\underline{s})를 탐색하기 위한 것이다.

[수학식 106]

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

제 1방법에 의해 얻은 \underline{s} 로부터, 이상이득은 다음 수학식 107에 의해 나타내는 것과 같다.

[수학식 107]

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

그러므로, 제 2방법으로서 수학식 108을 최소화하는 이러한 g 가

[수학식 108]

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

탐색된다.

E 는 2차함수이므로, Eg 를 최소화하는 이러한 g 는 E 를 최소화한다.

[수학식 109]

$$\underline{e} = \underline{r} - g \underline{s}_{syn}$$

이것은 제 2단의 제 2부호화부(120₂)의 참조로서 제 1단에서와 같이 양자화 된다.

즉, 단자(305, 307)로 공급된 신호가 제 1단의 제 2부호화부(120₁)의 청각가중 합성필터(312)로부터 제 2단의 제 2부호화부(120₂)의 청각가중 합성필터(322)에 직접 공급된다. 제 1단의 제 2부호화부(120₁)에 의해 구해진 양자화 오차벡터(\underline{e})는 제 2단의 제 2부호화부(120₂)의 감산기(323)에 공급된다.

도 19의 단계(S5)에서, 제 2단의 제 2부호화부(120₂)에서 제 1단과 유사한 처리가 실행된다. 즉, 5비트 형상인덱스 출력의 스테레오 코덱(320)으로부터의 대표값 출력은 이득회로(321)에 보내져서 3비트 이득 인덱스출력의 이득코덱(325)으로부터 이득이 코덱(320)의 대표값 출력에 곱해진다. 가중합성필터(322)의 출력이 청각가중 합성필터(322)와 제 1단의 양자화 오차 벡터(\underline{e})사이의 차가 구해지는 감산기(323)에 보내진다. 이 차는 양자화 오차에너지(E)를 최소화하는 형상벡터(\underline{s})와 이득(g)을 탐색하기 위하여 거리계산을 위한 거리계산회로(324)에 보내진다.

제 1단의 제 2부호화부(120₁)의 스테레오 코덱(310)의 형상인덱스 출력과 이득코덱(315)의 이득인덱스 출력과 제 2단의 제 2부호화부(120₂)의 스테레오 코덱(320)의 형상인덱스 출력과 이득코덱(325)의 이득인덱스 출력이 인덱스출력 전환회로(330)에 보내진다. 제 1단과 제 2단의 제 2부호화부(120₁, 120₁)의 스테레오 코덱(310, 320)과 이득코덱(315, 325)의 인덱스데이터가 합쳐져서 출력된다. 15비트가 출력되면, 제 1단의 제 2부호화부(120₁)의 스테레오 코덱(310)과 이득코덱(315)의 인덱스데이터가 출력된다.

단계(S6)에 나타낸 것같이 제로입력 응답출력을 계산하기 위하여 필터상태가 갱신된다.

본 실시예에서, 제 2단의 제 2부호화부(120₂)의 인덱스 비트의 수가 형상벡터에 대하여 5만큼 작으면, 이득에 대한 것은 3만큼 작다. 코덱에서 이 경우에 적당한 형상과 이득이 제시되지 않으면, 양자화오차는 감소되는 대신에 증가하려고 한다.

이러한 문제가 발생하는 것을 방지하기 위하여 이득에서 0이 제공되지만, 이득에 대하여는 3비트뿐이다. 이들 중 하나가 0으로 설정되면, 양자화실행이 현저하게 저하된다. 이것을 고려하여 큰 비트수가 할당되는 형상벡터에 대하여 모두 0의 벡터가 제공된다. 제로벡터를 배제하여 진술한 탐색을 행하고, 양자화오차가 최종적으로 증가하여 버리면, 제로벡터가 선택된다. 이득은 임의이다. 이것에 의해, 제 2단의 제 2의 부호화부(120₂)에서 양자화오차가 증가하는 것을 방지할 수 있다.

도 18을 참조하여 2단구성의 경우를 예로 들고 있지만, 단수를 2보다 크게 할 수 있다. 이 경우, 제 1단의 페루프 루프 탐색에 의한 벡터양자화가 종료하면, 제 N단(2≤N)에서는 N-제 1단의 양자화오차를 기준입력으로서 양자화를 행하고, 제 N단의 양자화오차는 N+ 제 1단의 기준입력으로 사용된다.

제 2부호화부에 다단의 벡터양자화기를 이용함으로써, 도 18, 도 19에 나타낸 것같이, 동일 비트수의 직접 벡터양자화나 공액코덱 등을 이용한 것과 비교하여 계산량이 적게 된다. 특히, CELP부호화에서는 합성에 의한 분석법을 이용한 페루프탐색을 이용한 시간축 파형의 벡터양자화를 행하고, 탐색동작의 회수가 적은 것이 중요하다. 또, 제 2단의 제 2부호화부(120₁, 120₂)의 양측 인덱스출력을 이용하는 것과 제 2단의 제 2의 부호화부(120₂)의 출력을 사용하지 않고 제 1단의 제 2의 부호화부(120₁)의 출력만을 이용하는 것 사이에서 전환함으로써 비트수가 쉽게 전환될 수 있다.

제 1단과 제 2단의 제 2부호화부(120₁, 120₂)의 인덱스출력이 결합하여 출력되면, 복호기는 인덱스출력의 한 개를 선택함으로써 구조에 쉽게 대응할 수 있다. 즉, 복호기는 2kbps에서 복호화 동작을 사용하여 예를 들면 6kbps로 부호화된 파라미터를 복호화 함으로써 구조에 쉽게 대응할 수 있다. 또한, 제로벡터가 제 2단의 제 2의 부호화부(120₂)의 형상코덱에 포함되어 있으면, 0이 이득에 가해질 때 성능에서 보다 적게 저하되어서 양자화오차가 증가되는 것을 방지할 수 있게 된다.

스테레오 코덱(형상벡터)의 코드벡터는 예를 들면 다음의 방법에 의해 생성될 수 있다.

스테레오 코덱의 코드벡터는 예를 들면 소위 가우스잡음에 의해 클리핑에 의해 생성될 수 있다. 특히, 코덱은 가우스잡음을 생성하고 적당한 임계값으로 가우스잡음을 클리핑하고 클리핑된 가우스잡음을 정규화 함으로써 생성될 수 있다.

그러나, 음성에는 여러 가지의 형태가 있다. 예를 들면, "사, 시, 스, 세, 소"와 같은 잡음에 근접한 자음의 음성에 가우스잡음이 대응할 수 있는 반면, "파, 피, 푸, 페, 포"와 같이 급격하게 일어서는 자음의 음성에는 대응할 수 없다.

본 발명에 따르면, 가우스잡음은 몇몇의 코드벡터에 적용할 수 있는 반면, 코드벡터의 나머지는 학습에 의해 다루어져서, 급격하게 일어서는 자음과 잡음에 가까운 자음을 가지는 2개의 자음이 대응될 수 있다. 만약, 예를 들면, 임계값이 증가하면, 몇 개의 보다 큰 피크를 가지는 이러한 벡터가 구해지는 반면, 임계값이 감소하면, 코드벡터는 가우스잡음에 근접한다. 그래서, 임계값을 클리핑하는 데에 변화를 증가함으로써 "파, 피, 푸, 페, 포"와 같이 급격하게 일어나는 부분을 가지는 자음과 "사, 시, 스, 세, 소"와 같은 잡음에 근접한 자음에 대응할 수 있음으로써 명확도가 증가한다. 도 20은 실선과 파선에 의해 각각 가우스잡음과 클리핑된 잡음의 모양을 나타낸다. 도 20a, 20b는 1.0과 같게 즉 큰 임계값으로 임계값을 클리핑하는 잡음과 0.4와 같게 즉 작은 임계값으로 임계값을 클리핑하는 잡음을 나타낸다.

도 20a, 20b로부터 임계값이 크게 선택되면, 몇 개의 큰 피크를 가지는 벡터가 얻어지는 반면, 임계값이 작은 값으로 선택되면 잡음은 가우스잡음 자체에 근접하게 된다.

이것을 알아보기 위하여, 최초코드북은 가우스잡음을 클리핑함으로써 준비되는 반면, 비학습 코드벡터의 적당한 수가 설정된다. "사, 시, 스, 세, 소"와 같은 잡음에 근접한 자음에 대응하기 위하여 분산값이 증가하는 순서로 선택된다. 학습에 의해 구해진 벡터는 학습을 위해서 LBG알고리즘을 사용한다. 가장 근접한 인접조건하에서 부호화는 고정코드벡터와 학습에서 구해진 코드벡터를 사용한다. 중심조건에서, 학습되는 코드벡터만이 갱신된다. 그래서, 학습된 코드벡터는 "파, 피, 푸, 페, 포"와 같이 급격하게 일어서는 자음에 대응할 수 있다.

이들 코드벡터에 대하여 통상의 학습법에 의해서 최적의 이득이 학습될 수 있다.

도 21은 가우스잡음을 클리핑에 의한 코드북의 구조에 대한 처리흐름을 나타낸다.

도 21에서, 학습의 횟수(n)가 초기화하기 위하여 단계(S10)에서 n=0으로 설정된다.

오차(D₀=∞)로서 학습 횟수의 최대값(n_{max})이 설정되고, 학습종료조건을 설정하는 임계값(ε)이 설정된다.

다음 단계(S11)에서, 가우스잡음의 클리핑에 의한 초기 코드북이 생성된다. 단계(S12)에서 코드벡터의 일부가 비학습 코드벡터로서 고정된다.

다음의 단계(S13)에서 상기 코드북을 이용하여 부호화를 행한다. 단계(S14)에서 오차가 산출된다. 단계(S15)에서, (D_{n-1} - D_n/D_n < ε, or n = n_{max})인가 아닌가가 판단된다. 결과가 YES이면, 처리가 종료한다. 결과가 NO이면, 처리는 단계(S16)로 이동한다.

단계(S16)에서 부호화에 사용되지 않는 코드벡터가 처리된다. 다음의 단계(S17)에서, 코드북이 갱신된다. 단계(S18)에서, 학습 횟수(n)는 단계(S13)로 돌아가기 전에 증가된다.

도 3의 음성부호기에서, 유성음/무성음(V/UV) 판정부(115)의 구체적인 예가 설명된다.

V/UV판정부(115)는 직교전송회로(145)의 출력과 고정밀도의 피치탐색부(146)로부터의 최적피치와 스펙트럼평가부(148)에서의 스펙트럼 진폭데이터와 개방루프 피치탐색부(141)에서의 정규화 자기상관최대값(r(p))과 영교차 카운터(412)로부터 영교차 카운터값에 의거하여 상기 프레임의 V/UV판정이 행해진다.

MBE의 경우에서 m번째의 조파의 진폭을 나타내는 파라미터 혹은 진폭 |A_m|은

$$\therefore |A_m| = \frac{\sum_{j=a_m}^b |S(j)||E(j)|}{\sum_{j=a_m}^b |E(j)|^2}$$

에 의해 표시된다.

이 식에서, $|S(j)|$ 는 LPC나머지를 DFT하여 얻어진 스펙트럼이고, $|E(j)|$ 는 기저신호의 스펙트럼이고, 구체적으로는 256 포인트의 해밍창이고, a_m, b_m 은 제 m조파에 차례대로 대응하는 제 m대역에 대응하는 주파수의, 인덱스(j)에 의해 표시되는 상부 및 하부 극한값이다. 대역마다 V/UV판정을 위하여 신호대잡음비(NSR)가 사용된다. 제 m밴드의 NSR이

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

에 의해 표시된다.

NSR값이 0.3과 같이 리세트 임계값보다 크면, 즉 오차가 크면, 대역에서 $|A_m| |E(j)|$ 에 의한 $|S(j)|$ 의 근사가 좋지 않은 것으로 즉, 여기신호 $|E(j)|$ 가 기저로서 적당하지 않은 것으로 판단된다. 그래서 대역을 무성음(UV)으로 판단한다. 한편, 근사가 양호하게 이루어진 것으로 판단하면 유성음(V)으로 판단된다.

