(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0008966
(43) 공개일자 2011년01월27일

(51) Int. Cl.
G10L 19/00 (2006.01) *H04N 7/24* (2011.01)
G11B 20/10 (2006.01)
(21) 출원번호 10-2009-0066537
(22) 출원일자 2009년07월21일
심사청구일자 2009년07월21일

(71) 출원인
세종대학교산학협력단
서울 광진구 군자동 98 세종대학교
(72) 발명자
김무영
서울특별시 강남구 도곡2동 도곡렉슬아파트 210동 402호
이윤주
서울특별시 성북구 석관1동 두산아파트 121동 2102호
(74) 대리인
김재홍

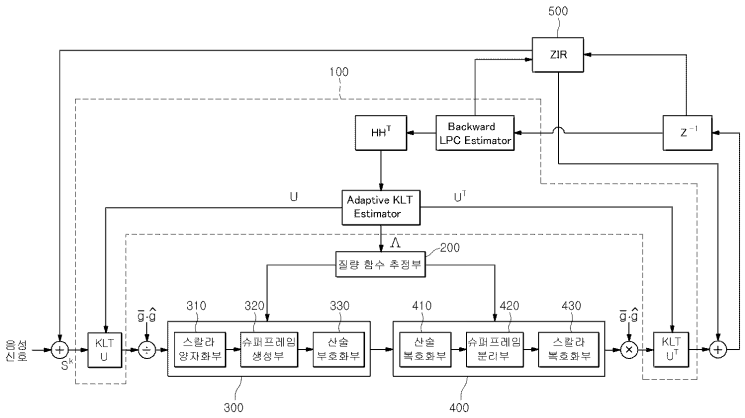
전체 청구항 수 : 총 21 항

(54) 음성신호 양자화 장치 및 방법

(57) 요 약

음성신호에 대한 양자화 장치는 (a) 입력되는 음성신호를 KLT 도메인으로 변환하는 KLT부, (b) 상기 KLT 도메인으로 변환된 음성신호를 양자화하여 상기 음성신호에 대응하는 인덱스를 포함하는 프레임을 생성하고, 상기 음성신호를 코드 벡터로 변환하는 스칼라 양자화부, (c) 상기 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여, 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어 변환하고, 상기 부호어에 대한 인덱스 정보를 전송하는 산술 부호화부, (d) 상기 산술 부호화부로부터 수신한 상기 인덱스 정보에 대응하는 부호어를 상기 인덱스를 포함하는 프레임으로 역변환하는 산술 복호화부 및 (e) 상기 역변환된 프레임에 포함되어 있는 상기 인덱스에 기초하여 이산화된 코드 벡터를 출력하는 스칼라 복호화부를 포함하되, 상기 스칼라 양자화부에 입력되는 상기 KLT 도메인으로 변환된 음성신호는 역방향 적응 이득 및 양자화 이득에 의해 정규화된다.

대 표 도



특허청구의 범위

청구항 1

음성신호에 대한 양자화 장치에 있어서,

(a) 입력되는 음성신호를 KLT 도메인으로 변환하는 KLT부,

(b) 상기 KLT 도메인으로 변환된 음성신호를 양자화하여 상기 음성신호에 대응하는 인덱스를 포함하는 프레임을 생성하고, 상기 음성신호를 코드 벡터로 변환하는 스칼라 양자화부,

(c) 상기 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여, 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어 변환하고, 상기 부호어에 대한 인덱스 정보를 전송하는 산술 부호화부,

(d) 상기 산술 부호화부로부터 수신한 상기 인덱스 정보에 대응하는 부호어를 상기 인덱스를 포함하는 프레임으로 역변환하는 산술 복호화부 및

(e) 상기 역변환된 프레임에 포함되어 있는 상기 인덱스에 기초하여 이산화된 코드 벡터를 출력하는 스칼라 복호화부

를 포함하되,

상기 스칼라 양자화부에 입력되는 상기 KLT 도메인으로 변환된 음성신호는 역방향 적응 이득 및 양자화 이득에 의해 정규화된 것인 양자화 장치.

청구항 2

제 1 항에 있어서,

(f) 상기 스칼라 양자부와 상기 산술 부호화부 사이에 상기 음성신호의 차수(dimension)에 따라 상기 인덱스를 구분하고, 상기 인덱스를 포함하는 단일 프레임을 복수개 포함하는 슈퍼 프레임(superframe)을 생성하는 슈퍼프레임 생성부

를 더 포함하는 양자화 장치.

청구항 3

제 1 항 또는 제 2 항에 있어서,

상기 인덱스는 상기 KLT 도메인으로 변환된 음성신호의 이득(gain) 인덱스와 형태(shape) 인덱스를 포함하는 것인 양자화 장치.

청구항 4

제 3 항에 있어서,

(g) 상기 KLT 도메인으로 변환된 음성신호의 확률 밀도 함수(probability density function)에 기초하여, 상기 이득 인덱스와 형태 인덱스의 확률 질량 함수(probability mass function)를 측정하는 질량 함수 추정부

를 더 포함하는 것인 양자화 장치.

청구항 5

제 3 항에 있어서,

상기 KLT부는,

이전 프레임에서 양자화된 음성신호에 기초하여 현재 프레임의 선형 예측(Linear Predictive) 계수를 측정하고, 상기 선형 예측 계수에 기초하여 생성된 고유 벡터 행렬을 이용하여 상기 입력되는 음성신호를 상기 KLT 도메인으로 변환하는 것인 양자화 장치.

청구항 6

제 3 항에 있어서,

상기 KLT 도메인으로 변환된 음성신호는 가우시안(Gaussian) 확률 밀도 함수(probability density function)를 갖되, 상기 가우시안 확률 밀도 함수는 $(-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i})$ 로 범위가 제한되며,

상기 λ_i는 상기 음성신호의 i번째 고유값인 것인 양자화 장치.

청구항 7

제 3 항에 있어서,

(h) 상기 산술 복호화부와 상기 스칼라 복호화부 사이에 상기 슈퍼 프레임(superframe)을 상기 KLT 도메인으로 변환된 음성신호의 이득(gain) 인덱스와 형태(shape) 인덱스를 포함하는 복수개의 단일 프레임으로 분리하는 슈퍼 프레임 분리부

를 더 포함하는 양자화 장치.

청구항 8

제 3 항에 있어서,

상기 산술 부호화부는

상기 음성신호의 차수(dimension)에 따라 상기 슈퍼 프레임에 포함되어 있는 복수개의 단일 프레임을 선택적으로 이용하여 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어로 변환하는 것인 양자화 장치.

청구항 9

제 4 항에 있어서,

상기 산술 부호화부는

상기 확률 질량 함수에 기초하여 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어로 변환하는 것인 양자화 장치.

