



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0065439
(43) 공개일자 2025년05월12일

(51) 국제특허분류(Int. Cl.)
H03M 7/40 (2006.01) H04N 19/124 (2014.01)
H04N 19/182 (2014.01)
(52) CPC특허분류
H03M 7/4006 (2013.01)
H04N 19/124 (2015.01)
(21) 출원번호 10-2025-7014556(분할)
(22) 출원일자(국제) 2010년10월01일
심사청구일자 없음
(62) 원출원 특허 10-2023-7033869
원출원일자(국제) 2010년10월01일
심사청구일자 2023년11월02일
(85) 번역문제출일자 2025년04월30일
(86) 국제출원번호 PCT/EP2010/064644
(87) 국제공개번호 WO 2011/042366
국제공개일자 2011년04월14일
(30) 우선권주장
09305961.6 2009년10월09일
유럽특허청(EPO)(EP)

(71) 출원인
돌비 인터네셔널 에이비
아일랜드 디02 브이케이60 더블린 그랜드 커널 독
랜즈 블록 씨 서 존 로저슨스 키 77
(72) 발명자
부에브볼트, 올리베르
독일 30625 하노버 카를 뤼헤르트 알리 74 리서치
앤드 이노베이션 도이취 톱슨 오후게
(74) 대리인
양영준, 백만기

전체 청구항 수 : 총 1 항

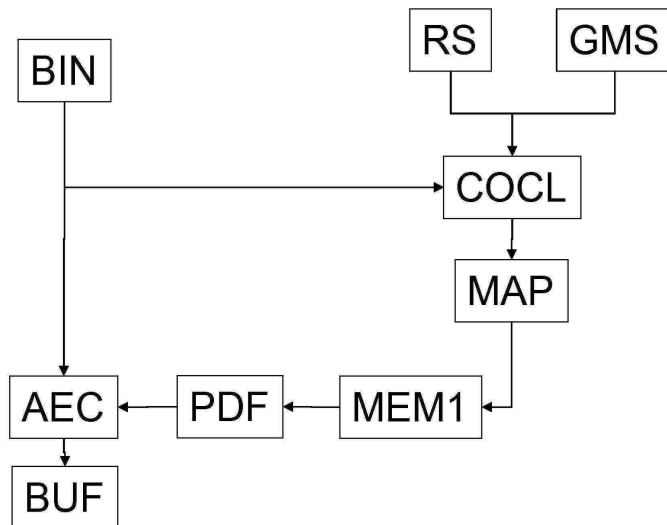
(54) 발명의 명칭 산술 인코딩 또는 산술 디코딩 방법 및 장치

(57) 요약

본 발명은 선행 스펙트럼 계수를 사용하여 현재 스펙트럼 계수를 산술 인코딩하는 방법 및 장치를 제안한다. 상기 선행 스펙트럼 계수는 이미 인코딩되어 있고, 상기 선행 스펙트럼 계수 및 상기 현재 스펙트럼 계수 양쪽 모두는 비디오, 오디오 또는 음성 신호 샘플 값의 시간-주파수 변환을 양자화하는 것으로부터 얻어지는 하나 이상

(뒷면에 계속)

대표도 - 도1



의 양자화된 스펙트럼에 포함되어 있다. 상기 방법은 선행 스펙트럼 계수를 처리하는 단계, 처리된 선행 스펙트럼 계수를 사용하여 적어도 2개의 상이한 컨텍스트 클래스 중 하나인 컨텍스트 클래스를 결정하는 단계, 결정된 컨텍스트 클래스 및 적어도 2개의 상이한 컨텍스트 클래스로부터 적어도 2개의 상이한 확률 밀도 함수로의 매핑을 사용하여 확률 밀도 함수를 결정하는 단계, 및 결정된 확률 밀도 함수에 기초하여 현재 스펙트럼 계수를 산술 인코딩하는 단계를 포함하고, 선행 스펙트럼 계수를 처리하는 단계는 컨텍스트 클래스를 결정하는 데 사용하기 위해 선행 스펙트럼 계수의 절대값을 비균등 양자화하는 단계를 포함한다.

(52) CPC특허분류

H04N 19/182 (2015.01)

명세서

청구범위

청구항 1

제1항

발명의 설명

기술 분야

[0001] 본 발명은 멀티미디어 데이터의 산술 인코딩 및 디코딩에 관한 것이다.

배경 기술

[0002] 산술 코딩은 무손실 데이터 압축 방법이다. 산술 코딩은 확률 밀도 함수(PDF)에 기초하고 있다. 압축 효과를 달성하기 위해, 코딩이 기초하고 있는 확률 밀도 함수가 데이터가 실제로 따르고 있는 실제 확률 밀도 함수와 동일하거나 적어도 비슷해야만 한다 - 실제 확률 밀도 함수와 비슷할수록 좋다 -.