각각의 대역(조파)의 NSR이 1개의 조파로부터 다른 것까지 조파의 유사 도를 나타내고 있다. NSR의 이득가중 조파의 합계는

$$NSR_{all} = (\sum_m |A_m| NSR_m) / (\sum_m |A_m|)$$

에 의해 NSR_{all} 로서 정의된다. 이 스펙트럼유사도(NSR_{all})가 어느 임계값보다 큰가 작은가에 의해 V/UV판정에 이용되는 기본규칙이 결정된다. 여기에서 임계값은 $TH_{NSR} = 0.3$ 으로 설정된다. 이 기본규칙은 프레임파워, 영교차, LPC나머지의 자기상관의 최대치에 관한 것이고, $NSR < TH_{NSR}$

인 때에 이용되는 기본규칙에서는 규칙이 적용되면 프레임이 V, 적용되지 않으면 프레임이 UV로 된다.

구체적인 규칙은 다음과 같다.

$NSR_{all} < TH_{NSR}$ 에 대하여,

$numZero XP < 24, firmPow > 340$ 그리고 $r0 > 0.32$ 이면, 프레임은 V이다.

$NSR_{all} \geq TH_{NSR}$ 에 대하여,

$numZero XP > 30, firmPow < 900$ 그리고 $r0 > 0.23$ 이면, 프레임은 UV이다.

여기에서 각 변수는 다음과 같이 정의된다.

$numZero XP$: 프레임당 영교차 수

$firmPow$: 프레임 파워

$r0$: 자기상관의 최대치

상기와 같이 주어진 구체적인 규칙의 세트를 나타내는 규칙은 V/UV판정을 위한 것이다.

도 4의 음성신호 복호기의 주요부분과 동작의 구성은 보다 상세히 설명한다.

스펙트럼 엔벨로프의 역벡터양자화기(212)에서, 음성부호기의 벡터양자화기에 대응하는 역벡터양자화기 구조가 사용된다.

예를 들면, 벡터양자화가 도 12에 나타난 구조에 의해 실시되면, 복호기 측은 형상코드북(CB0, CB1)과 이득코드북(DBg)으로부터 코드벡터($\underline{s}_0, \underline{s}_1$)와 이득(g)을 읽어내고, 44차원과 같은 $g(\underline{s}_0 + \underline{s}_1)$ 고정차원의 벡터로서 취하여 원래 조파 스펙트럼의 벡터의 차원수에 대응하는 가변차원벡터가 변환되도록 한다(고정/가변 차원변환).

도 14 ~ 도 17에 나타난 것같이 부호기가 고정차원 코드벡터를 가변차원 코드벡터에 가산하는 벡터양자화기의 구조를 가지면, 가변차원에 대한 코드북(도 14의 코드북(CB0))으로부터 읽어낸 코드벡터가 고정/가변차원 변환되고 조파의 저역으로부터 차원수에 대응하는 고정차원(도 14의 코드북(CB1))에 대한 코드북으로부터 읽어낸 고정차원에 대한 코드벡터의 수에 가산된다. 결과합계가 취해진다.

도 4의 LPC합성필터(214)는 이미 설명한 것같이 유성음(V)에 대하여 합성필터(236)로, 무성음(UV)에 대하여 합성필터(237)로 분리된다. LSP가 V/UV구별 없이 합성필터를 구분하지 않고 20샘플마다 즉, 2.5msec마다 계속하여 보간되면, 전체 다른 성질의 LSP가 V→UV, UV→V천이부에서 보간된다. UV 및 V의 LPC가 각각 V 및 UV의 나머지로 사용되므로, 이상한 소리가 발생되게 된다. 이러한 좋지 않은 효과가 생기는 것을 방지하기 위하여, LPC합성필터가 V 및 UV로 분리되고, LPC계수 보간이 V 및 UV에 대하여 독립하여 실행된다.

이러한 경우에 LPC필터(236, 237)의 계수 보간의 방법을 설명한다. 특히, LSP보간이 도 22에 나타난 것처럼 V/UV에 의거하여 전환된다.

10차 LPC분석의 예를 취하면, 도 22에서 등간격 LSP가 평탄한 필터특성과 이득이 1, 즉

$$a_0 = 1, a_1 = a_2 = \dots = a_{10} = 0, 0 \leq \alpha \leq 10$$

에 대하여 α 파라미터에 대응하는 이러한 LSP이다.

이러한 10차 LPC분석, 즉 10차 LSP가 도 23에 나타난 것같이 0과 π 사이에서 11개로 갈래 분리된 부분에서 등간격으로 배열된 LSP로 완전히 편평한 스펙트럼에 대응하는 LSP이다. 이러한 경우에서, 합성필터의 전체대역이득이 이 시각에서 최소 스루(through)특성을 가진다.

도 24는 이득변화의 방법을 개략적으로 나타낸다. 구체적으로 도 15는 $1/H_{uv(z)}$ 이득과 이득 $1/H_{v(z)}$ 이 무성음(UV)부로부터 유성음(V)부로의 천이동안 어떻게 변화하는가를 나타낸다.

보간의 단위에 대하여, $1/H_{v(z)}$ 의 계수에 대하여 2kbps의 비트율에 대하여 10msec(80샘플)이고, $1/H_{uv(z)}$ 의 계수에 대하여 6kbps의 비트율에 대하여 5msec(40샘플)이다. UV에 대하여, 제 2부호화부(120)는 합성법에 의한 분석을 사용하여 과형매칭을 실시하기 때문에, 인접하는 V부의 LSP에서 보간이 등간격LSP에서 보간을 실행하지 않고 실행될 수 있다. 제 2부호화부(120)에서 UV부의 부호화에서 제로입력응답이 V→UV천이부에서 $1/A(z)$ 가중합성필터(122)의 내부상태를 클리어함으로써 0으로 설정된다.

LPC합성필터(236, 237)의 출력이 각각의 독립적으로 설치된 포스트필터(238u, 238v)에 보내진다. 포스트필터의 강도와 주파수응답은 V 및 UV에 대하여 다른 값으로 설정된다.

LPC나머지신호의 V 및 UV부 사이에서 연결부의 윈도우잉(windowing) 즉, LPC합성필터입력으로서 여기가 설명된다. 윈도우잉은 도 4에 나타난 무성음합성부(211)의 윈도우잉회로(223)와 유성음합성부(211)의 사인파합성회로(215)에 의해 실행된다. 여기의 V부 합성법은 본 출원인에 의해 출원된 JP특허출원 No. 4-91422에 상세히 설명되어 있고, 여기의 V부의 패스트합성법이 본 출원인에 의해 유사하게 출원된 JP특허출원 NO. 6-198451에 상세히 설명되어 있다.

유성음부(V)에는 인접하는 프레임의 스펙트럼을 이용하여 스펙트럼을 보간하여 사인파합성하기 위해, 도 25에 나타난 것같이 제 n프레임과 제 n+1프레임과의 사이에 모든 과형을 만들 수 있다. 그러나, 도 25의 제 n+1프레임과 제 n+2프레임과 같이, V와 UV에 걸리는 신호부분 혹은 V와 UV에 걸리는 부분에는 UV부분은 프레임 중에 ± 80 샘플(전체 160샘플의 총수가 1프레임간격이다)의 데이터만을 부호화 및 복호화하고 있다. 그 결과, 도 26에 나타난 것같이 V측에서는 프레임과

프레임과의 사이의 중심점(CN)을 넘어서 윈도잉을 행하고, UV측에서는 중심점(CN)까지의 윈도잉을 행하고, 인접부분을 중첩시키고 있다. UV→V의 천이부에 대하여 그 역을 행하고 있다. 또한, V측의 윈도잉은 도 26에 파선으로 나타낸 것과 같이 할 수 있다.

유성음(V)부분에서의 잡음합성 및 잡음가산에 대하여 설명한다. 이것은 도 4의 잡음합성회로(216), 가중중첩회로(217), 및 가산기(218)를 이용하여 유성음부분의 여기와 LPC합성필터입력으로서 다음의 파라미터를 고려한 잡음을 LPC나머지 신호의 유성음부분에 더함으로써 행해진다.

즉, 상기 파라미터로서는 피치랙(Pch), 유성음의 스펙트럼진폭(Am[i]), 프레임내의 최대 스펙트럼진폭(Amax), 및 나머지 신호의 벡터(Lev)를 열거할 수 있다. 여기에서, 피치랙(Pch)은 소정의 샘플링주파수(fs), fs=8kHz와 같이 피치주기내의 샘플 수이고, 스펙트럼진폭 Am[i]의 I는 fs/2의 대역내의 조파의 수를 I=Pch/2로 하는 동안 0<i<I의 범위내의 정수이다.

잡음합성회로(216)에 의한 처리는 예를 들면 다중대역 부호화(MBE)의 무성음의 합성과 동일한 방법으로 행해진다. 도 27은 잡음합성회로(216)의 구체적인 예를 나타내고 있다.

즉 도 27에 있어서, 백색잡음 발생회로(401)는 가우스잡음을 출력하여서, STFT처리부(402)에 의해 STFT(short-term Fourier transform)처리를 실시함으로써 잡음의 주파수축상의 전력스펙트럼을 얻는다. 가우스잡음은 소정의 길이(예를 들면 256샘플)를 가지는 해밍창과 같이 적당한 윈도잉함수에 의해 윈도잉으로 된 시간축 백색잡음 신호파형이다. STFT처리부(402)에서의 전력스펙트럼을 진폭처리를 위하여 곱셈기(403)에 보내고, 잡음진폭제어회로(410)에서의 출력과 승산되고 있다. 증폭기(403)에서의 출력은 ISTFT처리부(404)에 보내지고, 위상은 원래의 백색잡음의 위상을 이용하여 역 STFT(ISTFT)처리를 실시함으로써 시간축 상의 신호로 변환한다. ISTFT처리부(404)에서의 출력은 가중중첩 가산회로(217)에 보내진다.

도 27의 예에 있어서는 백색잡음 발생부(401)에서 시간영역의 잡음을 발생하여 그것을 STFT 등의 직교변환을 행하므로 주파수영역의 잡음을 얻었다. 그러나, 잡음발생부에서 직접적으로 주파수영역의 잡음을 발생하도록 하여도 좋다. 즉, 주파수영역의 잡음을 직접 발생함으로써 STFT나 ISTFT 등의 직교변환처리가 제거될 수 있다.

구체적으로는 ±x의 범위가 발생하고 그것을 FFT스펙트럼의 실부와 허부로서 취급하도록 하는 방법이나, 0에서 최대값(max)까지의 범위의 정의 난수를 발생하고 그것을 FFT스펙트럼의 진폭으로서 취급한다. -π에서 π까지의 난수를 발생하고 그것을 FFT스펙트럼의 위상으로서 취급하는 방법 등이 고려된다.

이것은 도 27의 STFT처리부(402)를 제거할 수 있고, 구성의 간략화 혹은 처리량의 저감화가 도모될 수 있다.

잡음진폭제어회로(410)는 예를 들면 도 28과 같은 기본구성을 가지고, 상기 도4의 스펙트럼 엔벨로프의 양자화기(212)에서 단자(411)를 통하여 주어지는 V(유성음)의 스펙트럼진폭Am[i]에 의거하여 곱셈기(403)에서 곱셈계를 제어함으로써 합성된 잡음진폭 Am_noise[i]를 구할 수 있다. 즉, 도 28에서, 스펙트럼 진폭회로 Am[i]와 피치랙(Pch)이 입력되는 최적 noise_mix값의 산출회로(416)에서의 출력을 잡음의 가중회로(417)에서 가중하고, 얻어진 출력을 곱셈기(418)에 보내어 스펙트럼진폭 Am[i]와 곱셈함으로써 잡음진폭Am_noise[i]를 얻고 있다.

잡음합성가산의 제 1구체적인 예로서, 잡음진폭(Am_noise[i])이 상기 4개의 파라미터내의 2개, 즉 피치랙(Pch) 및 스펙트럼진폭(Am)의 함수가 되는 경우에 대하여 설명한다.