청구항 10

제 9 항에 있어서,

상기 산술 부호화부는

스텝 크기 △에 기초하여 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어로 변환하되,

상기 스텝 크기 △는 유효 스텝 크기 g△가 미리 설정된 최소 스텝 크기 $\triangle_{min}$ 보다 작은 경우에는 상기 최소 스텝 크기 $\triangle_{min}$를 이득으로 나눈 값이고, 상기 유효 스텝 크기 g△가 미리 설정된 최대 스텝 크기 $\triangle_{max}$보다 큰 경우에는 상기 최대 스텝 크기 $\triangle_{max}$를 상기 이득으로 나눈 값인 양자화 장치.

청구항 11

제 5 항에 있어서,

상기 KLT부는

현재 프레임의 선형 예측 계수에 기초하여 고유값과 고유 벡터 행렬을 각각 검출하고, 상기 고유값과 고유 벡터 행렬을 이용하여 상기 이산화된 코드 벡터를 역KLT 변환하는 것인 양자화 장치.

청구항 12

음성신호에 대한 양자화 방법에 있어서,

a) 입력되는 음성신호를 KLT 도메인으로 변환하는 단계,

b) 상기 KLT 도메인 변환된 음성신호를 역방향 적응 이득 및 양자화 이득에 의해 정규화시키는 단계,

c) 상기 정규화된 음성신호를 양자화하여 상기 음성신호에 대응하는 인덱스를 포함하는 프레임을 생성하고, 상기 음성신호를 코드 벡터로 변환하는 단계,

d) 상기 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여, 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어로 변환하고, 상기 부호어에 대한 인덱스 정보를 전송하는 단계,

e)상기 수신한 인덱스 정보에 대응하는 부호어를 상기 인덱스를 포함하는 프레임으로 역변환하는 단계 및

f) 상기 역변환된 프레임에 포함되어 있는 상기 인덱스에 기초하여 이산화된 코드 벡터를 출력하는 단계

를 포함하는 양자화 방법.

청구항 13

제 10 항에 있어서,

g) 상기 단계 c) 와 상기 단계 d) 사이에서 상기 음성신호의 차수(dimension)에 따라 상기 인덱스를 구분하고, 상기 인덱스를 포함하는 단일 프레임을 복수개 포함하는 슈퍼 프레임(superframe)을 생성하는 단계

를 더 포함하는 양자화 방법.

청구항 14

제 12 항 또는 제 13 항에 있어서,

상기 인덱스는 상기 KLT 도메인으로 변환된 음성신호의 이득(gain) 인덱스와 형태(shape) 인덱스를 포함하는 것인 양자화 방법.

청구항 15

제 14 항에 있어서,

h) 상기 KLT 도메인으로 변환된 음성신호의 확률 밀도 함수(probability density function)에 기초하여, 상기 이득 인덱스와 형태 인덱스의 확률 질량 함수(probability mass function)를 측정하는 단계

를 더 포함하는 것인 양자화 방법.

청구항 16

제 15 항에 있어서,

상기 d) 단계는

상기 확률 질량 함수에 기초하여 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어로 변환하는 것인 양자화 방법.

청구항 17

제 14 항에 있어서,

상기 KLT 도메인으로 변환된 음성신호는 가우시안(Gaussian) 확률 밀도 함수(probability density function)를 갖되, 상기 가우시안 확률 밀도 함수는 $(-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i})$ 로 범위가 제한되며,

상기 λ_i는 상기 음성신호의 i번째 고유값인 것인 양자화 방법.

청구항 18

제 16 항에 있어서,

상기 d) 단계는

스텝 크기 △에 기초하여 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어로 변환하되,

상기 스텝 크기 △는 유효 스텝 크기 g△가 미리 설정된 최소 스텝 크기 $\triangle_{min}$ 보다 작은 경우에는 상기 최소 스텝 크기 $\triangle_{min}$를 이득으로 나눈 값이고, 상기 유효 스텝 크기 g△가 미리 설정된 최대 스텝 크기 $\triangle_{max}$보다 큰 경우에는 상기 최대 스텝 크기 $\triangle_{max}$를 상기 이득으로 나눈 값인 양자화 방법.

청구항 19

제 14 항에 있어서,

상기 a) 단계는

이전 프레임에서 양자화된 음성신호에 기초하여 현재 프레임의 선형 예측(Linear Predictive) 계수를 측정하고, 상기 선형 예측 계수에 기초하여 생성된 고유 벡터 행렬을 이용하여 상기 입력되는 음성신호를 상기 KLT 도메인으로 변환하는 것인 양자화 방법.

청구항 20

제 14 항에 있어서,

i) 상기 e) 단계와 상기 f) 단계 사이에서 상기 슈퍼 프레임(superframe)을 상기 KLT 도메인으로 변환된 음성신호의 이득(gain) 인덱스와 형태(shape) 인덱스를 포함하는 복수개의 단일 프레임으로 분리하는 단계

를 더 포함하는 양자화 방법.

청구항 21

제 19 항에 있어서,

j) 현재 프레임의 선형 예측 계수에 기초하여 고유값과 고유 벡터 행렬을 각각 검출하고, 상기 고유값과 상기 고유 벡터 행렬을 이용하여 상기 이산화된 코드 벡터를 역KLT 변환하는 단계

를 더 포함하는 양자화 방법.

명 세 서

발명의 상세한 설명

기 술 분 야

[0001] 본 발명은 음성신호 양자화 장치 및 방법에 관한 것으로서, 보다 상세하게는, 상기 음성신호의 이득(gain) 인덱스와 형태(shape) 인덱스를 포함하는 프레임을 생성하고, 상기 프레임에 포함되어 있는 인덱스의 발생 확률 분포 및 누적 확률 분포에 기초하여, 상기 인덱스를 비트열(bit stream)로 변환함으로써, 음성신호에 대한 높은 부호화 효율을 제공할 수 있는 엔트로피 제한적 양자화 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 용량이 큰 멀티미디어 데이터의 효율적인 압축을 위해, 해상도 제한적 양자화(Resolution-Constrained Quantization, 이하 RCQ라고 함)와 엔트로피 제한적 양자화(Entropy-Constrained Quantization, 이하 ECQ라고 함)의 두 가지의 주요한 소스 부호화 패러다임(paradigm)이 사용되어 왔다. RCQ는 주어진 고정 비트율(bit-rate) 제한 하에서, 평균 왜곡(mean distortion)을 최소화하는 것을 목적으로 한다. 이때, 상기 평균 왜곡은 충분히 최소화 될 수 있더라도, 어떤 경우에 있어서는 왜곡 이상값(outliers)을 야기하는 대형의 Voronoi 셀(cell)들이 생기는 문제가 있다.

[0003] 또한, RCQ 소스 부호화 패러다임의 일종으로서 KLT(Karhunen-Loeve Transform, 이하 KLT라고 함) 기반의 분류 벡터 양자화(Classified Vector Quantization, 이하 KLT-CVQ라고 함) 방법이 있다. 상기 KLT 기반의 CVQ는 벡터 양자화의 장점을 충분히 활용하여 멀티미디어 데이터를 효율적으로 압축할 수 있는 장점이 있지만, RCQ와 마찬가지로 대형의 Voronoi 셀(cell)에 의한 왜곡이 발생하는 문제가 있다.