[0003] 산술 코딩이 적당한 확률 밀도 함수에 기초하고 있는 경우, 산술 코딩은 상당한 압축을 달성하여, 적어도 거의 최적의 코드를 얻을 수 있다. 따라서, 산술 코딩은 계수 시퀀스의 인코딩 및 디코딩을 위해 오디오, 음성 또는 비디오 코딩에서 빈번히 사용되는 기법으로서, 여기서 계수는 이진 표현으로 된 비디오 픽셀이나 오디오 또는 음성 신호 샘플 값의 양자화된 시간-주파수 변환이다.

[0004] 압축을 더욱 향상시키기 위해, 산술 코딩은 한 세트의 확률 밀도 함수에 기초할 수 있고, 여기서 현재 계수(current coefficient)를 코딩하는 데 사용되는 확률 밀도 함수는 상기 현재 계수의 컨텍스트에 의존한다. 즉, 동일한 양자화 값을 갖는 계수가 나오는 컨텍스트에 따라 상기 동일한 양자화 값을 코딩하는 데 상이한 확률 밀도 함수가 사용될 수 있다. 계수의 컨텍스트는 각자의 계수에 이웃하는 하나 이상의 이웃하는 계수들의 이웃, 예컨대 시퀀스에서 각자의 인코딩될 또는 디코딩될 계수에 인접하여 선행하는 하나 이상의 이미 인코딩된 또는 이미 디코딩된 계수들의 서브시퀀스에 포함되는 계수의 양자화 값에 의해 정의된다. 이웃이 처할 수 있는 상이한 가능한 상황들 각각은, 각각이 연관된 확률 밀도 함수에 매핑되는 상이한 가능한 컨텍스트를 정의한다.

[0005] 실제로, 상기 압축 향상은 이웃이 충분히 클 경우에만 명백하게 된다. 이것은 다수의 상이한 가능한 컨텍스트는 물론 대응하는 엄청난 수의 가능한 확률 밀도 함수 또는 대응하는 복잡한 매핑의 조합 급증(combinatory explosion)으로 잘 동작한다.

[0006] 컨텍스트 기반 산술 코딩 방식의 예는 USAC(Unified Speech and Audio Coding)에 대한 참조 모델을 제안한 ISO/IEC JTC1/SC29/WG11 N10215(2008년 10월, 대한민국 부산)에서 찾아볼 수 있다. 이 제안에 따르면, 이미 디코딩된 4-튜플이 컨텍스트를 위해 고려되고 있다.

[0007] USAC 관련 컨텍스트 기반 산술 코딩의 다른 예는 ISO/IEC JTC1/SC29/WG11 N10847(2009년 7월, 영국 런던)에서 찾아볼 수 있다.

[0008] 고차 조건부 엔트로피 인코딩에서의 복잡도 감소를 위해, 미국 특허 제5,298,896호는 컨디셔닝 심볼(conditioning symbols)의 비균등 양자화를 제안하고 있다.

발명의 내용

[0009] 엄청난 수의 처리될 컨텍스트에 대응하여, 저장, 검색 및 처리될 필요가 있는 엄청난 수의 확률 밀도 함수 또는 적어도 그에 대응하여 복잡한, 컨텍스트로부터 확률 밀도 함수로의 매핑이 있다. 이것은 인코딩/디코딩 지연 시간 및 메모리 용량 요구사항 중 적어도 하나를 증가시킨다. 인코딩/디코딩 지연 시간 및 메모리 용량 요구사항 중 적어도 하나를 감소시키면서 압축을 마찬가지로 잘 달성할 수 있게 하는 대안의 해결책이 기술 분야에 필요하다.

[0010] 이 필요성을 해결하기 위해, 본 발명은 제1항의 특징을 포함하는 인코딩 방법, 제2항의 특징을 포함하는 디코딩 방법, 제13항의 특징을 포함하는 산술 인코딩 장치, 제14항의 특징을 포함하는 산술 디코딩 장치, 및 제15항에

따른 저장 매체를 제안한다.