이와 같은 함수(f₁(Pch, Am[i]))가운데,

$$f_1(Pch, Am[i])=0 \quad (0 < i < Noise_b \times I)$$

$$f_1(Pch, Am[i])=Am[i] \times noise_mix \quad (Noise_b \times I \leq i < I)$$

$$noise_mix=K \times Pch/2.0$$

이다.

noise_mix값의 최대치는 noise_mix_max이고, 그 값이 클리핑된다. 일 예로서, K=0.02, noise_mix_max=0.3, Noise_b=0.7에서, Noise_b는 전체 대역으로부터 이 잡음이 가산되는 부분을 결정하는 정수이다. 본 실시예에서는 70%보다 높은 주파수영역, 즉 fs=8kHz의 때, 4000 × 0.7=2800Hz에서 4000Hz까지의 범위에서 잡음이 가산된다.

잡음합성가산의 제 2구체적인 예로서, 상기 잡음진폭(Am_noise[i])을 상기 4개의 파라미터내의 3개, 즉 피치랙(Pch) 및 스펙트럼진폭(Am) 및 최대스펙트럼진폭(Amax)의 함수(f₂(Pch, Am[i], Amax))로 하는 경우에 대하여 설명한다.

이들 함수(f₂(Pch, Am[i], Amax))가운데,

$$f_2(\text{Pch}, \text{Am}[i], \text{Amax})=0 \quad (0 < i < \text{Noise_b} \times I)$$

$$f_2(\text{Pch}, \text{Am}[i], \text{Amax})=\text{Am}[i] \times \text{noise_mix} \quad (\text{Noise_b} \times I \leq i < I)$$

$$\text{noise_mix}=K \times \text{Pch}/2.0$$

이다.

noise_mix값의 최대값은 noise_mix_max이고, 일 예로서, K=0.02, noise_mix_max=0.3, Noise_b=0.7이다.

만약 Am[i] × noise_mix > Amax × C × noise_mix 이면, f₂(Pch, Am[i], Amax)=Amax × C × noise_mix이고, 여기서 정수(C)는 C=0.3으로 설정하고 있다. 이 조건식에 의해 잡음레벨이 매우 크게 되는 것을 방지할 수 있기 때문에, 상기 K, noise_mix_max를 다시 크게 하여도 좋고, 고역의 레벨도 비교적 큰 때에 잡음레벨을 높일 수 있다.

잡음합성가산의 제 3구체적인 예로서, 상기 잡음진폭(Am_noise[i])을 상기 4개의 파라미터내의 4개 전체의 함수(f₃(Pch, Am[i], Amax, Lev))로 할 수 있다.

이와 같은 함수(f₃(Pch, Am[i], Amax, Lev))의 구체적인 예는 기본적으로는 상기 제 2구체예의 함수(f₂(Pch, Am[i], Amax))와 동일하다. 단, 나머지신호레벨(Lev)은 스펙트럼진폭(Am[i]의 rms(root mean square) 혹은 시간축 상에서 측정된 신호레벨이다. 상기 제 2구체적인 예와의 다름은 K의 값과 noise_mix_max의 값을 Lev함수로 하는 점이다. 즉, Lev가 작거나 크면, K, noise_mix_max의 값은 각각 크거나 작게 설정된다. 또한, Lev는 K, noise_mix_max에 역비례하도록 설정될 수 있다.

다음에, 포스트필터(238v, 238u)에 대하여 설명한다.

도 29는 도 4 실시예의 포스트필터(238v, 238u)로서 이용되는 포스트필터를 나타내고 있다. 포스트필터의 요부로서 스펙트럼정형필터(440)는 포먼트(formant)강조필터(441)와 고역강조필터(442)로 이루어진다. 스펙트럼정형필터(440)에서의 출력은 스펙트럼정형에 의한 이득변화를 보정하기 위한 이득조정회로(443)에 보내진다.

이 이득조정회로(443)의 이득(G)은 이득제어회로(445)에 의해 스펙트럼정형필터(440)의 입력(x)과 출력(y)을 비교하여 이득변화를 계산하고, 보정값을 산출하는 것으로 결정된다.

스펙트럼 정형필터(440)의 특성PF(z)은 LPC합성필터의 분모(Hv(z), Huv(z))의 계수, 즉 a파라미터를 a_i로 하면,

$$PF(z) = \frac{\sum_{i=0}^p a_i \beta^i z^{-i}}{\sum_{i=0}^p a_i \gamma^i z^i} (1 - kz^{-1})$$

로 표현된다.

이 식의 분수부분이 포먼트강조필터특성을 나타내고, (1-kz⁻¹)의 부분이 고역강조필터의 특성을 나타낸다. β, γ, k는 정수이고 일 예로서 β=0.6, γ=0.8, k=0.3이다.

이득조정회로(443)의 이득(G)은

$$G = \sqrt{\frac{\sum_{i=0}^{159} x^2(i)}{\sum_{i=0}^{159} y^2(i)}}$$

에 의해 주어진다.

상기 식에서, x(i)와 y(i)는 스펙트럼정형필터(440)의 입력과 출력을 각각 나타낸다.

상기 스펙트럼 정형필터(440)의 계수의 갱신주기는 도 30에 나타난 것같이, LPC합성필터의 계수인 a파라미터의 갱신주기와 동일하게 20샘플, 2.5msec이고, 이득조정회로(443)의 이득(G)의 갱신주기는 160샘플, 20msec이다.

이와 같이, 포스트필터의 스펙트럼 정형필터(440)의 계수의 갱신주기보다 이득조정회로(443)의 갱신주기를 길게 취함으로써 이득조정의 변동에 의한 악영향을 방지할 수 있게 된다.

즉, 일반의 포스트필터에 있어서는 스펙트럼정형필터의 계수의 갱신주기와 이득의 갱신주기를 동일하게 하고 있고, 이때 이득의 갱신주기가 20샘플, 2.5msec로 선택되면, 도 30에 나타난 것같이 1피치의 주기 중에서 변동하게 되고, 클릭잡음이 생기는 원인으로 된다. 본 예에 있어서는 이득의 전환주기를 보다 길게, 예를 들면 1프레임분의 160샘플, 20msec로 하게 되고, 급격한 이득의 변동을 방지할 수 있다. 역으로 스펙트럼 정형필터의 계수의 갱신주기가 160샘플, 20msec이면, 원활한 필터특성의 변화가 얻어지지 않고, 합성과형에 악영향이 생긴다. 그러나, 이 필터계수의 갱신주기를 20샘플, 2.5msec로 짧게 함으로써 보다 효과적인 포스트필터처리가 가능하게 된다.

인접하는 프레임간에서의 이득의 연결처리는 이전 프레임의 필터계수 및 이득과 현 프레임의 필터계수 및 이득이 $W(i) = I/120$ ($0 \leq i \leq 20$), 페이드인, 페이드아웃에 대하여 $1-W(i)$ ($0 \leq i \leq 20$)의 삼각 원도잉에 의해 곱해지고, 결과 곱이 가산된다. 도 31에서는 전 프레임의 이득(G1)이 현 프레임의 이득(G2)에 합쳐지는 모습을 나타내고 있다. 특히, 전 프레임의 이득, 필터계수를 사용하는 비율이 서서히 감소하고, 현 프레임의 이득, 필터계수의 사용이 서서히 증대한다. 또한, 도 31의 시각(T)에 있어서의 전 프레임에 대한 현 프레임의 필터의 내부상태가 동일상태, 즉 전 프레임의 최종상태에서 시작한다.

이상 설명한 것과 같은 신호부호 및 신호복호화장치는 예를 들면 도 32 및 도 33에 나타난 것과 같은 휴대통신단말기 혹은 휴대전화기 등에 사용되는 음성코드북으로서 사용할 수 있다.

도 32는 상기 도 1, 도 3에 나타난 것과 같은 구성을 가지는 음성부호화부(160)를 이용하는 휴대단말기의 송신측 구성을 나타내고 있다. 마이크로폰(161)에서 집음된 음성신호는 증폭기(162)에서 증폭되고, A/D(아날로그/디지털) 변환기(163)에서 디지털신호로 변환되어서, 도 1, 도 3에 나타난 것과 같은 구성을 가지는 음성부호화부(160)에 보내진다. 입력단자(101)에 상기 A/D변환기(163)에서의 디지털신호가 입력된다.

음성부호화부(160)에서는 상기 도 1, 도 3과 함께 설명한 것과 같은 부호화처리가 행해진다. 도 1, 도 2의 각 출력단자에서의 출력신호는 음성부호화부(160)의 출력신호로서 전송로 부호화부(164)에 보내지고, 공급된 신호에 채널부호화를 실행한다.

전송로 부호화부(164)의 출력신호는 변조를 위하여 변조회로(165)에 보내지고, D/A(디지털/아날로그)변환부(166), RF증폭기(167)를 거쳐서 안테나(168)에 보내진다.

도 33은 상기 도 2, 도 4에 나타난 것과 같은 구성을 가지는 음성복호화부(260)를 이용하여 휴대단말기의 수신측을 나타내고 있다. 이 도 33의 안테나(261)에서 수신된 음성신호는 RF증폭기(262)에서 증폭되고, A/D(아날로그/디지털) 변환기(263)를 거쳐서 복조회로(264)에 보내지고, 복조신호가 전송로 복호화부(265)에 보내진다. 복호부(265)의 출력신호는 상기 도 2, 도 4에 나타난 것과 같은 구성을 가지는 음성복호화부(260)에 보내진다. 음성복호화부(260)는 상기 도 2, 도 4와 함께 설명한 바와 같은 방법으로 신호를 복호화 한다. 도 2, 도 4의 출력단자(201)에서의 출력신호가 음성복호화부(260)에서의 신호로서 D/A(디지털/아날로그) 변환기(266)에 보내진다. 이 D/A변환기(266)에서의 아날로그 음성신호가 스피커(268)에 보내진다.

본 발명은 상기 실시예에만 한정되는 것은 아니다. 예를 들면 상기 도 1, 도 3의 음성분석측(부호화 측)의 구성이나, 도 2, 도 4의 음성합성측(복호화 측)의 구성에 대하여는 각부를 하드웨어적으로 기재하고 있지만, 예를 들면 DSP(디지털신호 프로세서) 등을 이용하여 소프트웨어 프로그램에 의해 실현할 수 있다. 복호화 측에서의 합성필터(236, 237) 또는 포스트필터(238v, 238u)는 유성음 또는 무성음으로 분리되지 않고 오직 LPC합성필터 또는 포스트필터에 의해 구성된다. 본 발명은 전송이나 기록재생에 한정되지 않고, 피치변환이나 스피드변환, 규칙음성합성, 혹은 잡음억압과 같은 여러 가지의 용도에 이용할 수 있다.

발명의 효과

이상의 설명에서 명백한 바와 같이, 본 발명에 의하면, 가변차원의 입력벡터를 벡터양자화 할 때에, 코드북에서 독출된 고정차원의 코드벡터를 원래의 입력벡터의 차원과 동일가변차원으로 변환하고, 이 고정/가변차원 변환된 가변차원의 코드벡터에 대하여 원래의 입력벡터와의 오차를 최소화하는 최적의 코드벡터를 코드북에서 선택하고 있기 때문에, 최적의 코드벡터를 코드북에서 선택하는 코드북검색의 때에는 원래의 가변차원의 입력벡터와의 사이의 오차 혹은 왜곡이 계산되고, 양자화벡터의 정밀도를 높일 수 있다.

여기에서, 코드북을 형상코드북과 이득코드북으로 구성하는 경우에, 적어도 이득코드북에서의 이득의 최적화를 가변차원의 형상벡터와 입력벡터에 의거하여 행하도록 하는 것을 든다. 이 경우 또한 원래의 가변차원의 입력벡터를 형상코드북의 고정차원으로 변환하고, 이 가변/고정차원 변환된 고정차원의 입력벡터와 형상코드북에 축적된 코드벡터와의 오차를 최소화하는 단수 혹은 복수의 코드벡터를 형상코드북에서 선택하고, 형상코드북에서 독출되어 고정/가변차원 변환된 가변차원의 코드벡터와 입력벡터에 의거하여 고정/가변차원 변환된 코드벡터에 대한 최적이득을 선택하는 것을 든다.

이와 같이 이득을 고정/가변차원 변환된 가변차원의 코드벡터에 대하여 적용함으로써 고정차원코드벡터를 이득배한 것을 고정/가변차원 변환하는 경우에 비해서 고정/가변차원변환에 의한 왜곡의 영향을 적게 억누를 수 있다.