[0004] 또한, KLT(Karhunen-Loeve Transform)를 엔트로피 제한적 양자화에 적용할 경우, 입력되는 음성신호의 차원(dimension)이 높으면, 벡터 양자화의 이점을 충분히 이용할 수 없는 문제점이 있었다. 예를 들어, 벡터 양자

화의 이점은 메모리, 스페이스 필링(space-filling) 및 형태(shape) 상의 이점으로 구성될 수 있다. 메모리 이점은 샘플(sample)들 간의 의존성을 나타내고, 스페이스 필링(space-filling)은 Voronoi 셀에서의 비활동(inertia)의 순간에 의해 야기된다. 또한, 형태(shape) 상의 이점은 스칼라 양자화로부터 얻어지는 k 차원(dimensional) 무게 중심(centriod) 밀도(density)가 최적화된 것이 아니라는 사실과 관계된다.

[0005] 또한, 산술 부호화(arithmetic coding)를 이용하여 입력되는 음성신호를 부호화하는 경우, 입력되는 음성신호의 차원이 낮으면, 다른 종류의 무손실 부호화(lossless coding)를 이용하여 음성신호를 부호화하는 경우보다 부호화 효율이 떨어지는 문제가 있었다.

발명의 내용

해결 하고자하는 과제

[0006] 본 발명의 일부 실시예는 음성신호의 이득 인덱스와 형태 인덱스를 포함하는 프레임을 생성하고, 산술 부호화를 이용하여 상기 프레임에 포함되어 있는 인덱스를 비트열로 변환함으로써, 음성신호에 대한 높은 부호화 효율을 제공할 수 있는 음성신호 양자화 장치 및 방법을 제공한다.

과제 해결수단

[0007] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 발명의 제 1 측면은 (a) 입력되는 음성신호를 KLT 도메인으로 변환하는 KLT부, (b) 상기 KLT 도메인으로 변환된 음성신호를 양자화하여 상기 음성신호에 대응하는 인덱스를 포함하는 프레임을 생성하고, 상기 음성신호를 코드 벡터로 변환하는 스칼라 양자화부, (c) 상기 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여, 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어 변환하고, 상기 부호어에 대한 인덱스 정보를 전송하는 산술 부호화부, (d) 상기 산술 부호화부로부터 수신한 상기 인덱스 정보에 대응하는 부호어를 상기 인덱스를 포함하는 프레임으로 역변환하는 산술 복호화부 및 (e) 상기 역변환된 프레임에 포함되어 있는 상기 인덱스에 기초하여 이산화된 코드 벡터를 출력하는 스칼라 복호화부를 포함하되, 상기 스칼라 양자화부에 입력되는 상기 KLT 도메인으로 변환된 음성신호는 역방향 적응 이득 및 양자화 이득에 의해 정규화된 것인 음성신호에 대한 양자화 장치를 제공할 수 있다.

[0008] 또한, 본 발명의 제 2 측면은 a) 입력되는 음성신호를 KLT 도메인으로 변환하는 단계, b) 상기 KLT 도메인 변환된 음성신호를 역방향 적응 이득 및 양자화 이득에 의해 정규화시키는 단계, c) 상기 정규화된 음성신호를 양자화하여 상기 음성신호에 대응하는 인덱스를 포함하는 프레임을 생성하고, 상기 음성신호를 코드 벡터로 변환하는 단계, d) 상기 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여, 상기 코드 벡터를 비트열(bit stream)로 이루어진 부호어로 변환하고, 상기 부호어에 대한 인덱스 정보를 전송하는 단계, e)상기 수신한 인덱스 정보에 대응하는 부호어를 상기 인덱스를 포함하는 프레임으로 역변환하는 단계 및 f) 상기 역변환된 프레임에 포함되어 있는 상기 인덱스에 기초하여 이산화된 코드 벡터를 출력하는 단계를 포함하는 음성신호에 대한 양자화 방법을 제공할 수 있다.

효 과

[0009] 전술한 본 발명의 과제 해결 수단에 의하면, 입력되는 음성신호의 차수에 따라 음성신호를 부호화 할 수 있다.

[0010] 또한, 전술한 본 발명의 과제 해결 수단에 의하면, 종래의 분류 벡터 양자화(CVQ) 장치 및 방법보다 향상된 SNR 및 R-D(Rate-Distortion) 특성을 나타낼 수 있다.

[0011] 또한, 전술한 본 발명의 과제 해결 수단에 의하면, 주관적 음질 평가인 음성 질 지각 평가(perceptual evaluation of speech quality, PESQ) 수치도 향상될 수 있다.

[0012] 또한, 본 발명의 또 다른 과제 해결 수단 중 하나에 의하면, 역적응 방법(backward adaptive manner)으로 선형 예측 계수를 추정하므로, 부호화기에서 복호화기로 LP 계수를 전송할 필요가 없다.

발명의 실시를 위한 구체적인 내용

[0013] 아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0014] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0015] 이하 첨부된 도면을 참고하여 본 발명을 상세히 설명하기로 한다.

[0016] 도 1은 본 발명의 일 실시예에 따른 음성신호에 대한 양자화 장치의 구성도이다.

[0017] 도 1에 도시된 바와 같이, 본 발명에 따른 음성신호에 대한 양자화 장치는 KLT부(100), 질량 함수 추정부(200), 부호기측 양자화부(300) 및 복호기측 양자화부(400)을 포함한다.

[0018] 도 1에는 도시되지 않았지만, KLT 기반의 분할 벡터 양자화 장치는 코드북 데이터베이스를 구비할 수 있다.

[0019] 코드북 데이터베이스는 KLT에 의해 얻어진 음성신호에 대한 코드 벡터를 KLT 도메인 통계치에 따라 복수개로 분류하여 구비하고 있다.

[0020] 코드북 데이터베이스는 KLT의 에너지 집중 특성(energy concentration property)을 이용하여 음성신호(speech signal)에 대한 KLT 도메인 통계치(statistics)가 좁은 클래스(narrow class)별로 이득(gain) 값을 갖는 코드북이 분류되게 설계(design)될 수 있다.

[0021] 도 2는 각 상관계수(correlation coefficient) a1에 대한 2차원 음성 신호에 대한 코드 벡터들의 분포도를 나타낸 것이다. 도 3은 도 2에서와 같은 상관계수 a_1에서 상기 2차원 음성 신호에 상당하는 KLT 변환된 신호에 대한 코드 벡터들의 분포도를 나타낸 것이다. 즉, 음성신호(speech signal)를 KLT 도메인(domain)으로 변환할 경우에, 도 3에 도시된 바와 같이 수평축(horizontal axis)을 따라 에너지가 집중된 도메인을 얻게 된다. 도 2에서 다른 통계치를 갖는 음성 신호들이 도 3의 KLT 도메인상에서 동일한 통계를 갖는다는 것을 알 수 있다. KLT 도메인 상에서 동일한 통계를 갖는다는 것은 동일한 고유값 세트(eigenvalue set)로 분류될 수 있다는 것을 의미한다.