- [0011] 추가적인 제안된 실시예의 특징은 종속 청구항에 명시되어 있다.
- [0012] 상기 산술 인코딩 또는 디코딩 방법은, 각각, 현재 스펙트럼 계수(current spectral coefficient)의 산술 인코딩 또는 디코딩을 위해 선행 스펙트럼 계수(preceding spectral coefficient)를 사용하며, 여기서 상기 선행 스펙트럼 계수는, 각각, 이미 인코딩 또는 디코딩되어 있다. 상기 선행 스펙트럼 계수 및 상기 현재 스펙트럼 계수 둘다는 비디오, 오디오 또는 음성 신호 샘플 값의 시간-주파수 변환을 양자화하는 것으로부터 얻어지는 하나 이상의 양자화된 스펙트럼에 포함되어 있다. 상기 방법은 선행 스펙트럼 계수를 처리하는 단계, 처리된 선행 스펙트럼 계수를 사용하여 적어도 2개의 상이한 컨텍스트 클래스(context class) 중 하나인 컨텍스트 클래스를 결정하는 단계, 결정된 컨텍스트 클래스 및 적어도 2개의 상이한 컨텍스트 클래스로부터 적어도 2개의 상이한 확률 밀도 함수로의 매핑을 사용하여 확률 밀도 함수를 결정하는 단계, 및 결정된 확률 밀도 함수에 기초하여, 각각, 현재 스펙트럼 계수를 산술 인코딩 또는 디코딩하는 단계를 더 포함한다. 선행 스펙트럼 계수를 처리하는 단계가 선행 스펙트럼 계수의 절대값(absolute)을 비균등 양자화(non-uniformly quantizing)하는 단계를 포함하는 것이 이 방법의 특징이다.
- [0013] 확률 밀도 함수를 결정하기 위해 컨텍스트에 대한 대안으로서 컨텍스트 클래스를 사용하는 것은 상이하지만 아주 유사한 확률 밀도 함수가 얻어지는 2개 이상의 상이한 컨텍스트를, 단일 확률 밀도 함수에 매핑되는 단일 컨텍스트 클래스로 그룹화하는 것을 가능하게 한다. 그룹화는 컨텍스트 클래스를 결정하기 위해 선행 스펙트럼 계수의 비균등 양자화된 절대값을 사용하는 것에 의해 달성된다.
- [0014] 예를 들어, 선행 스펙트럼 계수를 처리하는 단계가 컨텍스트 클래스를 결정하는 데 사용하기 위해 선행 스펙트럼 계수의 양자화된 절대값의 합을 구하는 단계를 포함하는 실시예가 있다. 유사하게, 처리 수단이 컨텍스트 클래스를 결정하는 데 사용하기 위해 선행 스펙트럼 계수의 양자화된 절대값의 합을 구하도록 적응되어 있는 대응하는 산술 인코딩 장치의 실시예는 물론 대응하는 산술 인코딩 장치의 실시예가 있다.
- [0015] 추가적인 장치 실시예에서, 처리 수단은, 선행 스펙트럼 계수를 처리하는 단계가 선행 스펙트럼 계수의 절대값이 제1 양자화 방식에 따라 양자화되는 제1 양자화, 제1 양자화 방식에 따라 양자화된 선행 스펙트럼 계수의 절대값의 분산이 구해지는 분산 구하기, 구해진 분산을 사용하여 적어도 2개의 상이한 비선형 제2 양자화 방식 중 하나를 선택하는 것, 및 제1 양자화 방식에 따라 양자화된 선행 스펙트럼 계수의 절대값이 선택된 비선형 제2 양자화 방식에 따라 추가로 양자화되는 제2 양자화를 더 포함하도록 적응된다. 추가적인 방법 실시예는 대응하는 단계들을 포함한다. 분산 구하기는 제1 양자화 방식에 따라 양자화된 선행 스펙트럼 계수의 절대값의 합을 구하는 것 및 구해진 합을 적어도 하나의 임계값과 비교하는 것을 포함할 수 있다.
- [0016] 추가의 실시예에서, 각각의 장치의 처리 수단은, 처리하는 단계로부터 제1 결과 또는 적어도 상이한 제2 결과가 얻어지도록 적응될 수 있다. 이어서, 컨텍스트 클래스를 결정하는 것은 그의 처리로부터 제1 결과가 얻어진 다수의 그 선행 스펙트럼 계수를 구하는 것 및 컨텍스트 클래스를 결정하기 위해 구해진 다수의 선행 스펙트럼 계수를 사용하는 것을 더 포함한다.
- [0017] 각각의 장치는 모드 전환 신호 및 리셋 신호 중 적어도 하나를 수신하는 수단을 포함할 수 있고, 이 때 장치는 적어도 하나의 수신된 신호를 사용하여 컨텍스트 클래스의 결정을 제어하도록 적응된다.
- [0018] 적어도 2개의 상이한 확률 밀도 함수를 결정하기 위해 적어도 2개의 상이한 확률 밀도 함수가 대표적인 데이터 집합을 사용하여 미리 결정될 수 있고, 탐색 테이블 또는 해시 테이블을 사용하여 매핑이 실현될 수 있다.