또, 원래의 가변차원의 입력벡터를 코드북의 고정차원으로 변환하고, 이 가변/고정차원 변환된 고정차원의 입력벡터와 코드북에 축적된 코드벡터와의 오차를 최소화하는 복수의 코드벡터를 형상코드북에 의해 일시적으로 선택하고, 이 일시적으로 선택된 코드벡터에 대하여 고정/가변차원변환을 행하여 가변차원에서 최적의 코드벡터를 선택하는 것을 든다.

이 경우, 일시적 선택에서의 탐색을 간략화 함으로써, 코드북검색에 요하는 처리량을 저감할 수도 있고, 또 가변차원에서 최종 선택함으로써 정밀도를 높일 수 있다.

이와 같은 벡터양자화를 음성부호화에 적용할 수 있고, 예를 들면 입력음성신호 또는 입력음성신호의 단기에측나머지를 사인파분석하여 조파 스펙트럼을 구하고, 부호화단위마다의 상기 조파 스펙트럼에 의거하는 파라미터를 입력벡터로서 벡터양자화하는 때에 적용할 수 있고, 정밀도가 높은 코드북탐색에 의한 음질의 향상을 도모할 수 있다.

(57) 청구의 범위

청구항 1.

각각의 코드북의 코드벡터의 하나의 인덱스를 출력하기 위하여 입력 벡터와 복수의 코드북에 저장된 코드벡터와 비교되는 벡터양자화 방법에 있어서,

음성신호로부터 유래된 주파수축상의 파라미터로 형성되어 있는 입력벡터와 각각의 복수의 코드북에 저장된 모든 코드벡터들 사이의 유사도를 구하고, 고정 차원에서 각각의 복수의 코드북에서 고도의 유사도를 나타내는 복수의 코드벡터를 예비선택하는 예비선택단계와,

상기 예비선택된 복수의 코드벡터로부터 가변차원에서 입력벡터로부터 최소 에러를 가지는 각각의 코드북에 대해서 코드벡터를 더 선택하는 최종 선택단계를 포함하여 구성된 것을 특징으로 하는 벡터양자화 방법.

청구항 2.

제 1항에 있어서,

상기 유사도는

입력벡터와 상기 코드벡터의 내적,

입력벡터와 상기 코드벡터의 가중된 내적,

각 코드벡터의 노름에 의해 분할된 상기 내적의 값,

각 코드벡터의 가중된 노름에 의해 분할된 상기 내적의 값,

각각의 코드벡터의 노름에 의해 분할된 상기 가중된 내적의 값, 및

각각의 코드벡터의 가중된 노름에 의해 분할된 상기 가중된 내적의 값 중 하나인 것을 특징으로 하는 벡터양자화 방법.

청구항 3.

입력음성신호가 소정의 부호화 단위로 시간축상에서 분할되고 소정의 부호화 단위로 부호화되는 음성부호화 방법에 있어서,

입력음성신호에서 유래된 신호의 사인파분석에 의해 조파의 스펙트럼성분을 구하는 단계와, 그리고

부호화용 입력벡터로서 부호화 단위에 기초한 조파들의 스펙트럼성분에서 유래하는 파라미터를 벡터양자화하는 단계로 구성되는 것을 특징으로 하며,

상기 벡터양자화단계는

음성신호로부터 유래된 주파수축상의 파라미터로 형성되어 있는 입력벡터와 각각의 복수의 코드북에 저장된 모든 코드벡터들 사이의 유사도를 구하고, 고정 차원에서 각각의 복수의 코드북에서 고도의 유사도를 나타내는 복수의 코드벡터를 예비선택하는 예비선택단계와,

상기 예비선택단계에서 선택된 예비선택된 복수의 코드벡터로부터 가변차원에서 입력벡터로부터 최소 에러를 가지는 코드벡터를 더 선택하는 최종 선택단계를 포함하여 구성된 것을 특징으로 하는 음성부호화 방법.

청구항 4.

제 3항에 있어서,

상기 유사도는

입력벡터와 상기 코드벡터의 내적,

입력벡터와 상기 코드벡터의 가중된 내적,

각 코드벡터의 노름에 의해 분할된 상기 내적의 값,

각 코드벡터의 가중된 노름에 의해 분할된 상기 내적의 값,

각각의 코드벡터의 노름에 의해 분할된 상기 가중된 내적의 값, 및

각각의 코드벡터의 가중된 노름에 의해 분할된 상기 가중된 내적의 값 중 하나인 것을 특징으로 하는 음성부호화 방법.

청구항 5.

입력음성신호가 소정의 부호화 단위에 대해 시간축상에서 분할되고 소정의 부호화 단위에 대해 부호화되는 음성부호화 장치에 있어서,

입력음성신호의 단기 예측나머지를 구하기 위한 예측부호화 수단과, 그리고

입력벡터로서 사인파분석에 의해 얻어진 조파들의 스펙트럼성분에서 유래하는 파라미터를 양자화하는 벡터양자화수단을 포함하는 사인파분석 부호화 수단을 포함하며,

상기 벡터양자화수단은

입력벡터와 각각의 복수의 코드북에 저장된 모든 코드벡터들 사이의 유사도를 구하고, 고정 차원에서 각각의 복수의 코드북에서 고도의 유사도를 나타내는 복수의 코드벡터를 예비선택하는 예비선택수단과,

상기 예비선택수단에 의해 선택된 복수의 코드북에서의 예비선택된 복수의 코드북으로부터 가변차원에서 입력벡터로부터 최소 에러를 가지는 코드벡터를 더 선택하는 최종 선택수단을 포함하여 구성된 것을 특징으로 하는 음성부호화 장치.

청구항 6.

제 5항에 있어서,

상기 유사도는

입력벡터와 상기 코드벡터의 내적,

입력벡터와 상기 코드벡터의 가중된 내적,

각 코드벡터의 노름에 의해 분할된 상기 내적의 값,

각 코드벡터의 가중된 노름에 의해 분할된 상기 내적의 값,

각각의 코드벡터의 노름에 의해 분할된 상기 가중된 내적의 값, 및

각각의 코드벡터의 가중된 노름에 의해 분할된 상기 가중된 내적의 값 중 하나인 것을 특징으로 하는 음성부호화 장치.

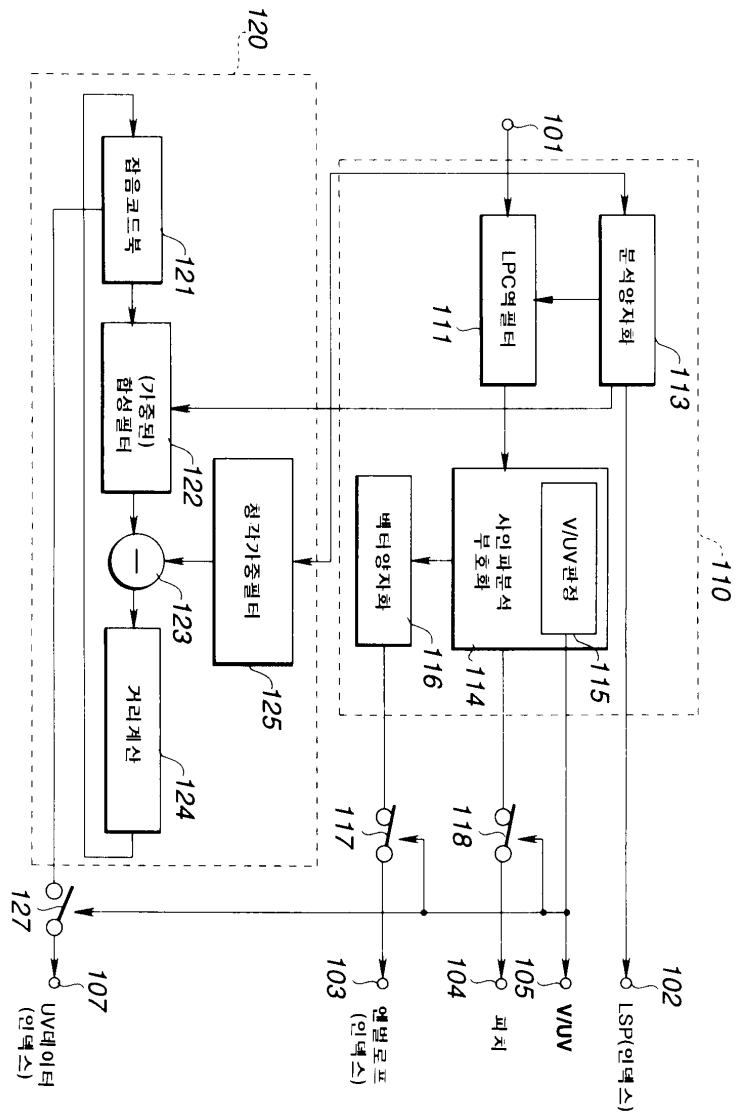
청구항 7.

제 6항에 있어서,

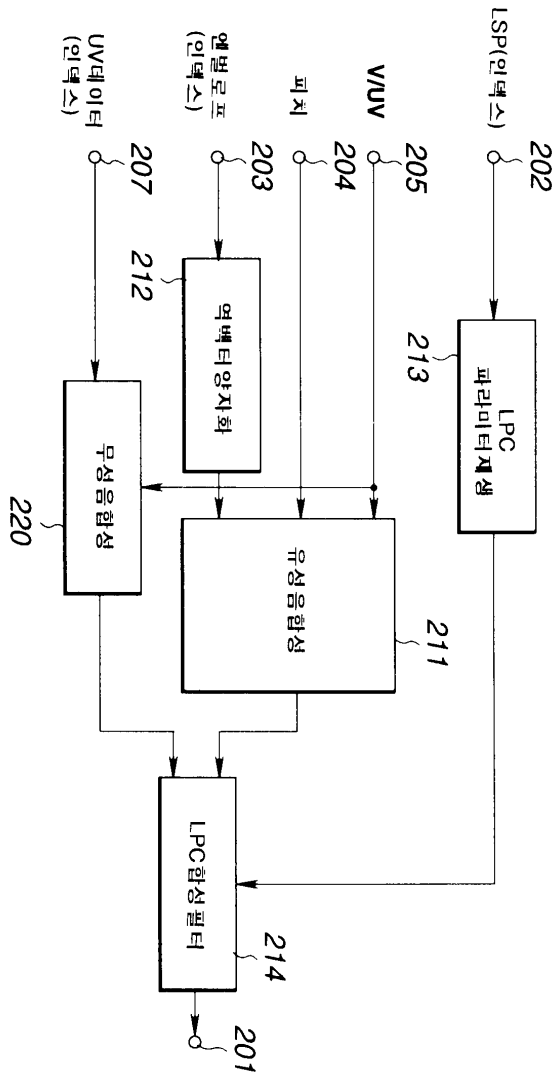
노름의 가중으로서, 입력벡터와 상기 코드벡터의 내적, 및 각각의 코드벡터의 노름 및 가중된 노름중 하나에 의해 분할된 입력벡터와 상기 코드벡터의 가중된 내적 중 하나가 사용되는 것을 특징으로 하는 음성부호화 장치.