[0022] 다시 도 1로 돌아와서 설명하면, KLT부(100)는 입력되는 음성신호를 KLT 도메인으로 변환하고, 이를 다시 역 KLT 변환할 수 있다. 상기 KLT 변환은 프레임 단위로 이루어진다. 이를 위하여 KLT부(100)는 먼저, 이전 프레임에서 양자화된 음성신호에 기초하여 현재 프레임의 선형 예측 계수를 측정할 수 있다.

[0023] 만약, 선형 예측 계수 분석이 원 신호 도메인(original signal domain)에서 수행되는 경우, LP 계수를 양자화하여 그것의 인덱스(index)를 다시 복호기측 양자화부(400)로 전송하여야만 한다. 이러한 부가 정보의 전송을 제거하기 위해, 본 실시예에서는 역 적응 방법(backward adaptive manner)을 이용하여, 선형 예측 계수 분석의 입력으로서 이전 프레임(frame)에서 합성된 신호를 이용한다.

[0024] 본 발명에 따른 양자화 장치는 부호기측 양자화부(300)와 복호기측 양자화부(400)를 포함하고, 영입력 응답부(500)에서는 LPC 필터(filter)의 동일한 제로 입력 응답(zero-input response, 이하 ZIR이라 함)이 얻어지기 때문에, KLT부(100)는 오직 제로 상태 응답(zero-state response, 이하 ZSR이라 함), 즉, s^k를 고유 벡터 행렬 U를 이용하여 KLT 도메인으로 변환한다.

[0025] 또한, KLT부(100)는 구해진 선형 예측 계수의 임펄스 응답 H를 이용하여 입력신호의 공분산(covariance) 행렬 $\sum=HH^T$를 구한다. 예를 들어, 5차원 (dimensional)인 경우에, 공분산 행렬 $\sum$는 하기 수학식 1과 같이 정의된다.

수학식 1

[0026]
$$\begin{bmatrix} 1 & A_1 & A_2 & A_3 & A_4 \\ A_1 & 1+A_1^2 & A_1+A_1A_2 & A_2+A_1A_3 & A_3+A_1A_4 \\ A_2 & A_1+A_1A_2 & 1+A_1^2+A_2^2 & A_1+A_1A_2+A_2A_3 & A_2+A_1A_3+A_2A_4 \\ A_3 & A_2+A_1A_3 & A_1+A_1A_2+A_2A_3 & 1+A_1^2+A_2^2+A_3^2 & A_1+A_1A_2+A_2A_3+A_3A_4 \\ A_4 & A_3+A_1A_4 & A_2+A_1A_3+A_2A_4 & A_1+A_1A_2+A_2A_3+A_3A_4 & 1+A_1^2+A_2^2+A_3^2+A_4^2 \end{bmatrix}$$

[0027] 상기 수학식 1에서, $A_1 = a_1$ 이고, $A_2 = a_1^2 + a_2$ 이고, $A_3 = a_1^3 + 2a_1a_2 + a_3$ 이고, $A_4 =$

$a_1^4+3a_1^2a_2+2a_1a_3+a_2^2+a_4$ 이다. a_1내지 a_4는 선형 예측 계수이다. 따라서, 선형 예측 계수를 이용하여 공분산 행렬 ∑를 구할 수 있다.

[0028] 또한, KLT부(100)는 공분산 행렬 ∑에 대한 고유값(eigenvalue) λ_i을 수학식 2를 이용하여 산출할 수 있다.

수학식 2

[0029] $\det(\sum-\lambda_i I) = 0$

[0030] 또한, KLT부(100)는 고유값에 기초하여 수학식 3에 의하여 고유 벡터를 구할 수 있다.

수학식 3

[0031] $(\sum- \lambda_i I)P_i = 0$

[0032] 상기 수학식 2 및 3에서 I는 단위(identity) 행렬이다. 이 행렬은 대각선 행렬 값이 모두 1이고, 나머지는 모두 0인 행렬이다. 수학식 3을 만족하는 고유 벡터는 정규화(normalization)하여 구할 수 있다.

[0033] 행렬 Λ의 대각선(diagonal)을 따라 공분산 행렬 ∑의 정돈된 고유값 λ_i들을 배열함으로써 행렬 Λ를 얻을 수 있다.

[0034] 한편, KLT부(100)는 구해진 고유벡터를 이용하여 고유 벡터 행렬 U를 수학식 4와 같이 구할 수 있다.

수학식 4

[0035] $U=[P_1 , P_2 , \cdots P_n]$

[0036] 상기 수학식 4에서 P_1 , P_2 , P_n은 n x 1 행렬이다.

[0037] KLT부(100)는 U를 입력되는 음성신호 s^k에 승산하여 입력된 음성신호를 KLT 도메인으로 변환한다. 여기서 s^k는 선형 예측 합성 필터의 k차 제로 입력 응답(k-dimensional zero input response)이다.

[0038] 질량 함수 추정부(200)는 공분산 행렬을 통해 구한 고유값 행렬 Λ에 기초하여 부호기측 양자화부(300)와 복호기측 양자화부(400)에서의 시변 확률 질량 함수(time-varying probability mass function, PMF)를 측정할 수 있다.

[0039] 또한, 질량 함수 추정부(200)는 이득(gain) 인덱스와 형태(shape) 인덱스의 확률 질량 함수를 측정할 수 있다. 이때, 상기 이득 인덱스의 확률 질량 함수는 훈련 데이터베이스(training database)에 기초하여 측정될 수 있고, 상기 형태 인덱스의 확률 질량 함수는 시변 시험 소스 통계값(time varying test source statistics)에 기초하여 적응적으로 측정될 수 있다.

[0040] 또한, 질량 함수 추정부(200)는 상기 KLT부(100)가 생성한 행렬 Λ에 기초하여 최적의 코드북을 산술 부호화를 위해 생성할 수 있다.

[0041] 부호기측 양자화부(300)은 스칼라 양자화부(310), 슈퍼프레임 생성부(320) 및 산술 부호화부(330)을 포함한다.

[0042] 스칼라 양자화부(310)는 상기 정규화된 KLT 도메인 음성신호를 양자화한다. 즉, 스칼라 양자화부(310)는 KLT부(200)에 의해 KLT 도메인으로 변환된 음성신호의 이득(gain) 인덱스(ϕ^g)와 형태(shape) 인덱스 ($\phi_1^s, \cdots, \phi_k^s$)를 포함하는 프레임 Φ를 생성한다.

[0043] 본 발명의 일 실시예에서는 사각 절단면(rectangular truncation)을 갖는 z-격자(z-lattice)와 동일한 특성을 갖는 스칼라 양자화(scalar quantization) 장치를 선택하였지만, 반드시 이에 한정되는 것은 아니다. 예를 들어, 고율(high-rate) 이론하에서 최적화된 엔트로피 제한적 양자화 장치를 디자인 하기 위해서는, 산술 부호화(arithmetic coding)과 같은 무손실 부호화(lossless coding)를 수반하는 격자 벡터 양자화(lattice vector quantization) 장치가 사용될 수도 있다.