도면

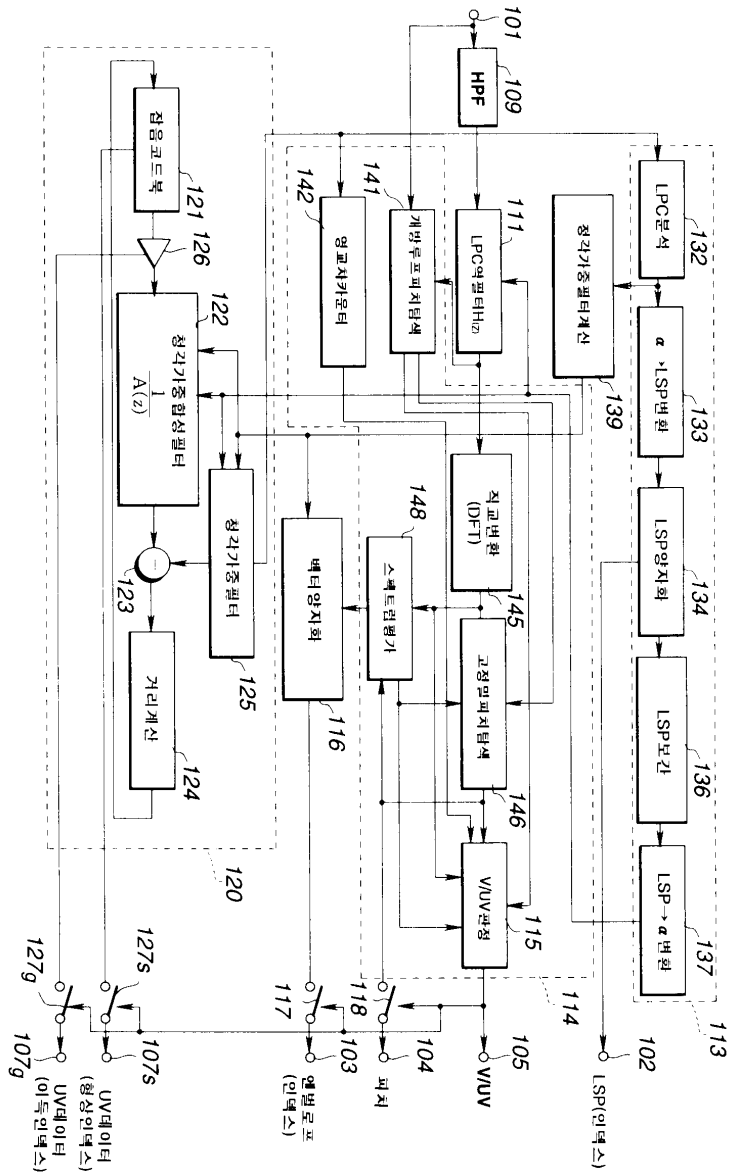
도면1



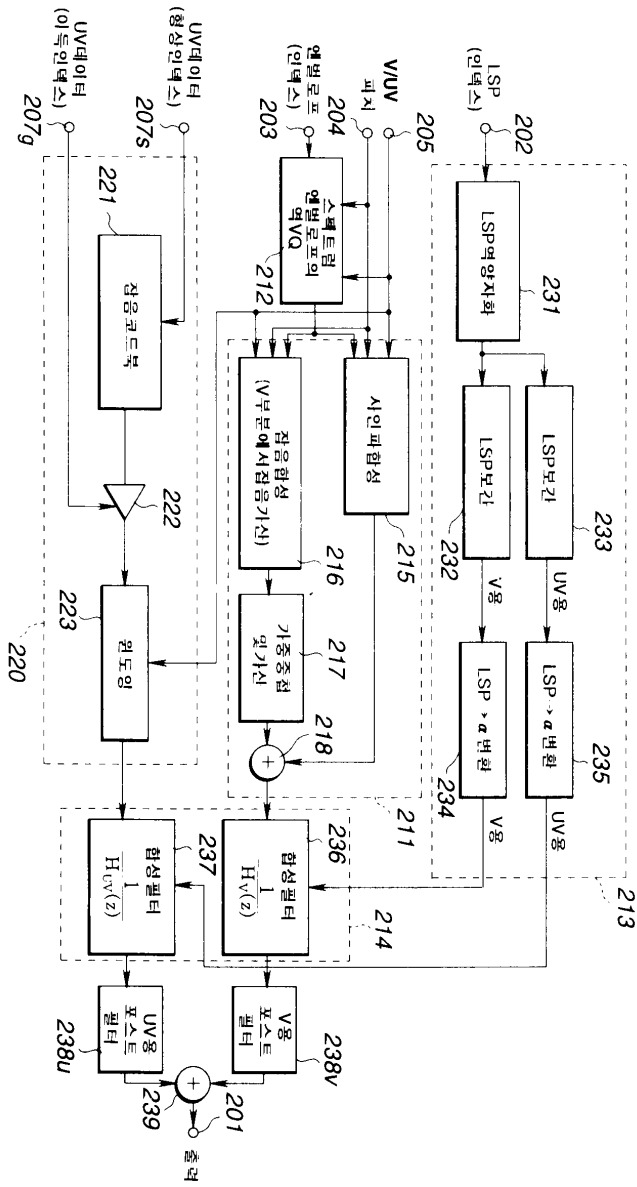
도면2



도면3



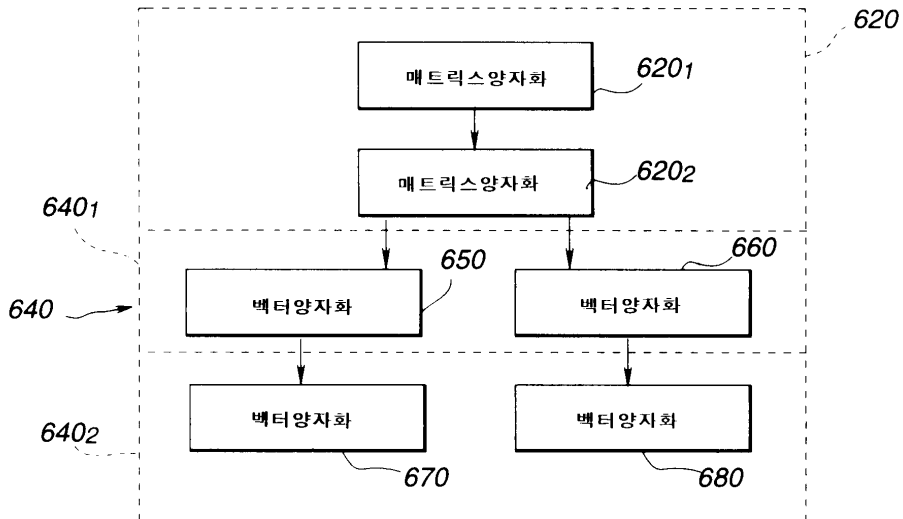
도면4



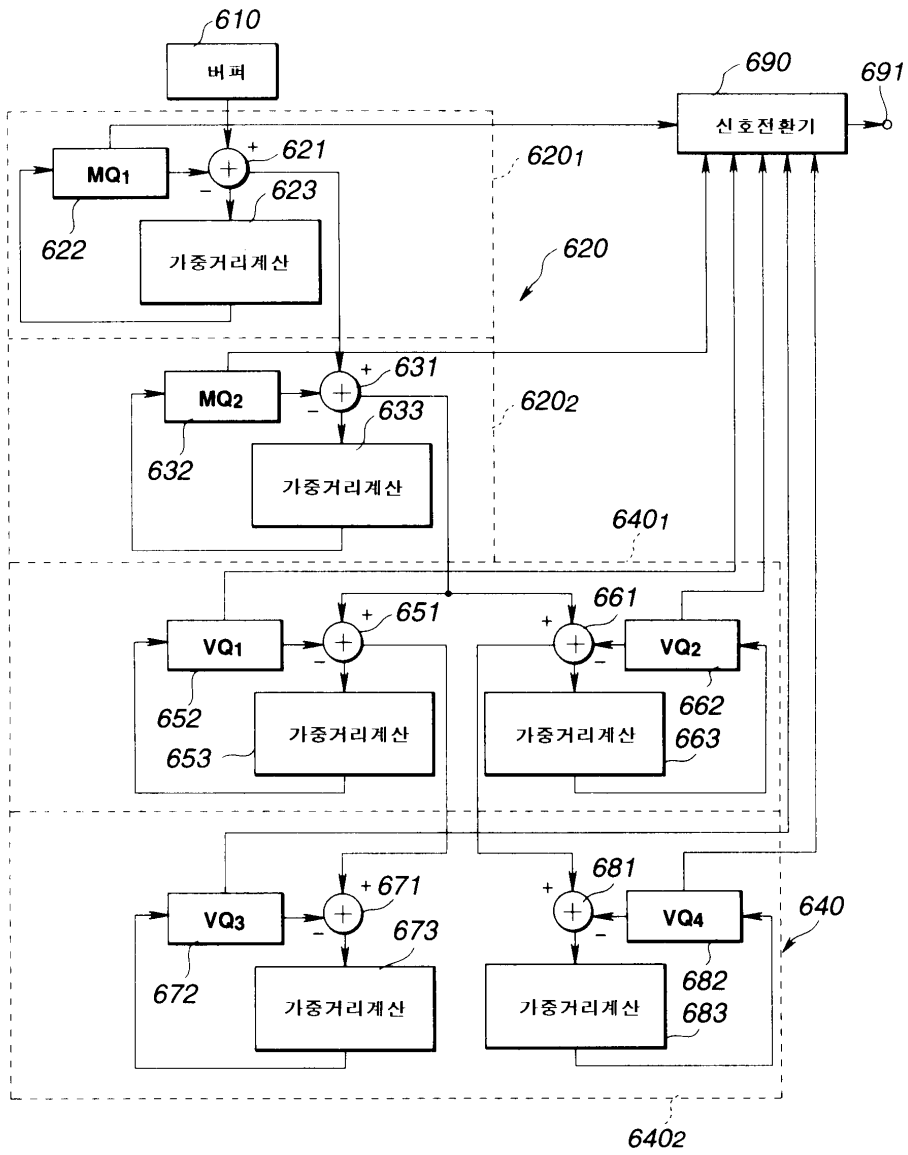
도면5

| | | | |
|--------------|--|--|-------------------------|
| | | 2Kbps | 6Kbps |
| V/UV편정출력 | 1bit / 20msec | 1bit / 20msec | |
| LSP양자화인덱스 | 32bits / 40msec | 48bits / 40msec | |
| | 피치데이터 | 피치데이터 | |
| | 8bits / 20msec | 8bits / 20msec | |
| | 인덱스 15bit/20msec | 인덱스 87bit/20msec | |
| 유성음(V)용 | 형상 (제1단) 이득 | 형상 (제1단) 이득 | 형상 (제1단) 이득 |
| | 5+5bits / 20msec | 5+5bits / 20msec | 5+5bits / 20msec |
| | 5bits / 20msec | 5bits / 20msec | 72bits / 20msec |
| | 인덱스 11bit/10msec | 인덱스 23bit/5msec | |
| 무성음(UV)용 | 형상 (제1단) 이득 | 형상 (제1단) 이득 | 형상 (제1단) 이득 |
| | 7bits / 10msec | 9bits / 5msec | 9bits / 5msec |
| | 4bits / 10msec | 6bits / 5msec | 6bits / 5msec |
| | | 5bits / 5msec | 5bits / 5msec |
| | | 3bits / 5msec | 3bits / 5msec |
| 유성음용 무성음용 | 40bits / 20msec 39bits / 20msec | 120bits / 20msec 117bits / 20msec | |