[0044] 한편, 역방향 적응 이득 $\tilde{g}$ 가 n차(예를 들어, 10차) 자기 회귀(auto regressive) 모델에 의해 추정된다. KLT

도메인의 입력신호가 상기 역방향 적응 이득 $\tilde{g}$에 의해 정규화되고, 다시 양자화 이득 $\hat{g}$에 의해 한번 더 정규화 된다.

[0045] 슈퍼프레임 생성부(320)는 KLT 도메인으로 변환된 음성신호의 이득(gain) 인덱스(ϕ^g)와 형태(shape) 인덱스($\phi_1^s, \cdots, \phi_k^s$)를 포함하는 단일 프레임($\Phi_1=\{\phi^g, \phi_1^s, \cdots, \phi_k^s\}$)을 복수개 포함하는 슈퍼 프레임(superframe)($\Psi=\{\Phi_1, \Phi_2, \Phi_3, \cdots\}$)을 생성할 수 있다.

[0046] 즉, 프레임 내의 샘플 개수(음성 신호의 차수)를 증가시킨다. 왜냐하면, 고유값 λ_i를 이용한 역 적응 방법에서는 차수가 적을수록 정확한 인덱스를 추정할 수 있으나, 스칼라 양자화와 산술 부호화를 수행하는 순차적 산술 부호화(sequential arithmetic coding)에서는 차수(dimension)가 클수록 성능과 구현성이 향상되기 때문이다.

[0047] 산술 부호화부(330)는 KLT 도메인으로 변환된 음성신호의 차수(dimension)에 따라 상기 슈퍼 프레임에 포함되어 있는 복수개의 단일 프레임을 선택적으로 이용하고, 상기 스칼라 양자화부(310)가 변환한 코드 벡터를 상기 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여 비트열(bit stream) Ψ로 이루어진 부호어(codeword)로 변환할 수 있다.

[0048] 산술 부호화부(330)는 확률 질량 함수에 기초하여 스칼라 양자화부(310)가 생성한 코드 벡터를 부호어로 변환하는데, 각 무게중심의 확률 질량 함수를 구하기 위해서는 각 무게중심의 확률 밀도 함수와 셀의 크기 Γ가 사용될 수 있다. 이때, KLT 도메인으로 변환된 음성신호의 확률 밀도 함수는 영 평균(zero mean)과 공분산 Λ를 갖는 가우시안(Gaussian) 형태라고 가정될 수 있다.

[0049] 나아가, 산술 부호화부(330)는 최소 왜곡을 갖는 최적의 부호어에 대한 인덱스 정보를 복호기측 양자화부(400)로 전송할 수 있다. 예를 들어, 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여, 입력되는 코드 벡터에 대한 누적분포 함수의 구간을 산술적으로 계산하면, 이에 대응하는 부호어를 자동적으로 얻을 수 있으므로, 입력되는 코드 벡터의 길이가 길어져도 용이하게 부호화 할 수 있다.

[0050] 복호기측 양자화부(400)은 산술 복호화부(410), 슈퍼프레임 분리부(420) 및 스칼라 복호화부(430)를 포함한다.

[0051] 산술 복호화부(410)는 산술 부호화부(330)로부터 최적의 코드 벡터의 인덱스를 수신하고, 상기 인덱스에 대응하는 부호어, 즉 비트열 Ψ을 이득(gain) 인덱스와 형태(shape) 인덱스를 포함하는 프레임 Φ으로 역변환할 수 있다.

[0052] 슈퍼프레임 분리부(420)는 단일 프레임을 복수개 포함하고 있는 슈퍼 프레임(superframe)을 단일 프레임으로 분리할 수 있다.

[0053] 스칼라 복호화부(430)는 산술 복호화부(410)가 역변환 프레임 Φ에 포함되어 있는 이득(gain) 인덱스와 형태(shape) 인덱스에 기초하여, 이산화된 코드 벡터를 출력할 수 있다.

[0054] 이산화된 코드 벡터는 고유값과 고유 벡터 행렬을 이용하여 KLT부(100)에 의해 역KLT 변환되며, 최종 합성음은 영입력 응답부(500)의 ZIR을 더해서 얻어질 수 있다.

[0055] 본 명세서에서는 부호기측 양자화부(300)와 복호기측 양자화부(400)가 하나의 시스템에 존재하는 것으로 설명하였지만, 이에 한정되지 않으며, 부호기측 양자화부(300)와 복호기측 양자화부(400)는 별도의 시스템으로 각각 존재할 수 있다.

[0056] 도 4는 본 발명의 일 실시예에 따른 음성신호에 대한 양자화 방법의 흐름도이다.

[0057] 도 4를 참조하면, 단계(S500)은 음성신호가 입력되는지 여부를 판단하는 단계이다.

[0058] 단계(S502)는 상기 단계(S500)에서 입력된 음성신호에 대한 선형 예측 계수를 측정하는 단계이다. 선형 예측 계수는 프레임 단위로 측정될 수 있다. 이때, 역 적응 방법(backward adaptive manner)을 이용하여, 이전 프레임에서 양자화된 음성신호에 기초하여 현재 프레임의 선형 예측(Linear Predictive) 계수를 측정할 수 있다.

[0059] 단계(S504)는 상기 단계(S500)에서 입력된 음성신호의 공분산 행렬 $\sum=HH^T$ 를 상기 수학식 1을 이용하여 구하는 단계이다.

[0060] 단계(S506)은 상기 단계(S504)에서 구해진 공분산 행렬 $\Sigma=HH^T$를 이용하여 입력되는 음성신호에 대한 고유값을 구하고, 구해진 고유값으로 고유 벡터를 구하는 단계이다. 고유값 및 고유 벡터를 구하는 과정은 전술한 바와 같이, 수학식 2 및 3을 이용할 수 있다.

[0061] 단계(S508)에서는 상기 단계(S506)에서 산출한 고유값을 이용하여 행렬 Λ를 구하고, 상기 고유 벡터를 이용하여 고유 벡터 행렬을 구할 수 있다. 또한, 단계(S508)에서 고유 벡터 행렬을 이용하여 입력되는 음성신호를 KLT 도메인으로 변환할 수 있다.

[0062] 단계(S510)은 상기 단계(S508)에서 KLT 도메인으로 변환된 음성신호를 정규화하는 단계이다. 단계(S510)에서 역방향 적응 이득 $\tilde{g}$가 n차(예를 들어, 10차) 자기 회귀(auto regressive) 모델에 의해 추정되고, KLT 도메인의 입력신호가 상기 역방향 적응 이득 $\tilde{g}$에 의해 정규화될 수 있다. 나아가, 역방향 적응 이득 $\tilde{g}$에 의해 정규화된 신호는 다시 양자화 이득 $\hat{g}$에 의해 한번 더 정규화 된다.