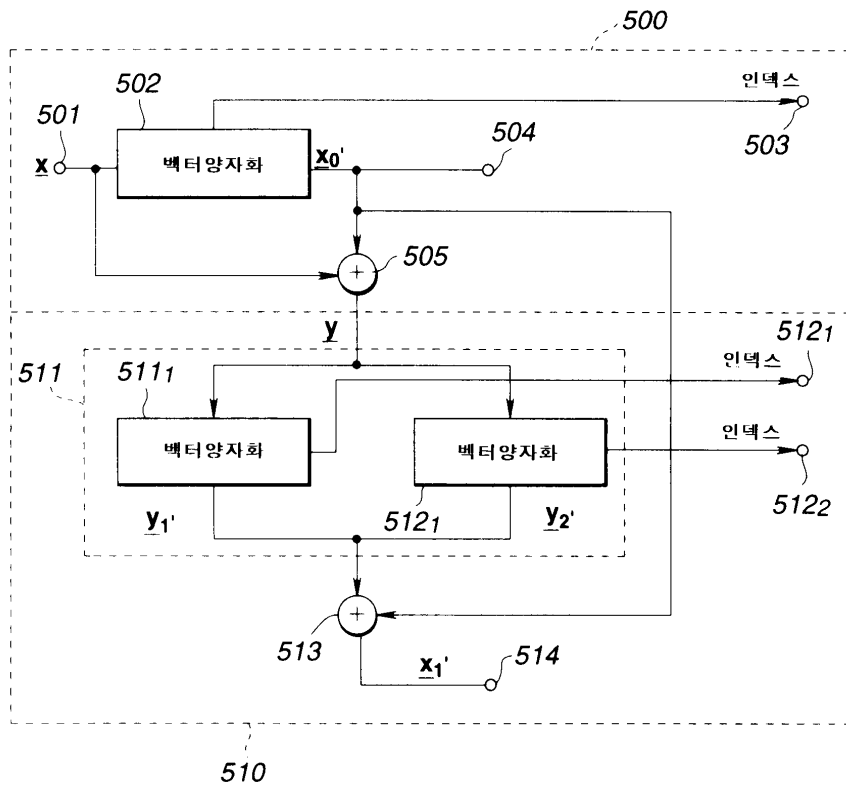
도면6



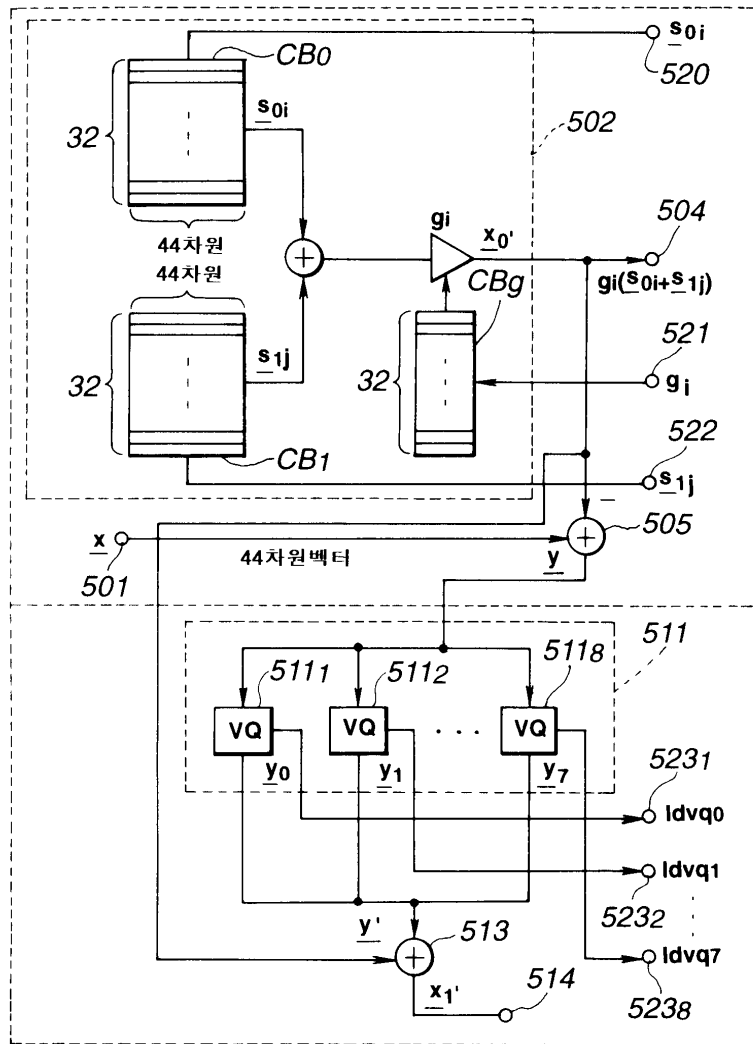
도면7



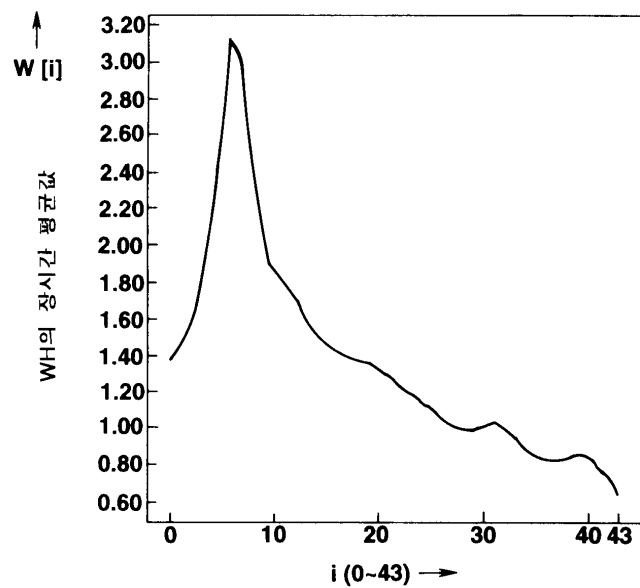
도면8



도면9



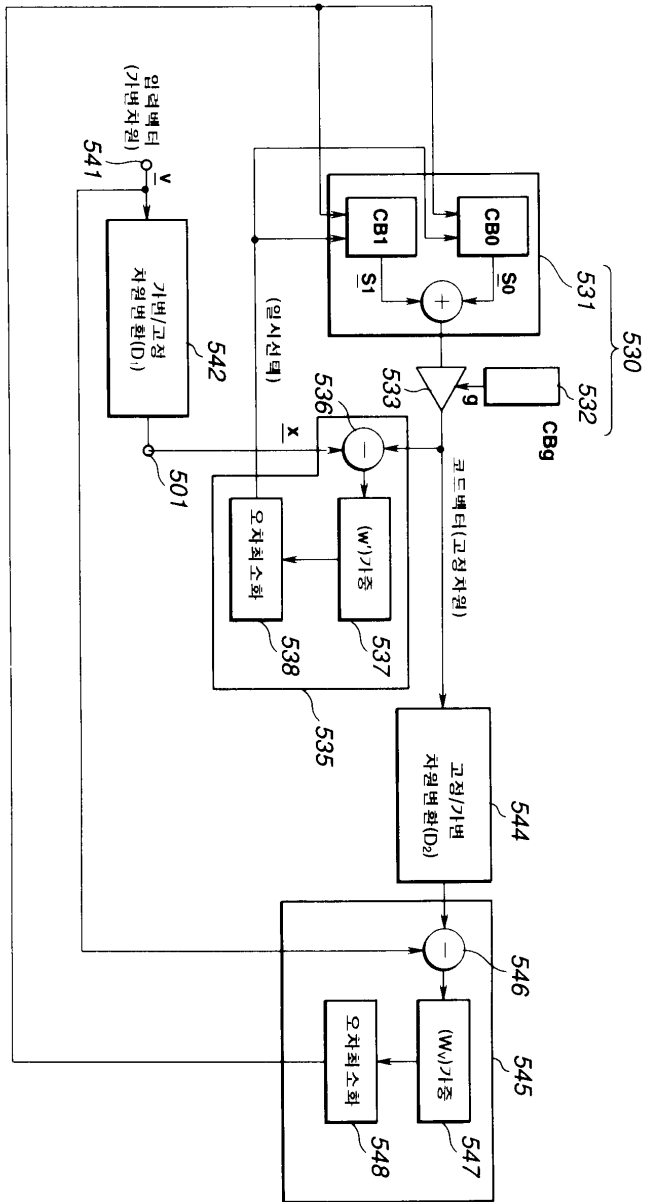
도면10



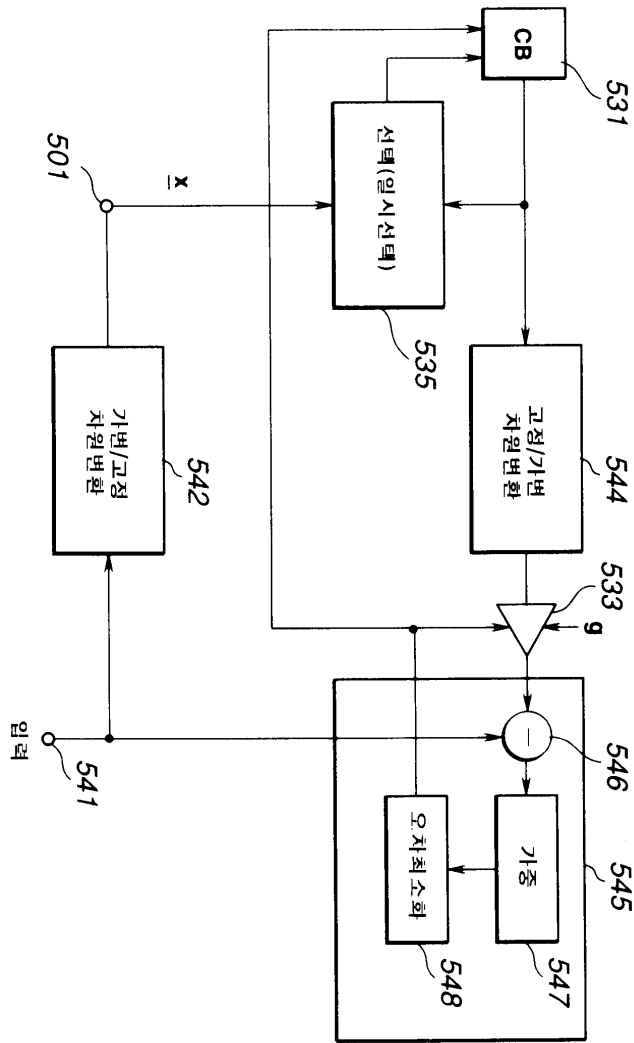
도면11

| 양자화값 | 차원의 수 | 비트수(bits) |
|-----------------------|-------|-----------|
| <u>y</u> ₀ | 4 | 10 |
| <u>y</u> ₁ | 4 | 10 |
| <u>y</u> ₂ | 4 | 10 |
| <u>y</u> ₃ | 4 | 10 |
| <u>y</u> ₄ | 4 | 9 |
| <u>y</u> ₅ | 8 | 8 |
| <u>y</u> ₆ | 8 | 8 |
| <u>y</u> ₇ | 8 | 7 |

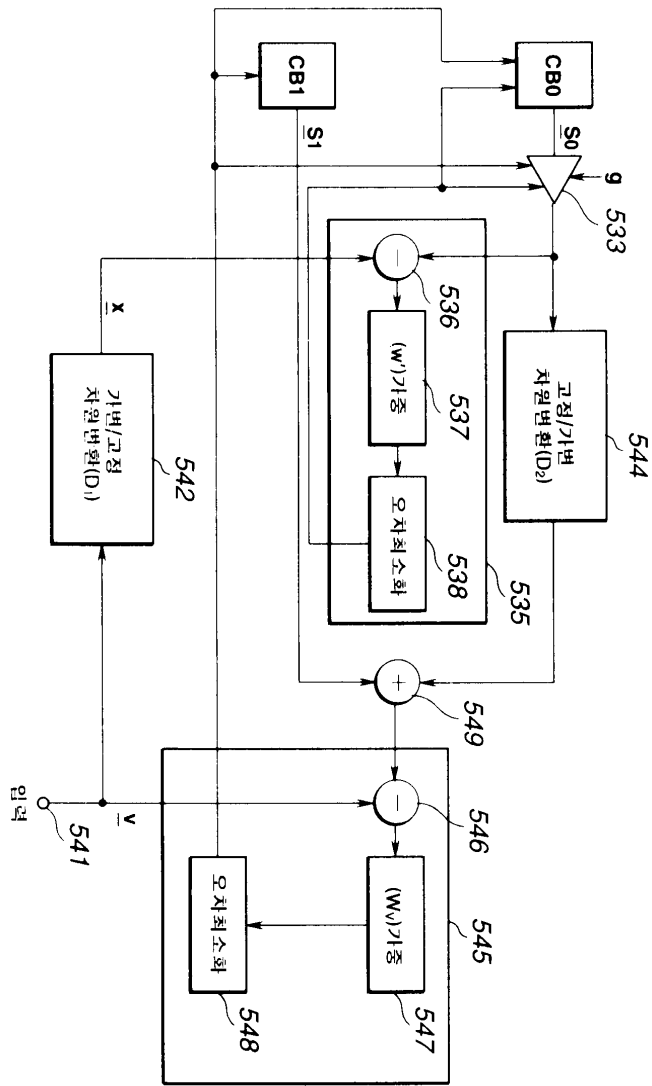
도면12



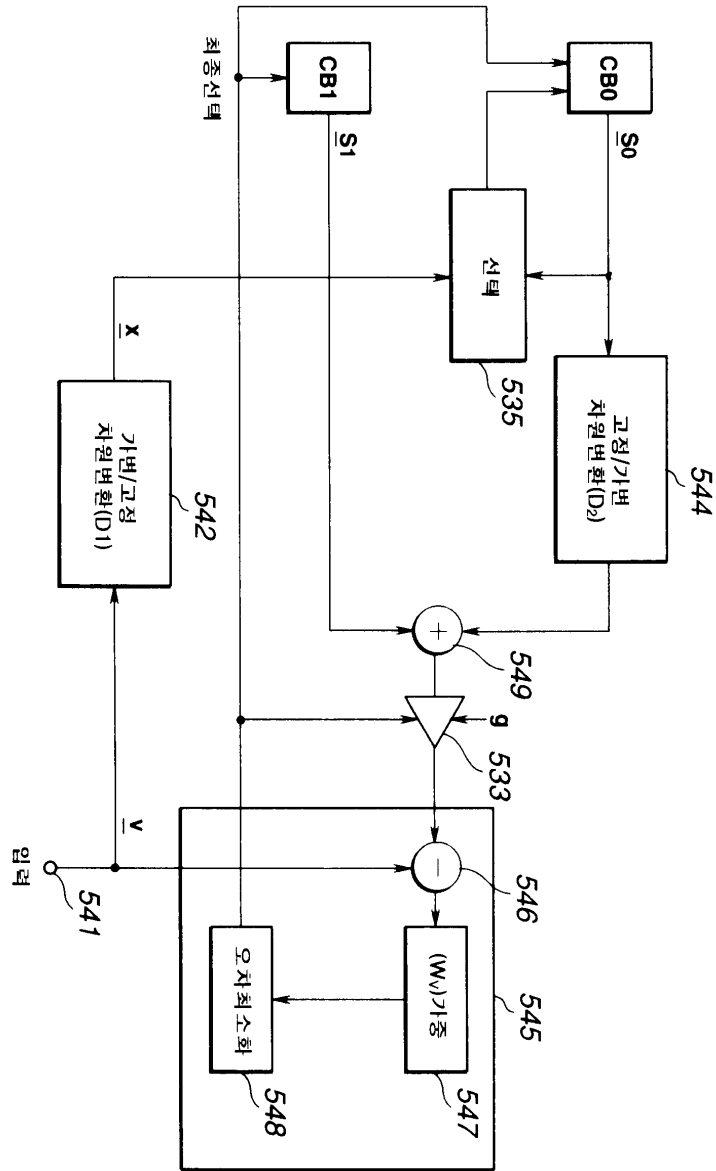
도면13



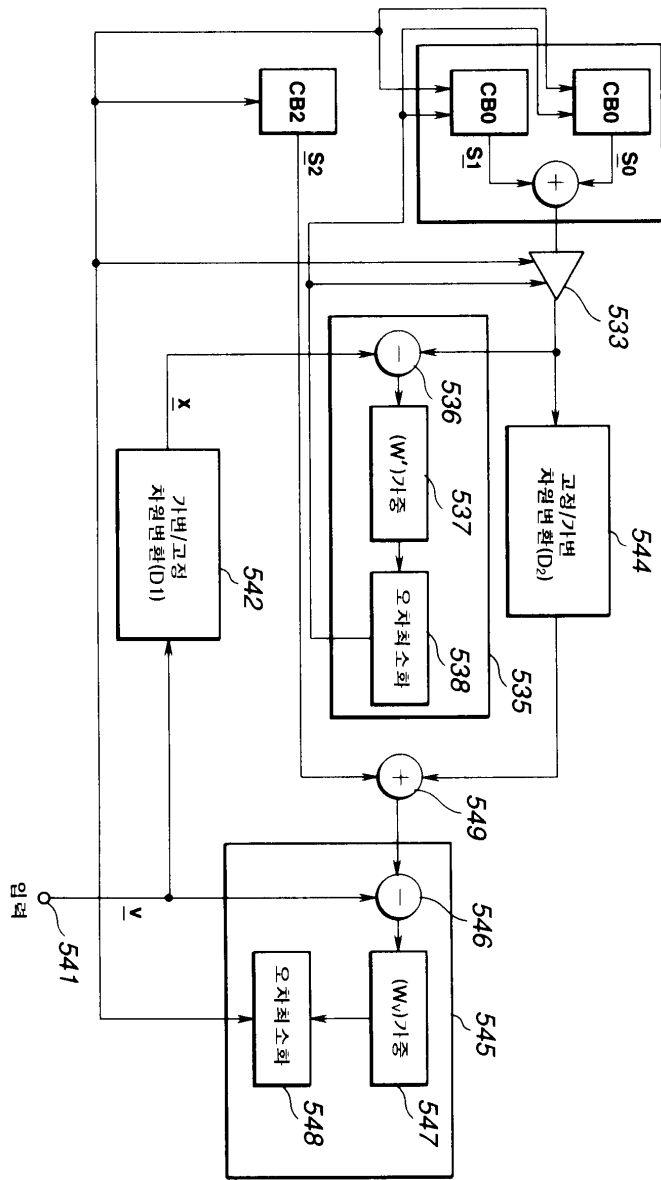
도면14



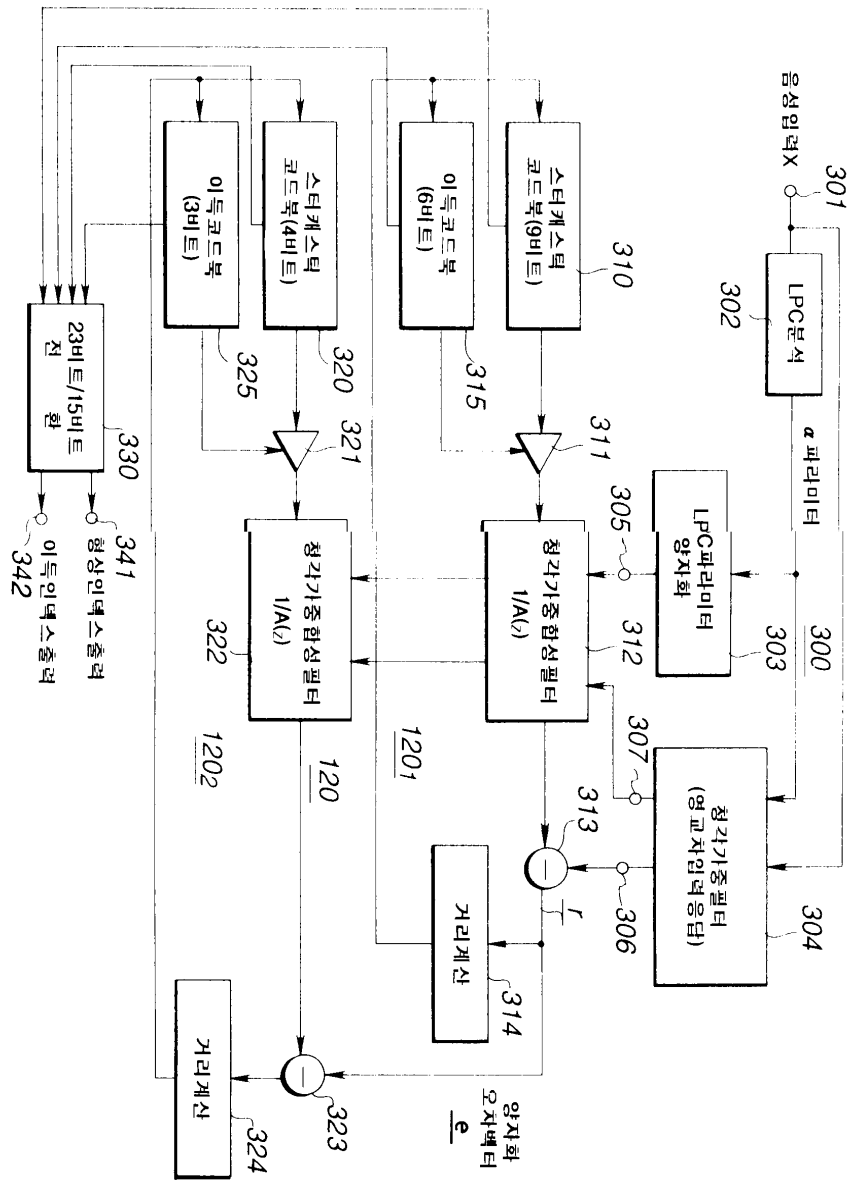
도면15



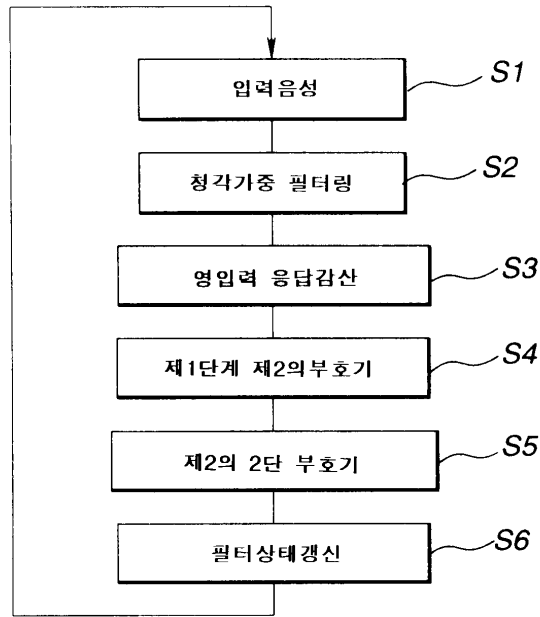
도면17



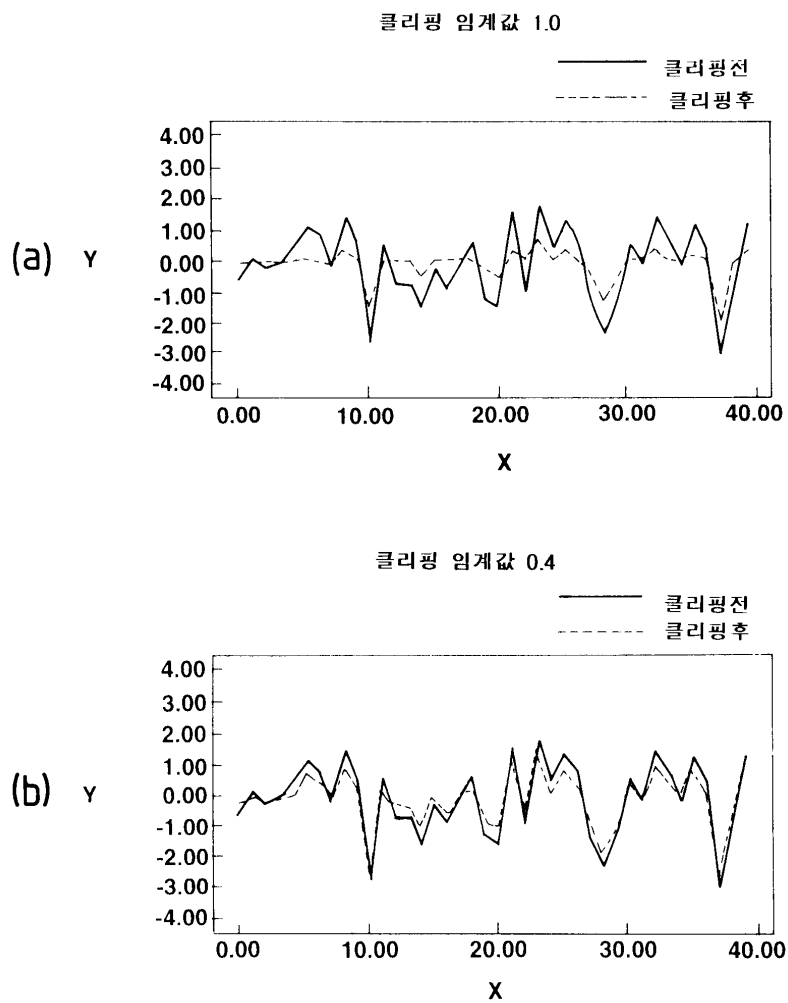
도면18



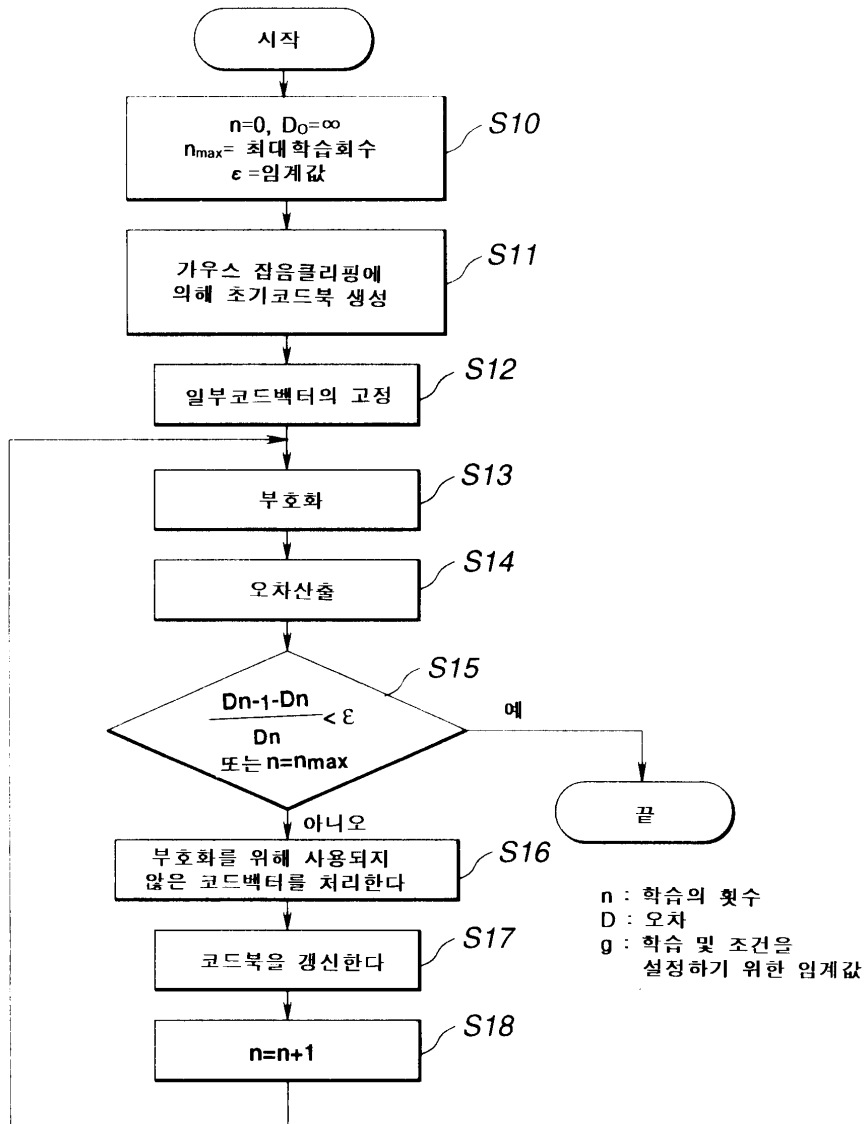
도면19



도면20



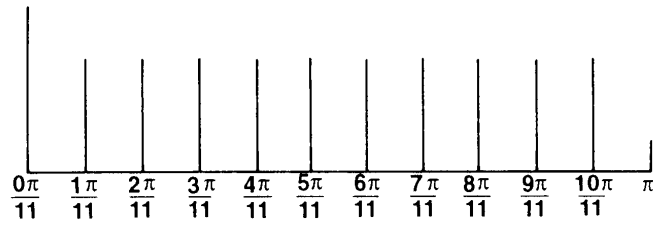
도면21



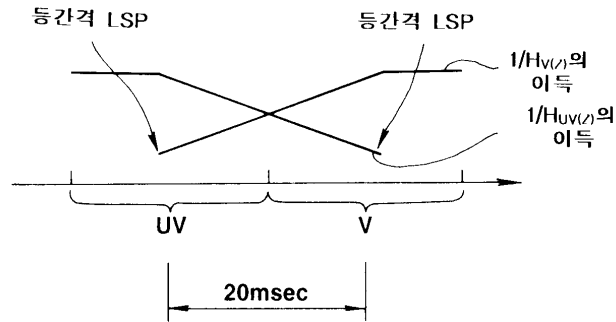
도면22

| | Hv(z) | | Huv(z) | |
|---------|--------|--------|--------|--------|
| | 이전 프레임 | 현재 프레임 | 이전 프레임 | 현재 프레임 |