[0063] 단계(S512)는 상기 단계(S510)에서 정규화된 KLT 도메인 음성신호를 양자화하는 단계이다. 즉, 단계(S512)에서 KLT 도메인으로 변환된 음성신호의 이득(gain) 인덱스(ϕ^g)와 형태(shape) 인덱스($\phi_1^s, \cdots, \phi_k^s$)를 포함하는 프레임 Φ를 생성할 수 있다.

[0064] 단계(S514)는 음성신호의 이득(gain) 인덱스(ϕ^g)와 형태(shape) 인덱스($\phi_1^s, \cdots, \phi_k^s$)를 포함하는 단일 프레임($\Phi_1=\{\phi^g, \phi_1^s, \cdots, \phi_k^s\}$)을 복수개 포함하는 슈퍼 프레임(superframe)($\Psi=\{\Phi_1, \Phi_2, \Phi_3, \cdots\}$)을 생성하는 단계이다.

[0065] 전술한 바와 같이, 고유값 λ_i를 이용한 역 적응 방법에서는 차수가 적을수록 정확한 인덱스를 추정할 수 있으나, 스칼라 양자화와 산술 부호화를 수행하는 순차적 산술부호화(sequential arithmetic coding)에서는 차수(dimension)가 클수록 성능과 구현성이 향상되기 때문에 프레임 내의 샘플 개수(음성 신호의 차수)를 증가시킨다.

[0066] 단계(S516)은 KLT 도메인 음성신호의 차수(dimension)에 따라, 상기 단계(S514)에서 생성된 슈퍼 프레임에 포함되어 있는 복수개의 단일 프레임을 선택적으로 이용하고, 상기 단계(S512)에서 변환된 코드 벡터의 발생 확률 분포 및 누적 확률 분포에 기초하여, 상기 코드 벡터를 비트열(bit stream) Ψ로 이루어진 부호어(codeword)로 변환하는 단계이다.

[0067] 단계(S518)은 상기 단계(S516)에서 생성된 부호어에 대한 인덱스 정보를 복호기측 양자화부(400)로 전송하는 단계이다.

[0068] 이후, 단계(S520)에서 복호기측 양자화부(400)는 상기 수신한 인덱스 정보에 기초하여 상기 인덱스에 대응하는 부호어, 즉 비트열 Ψ을 이득(gain) 인덱스와 형태(shape) 인덱스를 포함하는 프레임 Φ으로 역변환 할 수 있다.

[0069] 단계(S522)에서는 단일 프레임을 복수개 포함하고 있는 슈퍼 프레임(superframe)이 단일 프레임으로 분리될 수 있다.

[0070] 마지막으로, 단계(S524)에서는 상기 단계(S522)에서 분리된 단일 프레임을 이용하여 상기 역변환 된 프레임 Φ에 포함되어 있는 인덱스를 이용하여 이산화된 코드 벡터를 출력할 수 있다. 상기 이산화된 코드 벡터는 정규화된 것이므로 다시 복구될 수 있다. 또한, 상기 이산화된 코드 벡터는 KLT 변환되어 있으므로, KLT부(100)에 의해 역KLT 변환되어야 한다.

[0071] 본 발명의 일 실시예에 따른 양자화 장치 및 방법은, 연속적인 산술 부호화를 수반하는 엔트로피 제한적 양자화 장치이기 때문에, 추가적인 코드북을 필요로 하지 않으며, 분류 벡터 양자화(CVQ) 장치 및 방법보다 향상된 SNR 및 R-D 특성을 나타낼 수 있다.

[0072] 한편, 본 발명의 일 실시예에 따른 양자화 장치 및 방법은 전술한 바와 같이, KLT 도메인으로 변환된 음성신호가 영 평균(zero mean)과 공분산 Λ를 갖는 가우시안(Gaussian) 형태라고 가정하였다. 이때, 공분산 Λ는 역

적응 선형 예측 계수에 기초하여 산출되기 때문에, 공분산 Λ에 대한 잘못된 측정은 산술 부호화 과정에서 매우 긴 부호어의 비트열을 만들어 내는 문제점이 있었다.

[0073] 이러한 문제점을 완화시키기 위해, 범위가 제한된 가우시안(Gaussian) 소스(source)를 사용할 수 있다. 예를 들어, 확률 질량 함수를 계산하기 위해 소스(source) 확률 밀도 함수의 범위를 $(-3\sqrt{\lambda_i}, 3\sqrt{\lambda_i})$ 로 제한한다면, 공분산 Λ의 오차가 발생하더라도 산술 부호화 과정에서 발생할 수 있는 문제점을 완화할 수 있다. 여기서, λ_i는 입력신호의 i번째 고유값이다.

[0074] 또한, 본 발명의 일 실시예에 따른 양자화 장치 및 방법에서는 비율(rate)과 왜곡(distortion)에서의 오차를 줄이기 위해 새로운 방법이 사용될 수 있다. 예를 들어, 상기 제안된 방법은 하기의 수학식 5에서 알 수 있는 바와 같이, 비율(rate)이 증가함에 따라 페널티(penalty)를 부여하기 위해, 가중치 요소 χ를 포함할 수 있다.

수학식 5

[0075] $$\eta(D,R) = D + \chi R.$$

[0076] 또한, 본 발명의 일 실시예에 따른 양자화 장치 및 방법에서는 이득(gain)에 따라 스텝(step)의 크기 △를 선택할 수 있다. 전술한 바와 같이, 산술 부호화부(330)는 확률 질량 함수에 기초하여 인덱스를 비트열로 변환하는데, 이때 셀의 크기, 즉 스텝 크기 △가 사용된다. 기본적으로 스텝(step)의 크기 △는 입력신호의 차수나 프레임(frame)에 관계없이 상수 값이다. 그러나, 전술한 바와 같이, 본 발명의 일 실시예에 따른 양자화 장치 및 방법에서는 입력신호가 양자화되기 전에 정규화 되기 때문에, 상기 수학식 5는 하기의 수학식 6과 같이 변할 수 있다.

수학식 6

[0077] $$\begin{aligned}\eta(D,R) &= \|\kappa_i g^{-1} - \hat{d}_i \Delta\|^2 + \chi R \\ &= g^{-2}\|\kappa_i - \hat{d}_i g\Delta\|^2 + \chi R\end{aligned}$$

[0078] 여기서 g는 시변 이득(time varying gain)이다. 높은 에너지를 갖는 프레임(frame)에 대해서는, 이득 요소 g도 높은데, 상기 높은 이득 요소 g는 유효 스텝 크기(effective step size) g△와 이에 대응하는 양자화 에러(error)를 높이고, 스텝 크기 △는 신호의 에너지에 따라 시변(time varying)하는 문제점이 있었다. 이러한 문제를 방지하기 위해, 하기의 수학식 7과 같은 스텝 크기를 사용하는 새로운 이득(gain) 적응적 스텝 크기 선택 방법이 사용될 수 있다.

수학식 7

[0079] $$\tilde{\Delta} = \begin{cases} \Delta_{min}/g, & \text{if } g\Delta < \Delta_{min} \\ \Delta_{max}/g, & \text{if } g\Delta > \Delta_{max} \\ \Delta, & \text{elsewhere} \end{cases}$$

[0080] 여기서 $\triangle_{min}$과 $\triangle_{max}$는 각각 최소한의 스텝 크기, 최대한의 스텝 크기의 경계이다. 만약 이득이 너무 낮으면, 이것은 매우 낮은 에너지 프레임(frame)에 대응한다. 또한, 너무 작은 스텝 크기는 비율(rate)가 증가하는 것에 비해 효율적으로 왜곡(distortion)을 감소시키지 못한다. 이러한 프레임(frame)에 대해서 최소한의 스텝 크기 $\triangle_{min}$을 경계 값으로 설정할 수 있다.

[0081] 한편, 이득이 너무 높으면, 이것은 매우 높은 에너지 프레임(frame)에 대응하고, 너무 큰 스텝 크기는 비율 배당(rate allocation)이 매우 낮더라도 매우 높은 왜곡을 낳는다. 이러한 프레임(frame)에 대해서 최대한의 스텝 크기 $\triangle_{max}$를 경계 값으로 설정할 수 있다. 그러므로, 유효 스텝 크기(effective step size) g△는 $\triangle_{min}$과 $\triangle_{max}$ 사이에서 정해질 수 있다.

[0082] 또한, 본 발명의 일 실시예에 따른 양자화 장치 및 방법에 의하면, 주관적 음질 평가인 음성 질 지각 평가(perceptual evaluation of speech quality, PESQ) 수치도 향상될 수 있다.

[0083] 도 5는 본 발명의 일 실시예에 따른 양자화 장치 및 방법에 의한 SNR 성능을 기록한 테이블이다.

[0084] 또한, 도 6은 본 발명의 일 실시예에 따른 양자화 장치 및 방법의 음성 질 지각 평가(perceptual evaluation of speech quality, PESQ) 수치를 기록한 테이블이다.

[0085] 도 5 및 6을 참조하면, 본 발명의 일 실시예에 따른 적응적 엔트로피 제한적 양자화 장치(KLT-AECQ) 및 상기 양자화 장치에서 범위가 제한된 확률 밀도 함수 및 이득(gain) 적응적 스텝 크기 선택 방법을 추가적으로 이용한 강화된 적응적 엔트로피 제한적 양자화 장치(eKLT-AECQ)의 성능 이외에 KLT 도메인 분류 벡터 양자화(KLT-CVQ) 장치 및 KLT 도메인 엔트로피 제한적 양자화(KLT-ECQ) 장치의 성능을 기록하였다.

[0086] 본 발명의 일 실시예에 따른 적응적 엔트로피 제한적 양자화 장치(KLT-AECQ)는 연속적인 산술 부호화 과정을 수행함으로써, KLT 도메인 분류 벡터 양자화(KLT-CVQ) 장치 또는 KLT 도메인 엔트로피 제한적 양자화(KLT-ECQ) 장치보다 더 향상된 성능을 나타내었다.

[0087] 나아가, 범위가 제한된 확률 밀도 함수 및 이득(gain) 적응적 스텝 크기 선택 방법을 추가적으로 이용하고, 입력 신호의 차수(dimension)가 높을수록, SNR 및 음성 질 지각 평가(PESQ) 수치가 더욱 향상됨을 확인할 수 있다.

[0088] 본 발명의 일 실시예는 컴퓨터에 의해 실행되는 프로그램 모듈과 같은 컴퓨터에 의해 실행가능한 명령어를 포함하는 기록 매체의 형태로도 구현될 수 있다. 컴퓨터 판독 가능 매체는 컴퓨터에 의해 액세스될 수 있는 임의의 가용 매체일 수 있고, 휘발성 및 비휘발성 매체, 분리형 및 비분리형 매체를 모두 포함한다. 또한, 컴퓨터 판독 가능 매체는 컴퓨터 저장 매체 및 통신 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함한다. 통신 매체는 전형적으로 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈, 또는 반송파와 같은 변조된 데이터 신호의 기타 데이터, 또는 기타 전송 메커니즘을 포함하며, 임의의 정보 전달 매체를 포함한다.

[0089] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

[0090] 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면의 간단한 설명

[0091] 도 1은 본 발명의 일 실시예에 따른 음성신호에 대한 양자화 장치의 구성도.

[0092] 도 2는 KLT의 특성을 설명하기 위한 Voronoi- region 형태를 도시한 도면.

[0093] 도 3은 KLT의 특성을 설명하기 위한 Voronoi- region 형태를 도시한 도면.

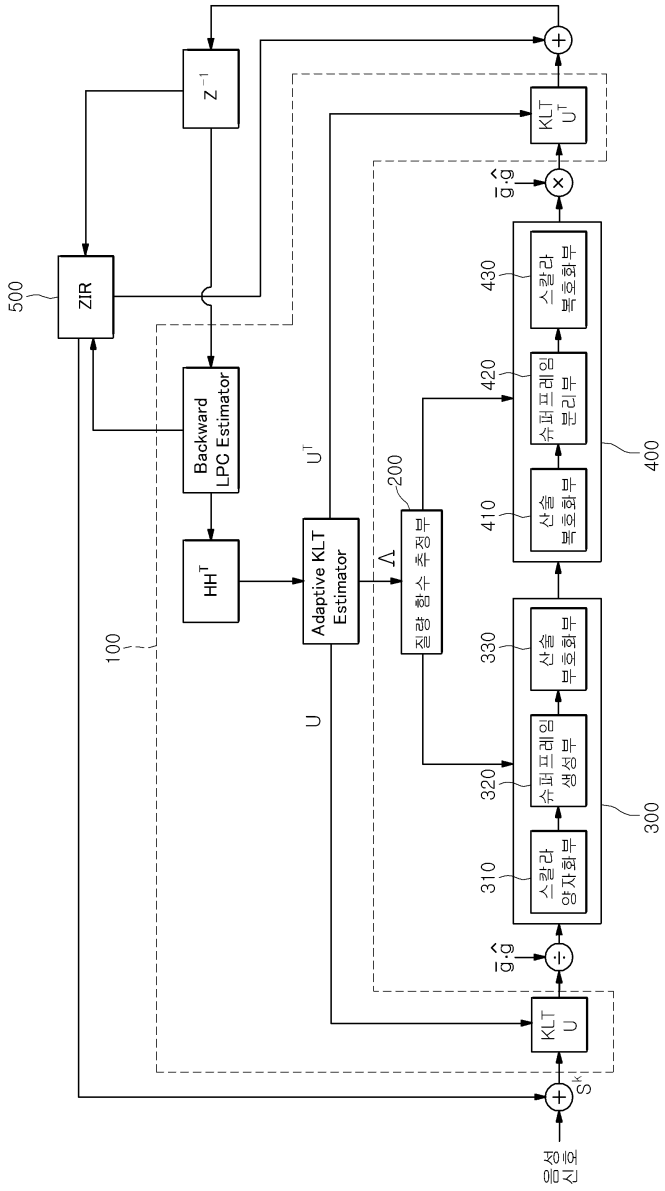
[0094] 도 4는 본 발명의 일 실시예에 따른 음성신호에 대한 양자화 방법의 흐름도.