| V → V | 송신 LSP | 송신 LSP | 등간격LSP | 등간격LSP |
| V → UV | 송신 LSP | 등간격LSP | 등간격LSP | 송신 LSP |
| UV → V | 등간격LSP | 송신 LSP | 송신 LSP | 등간격LSP |
| UV → UV | 등간격LSP | 등간격LSP | 송신 LSP | 송신 LSP |

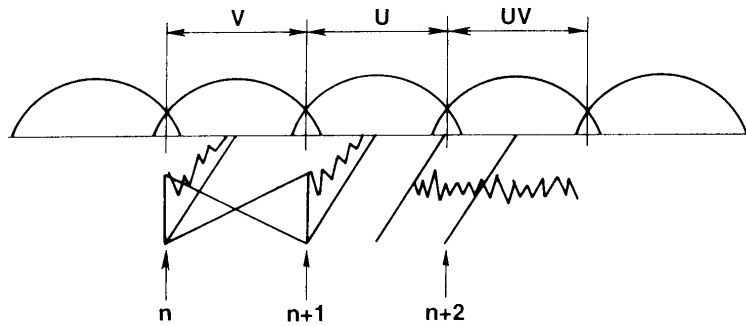
도면23



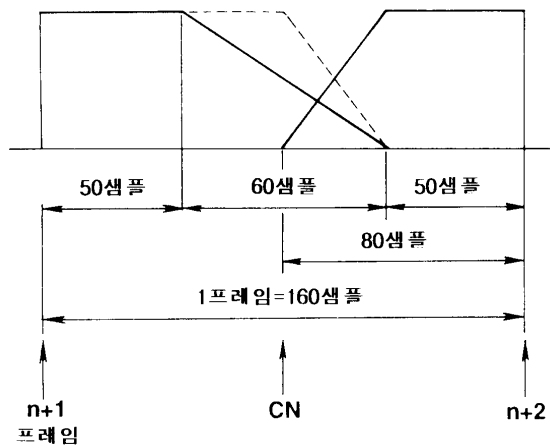
도면24



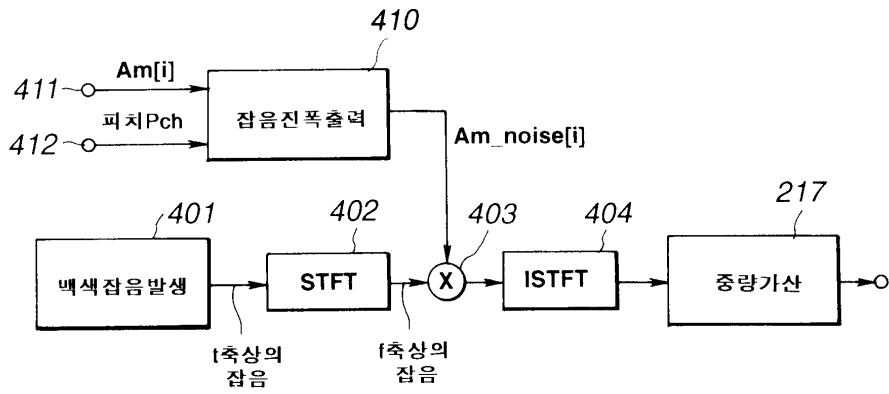
도면25



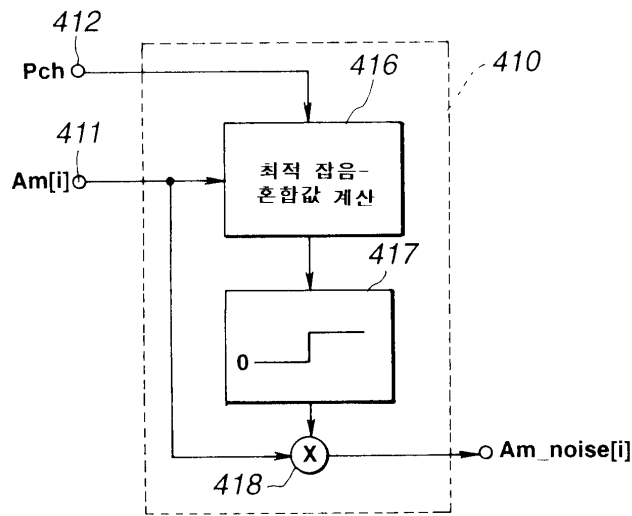
도면26



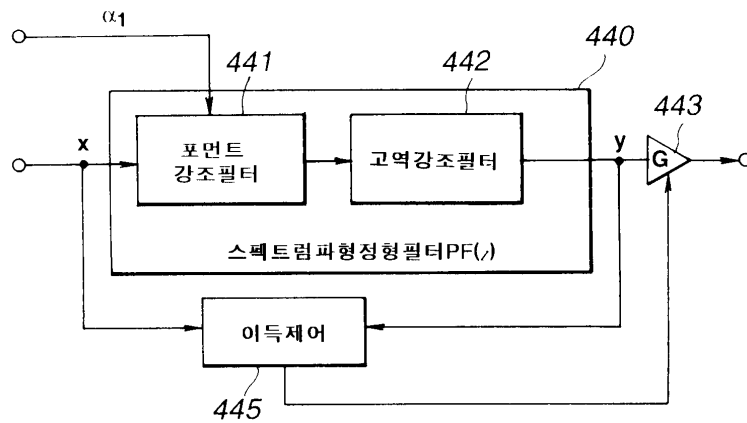
도면27



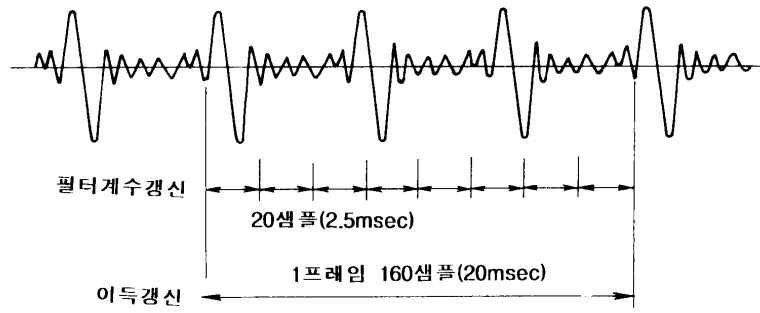
도면28



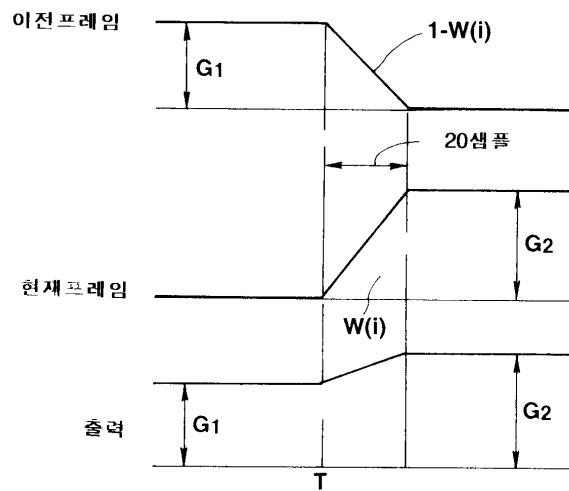
도면29



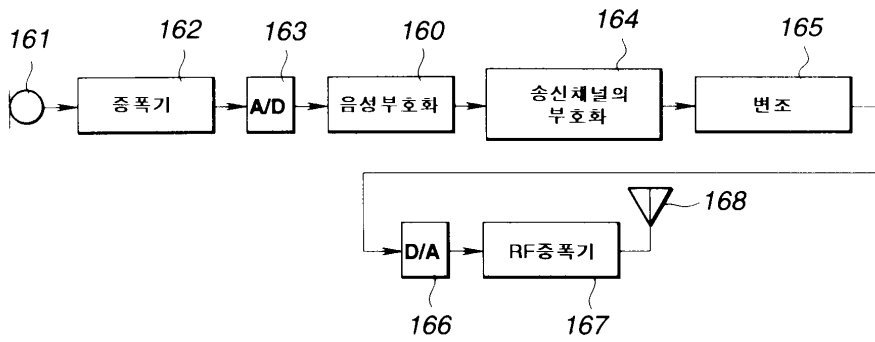
도면30



도면31



도면32



도면33