[0095] 도 5는 본 발명의 일 실시예에 따른 양자화 장치 및 방법에 의한 SNR 성능을 기록한 테이블.

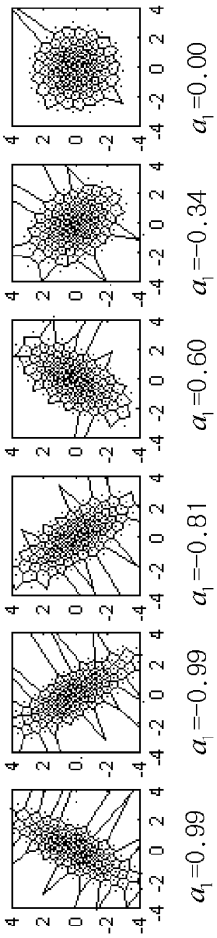
[0096] 도 6은 본 발명의 일 실시예에 따른 양자화 장치 및 방법의 음성 질 지각 평가(PESQ) 수치를 기록한 테이블.

도면

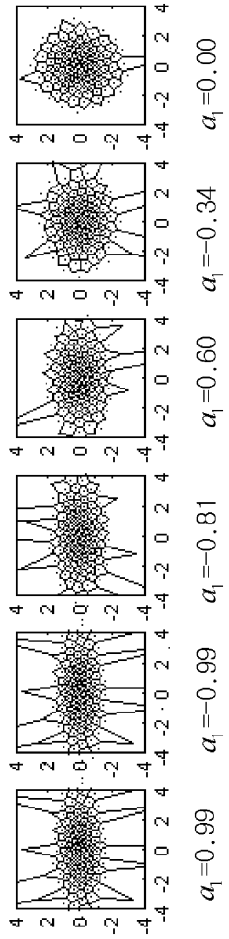
도면1


음성 신호
S^k
KLT U
$\hat{g}/\bar{g}$
스칼라 양자화부
310
슈퍼프레임 생성부
320
산술 부호화부
330
300
질량 함수 추정부
200
Λ
Adaptive KLT Estimator
U
U^T
HH^T
Backward LPC Estimator
ZIR
500
100
Z^{-1}
산술 복호화부
410
슈퍼프레임 분리부
420
스칼라 복호화부
430
400
KLT U^T
$\hat{g}\cdot\bar{g}$

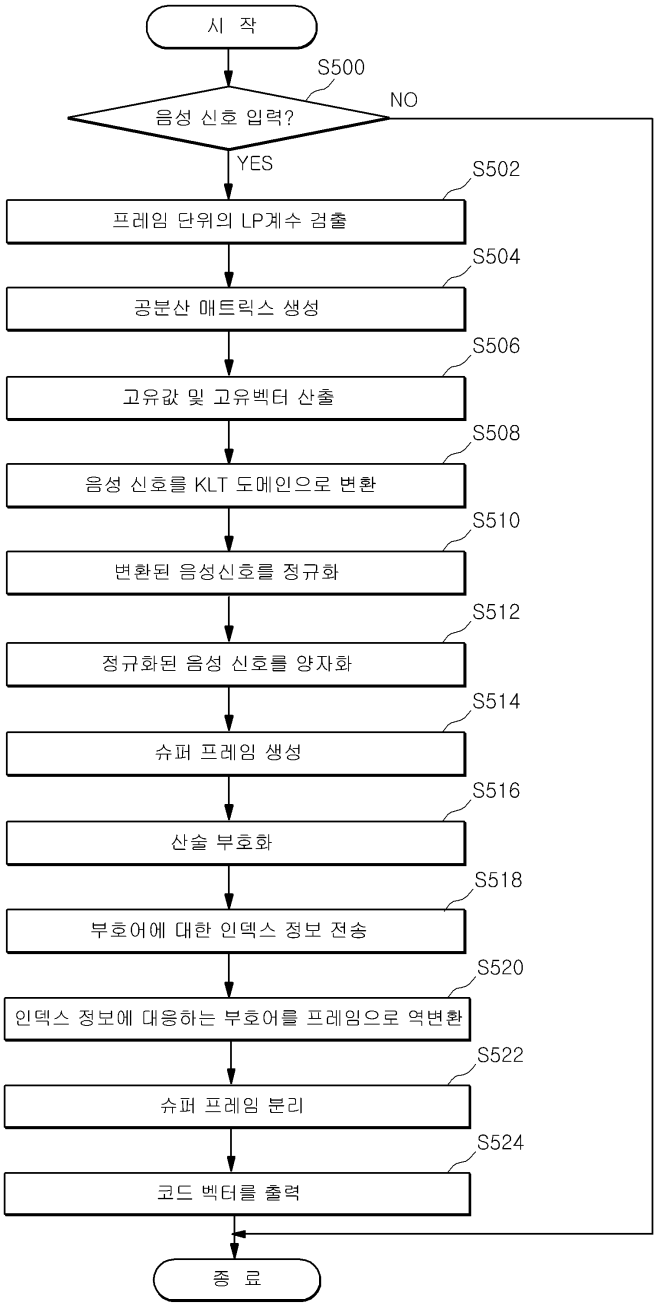
도면2



도면3



도면4


시 작
S500
음성 신호 입력?
NO
YES
S502
프레임 단위의 LP계수 검출
S504
공분산 매트릭스 생성
S506
고유값 및 고유벡터 산출
S508
음성 신호를 KLT 도메인으로 변환
S510
변환된 음성신호를 정규화
S512
정규화된 음성 신호를 양자화
S514
슈퍼 프레임 생성
S516
산술 부호화
S518
부호어에 대한 인덱스 정보 전송
S520
인덱스 정보에 대응하는 부호어를 프레임으로 역변환
S522
슈퍼 프레임 분리
S524
코드 벡터를 출력
종 료

도면5

	Female	Male	Total
KLT-CVQ [1] (k=5)	20.68	20.43	20.56
KLT-ECQ [11] (k=5)	15.86	14.25	15.05
KLT-AECQ (k=5)	21.16	19.73	20.45
eKLT-AECQ (k=5)	29.10	29.14	29.12
eKLT-AECQ (k=10)	30.45	30.39	30.42

도면6

	Female	Male	Total
KLT-CVQ [1] (k=5)	3.65	3.83	3.74
KLT-ECQ [11] (k=5)	3.13	3.47	3.30
KLT-AECQ (k=5)	3.65	3.84	3.74
eKLT-AECQ (k=5)	4.00	4.07	4.03
eKLT-AECQ (k=10)	4.20	4.27	4.23