도면의 간단한 설명

- [0019] 본 발명의 예시적인 실시예가 도면에 예시되어 있고, 이하의 설명에서 더 상세히 설명되어 있다. 예시적인 실시예는 특허청구범위에 한정된 본 발명의 범위 및 사상을 제한하기 위한 것이 아니라 단지 본 발명을 설명하기 위해 기술되어 있다.
- 도 1은 본 발명의 인코더의 일 실시예를 예시적으로 나타낸 도면이다.
- 도 2는 본 발명의 디코더의 일 실시예를 예시적으로 나타낸 도면이다.
- 도 3은 컨텍스트 클래스를 결정하는 컨텍스트 분류기의 제1 실시예를 예시적으로 나타낸 도면이다.
- 도 4는 컨텍스트 클래스를 결정하는 컨텍스트 분류기의 제2 실시예를 예시적으로 나타낸 도면이다.

도 5의 (a)는 주파수 영역 모드에서 인코딩될 또는 디코딩될 현재 스펙트럼 빈(current spectral bin)에 선행하는 선행 스펙트럼 빈(preceding spectral bin)의 제1 이웃을 예시적으로 나타낸 도면이다.

도 5의 (b)는 가중된 선형 예측 변환 모드(weighted linear prediction transform mode)에서 인코딩될 또는 디코딩될 현재 스펙트럼 빈에 선행하는 선행 스펙트럼 빈의 제2 이웃을 예시적으로 나타낸 도면이다.

도 6의 (a)는 주파수 영역 모드에서 인코딩될 또는 디코딩될 현재 최하위 주파수 스펙트럼 빈(current lowest frequency spectral bin)에 선행하는 선행 스펙트럼 빈의 제3 이웃을 예시적으로 나타낸 도면이다.

도 6의 (b)는 주파수 영역 모드에서 인코딩될 또는 디코딩될 현재 두번째 최하위 주파수 스펙트럼 빈(current second lowest frequency spectral bin)에 선행하는 선행 스펙트럼 빈의 제4 이웃을 예시적으로 나타낸 도면이다.

도 7의 (a)는 가중된 선형 예측 변환 모드에서 인코딩될 또는 디코딩될 현재 최하위 주파수 스펙트럼 빈에 선행하는 선행 스펙트럼 빈의 제5 이웃을 예시적으로 나타낸 도면이다.

도 7의 (b)는 가중된 선형 예측 변환 모드에서 인코딩될 또는 디코딩될 현재 두번째 최하위 주파수 스펙트럼 빈에 선행하는 선행 스펙트럼 빈의 제6 이웃을 예시적으로 나타낸 도면이다.

도 7의 (c)는 가중된 선형 예측 변환 모드에서 인코딩될 또는 디코딩될 현재 세번째 최하위 주파수 스펙트럼 빈에 선행하는 선행 스펙트럼 빈의 제7 이웃을 예시적으로 나타낸 도면이다.

도 7의 (d)는 가중된 선형 예측 변환 모드에서 인코딩될 또는 디코딩될 현재 세번째 최하위 주파수 스펙트럼 빈에 선행하는 선행 스펙트럼 빈의 제8 이웃을 예시적으로 나타낸 도면이다.

도 8은 인코딩될 또는 디코딩될 상이한 스펙트럼 빈의 이웃을 예시적으로 나타낸 도면으로서, 상기 상이한 스펙트럼 빈은 주파수 영역 모드에서 인코딩/디코딩의 시작 또는 리셋 신호의 발생 이후에 인코딩될 또는 디코딩될 제1 스펙트럼에 포함된다.

도 9는 가중된 선형 예측 변환 모드에서 인코딩될 또는 디코딩될 상이한 스펙트럼 빈의 추가적인 이웃을 예시적으로 나타낸 도면으로서, 상기 상이한 스펙트럼 빈은 가중된 선형 예측 변환 모드에서 인코딩/디코딩의 시작 또는 리셋 신호의 발생 이후에 인코딩될 또는 디코딩될 제2 스펙트럼에 포함된다.

발명을 실시하기 위한 구체적인 내용

- [0020] 본 발명은 그에 대응하여 적응된 처리 장치를 포함하는 임의의 전자 장치 상에서 실현될 수 있다. 예를 들어, 산술 디코딩 장치는 텔레비전, 휴대폰, 또는 개인용 컴퓨터, mp3 플레이어, 내비게이션 시스템, 또는 카 오디오 시스템에서 실현될 수 있다. 산술 인코딩 장치는, 몇가지 예를 들자면, 휴대폰, 개인용 컴퓨터, 능동 자동차 내비게이션 시스템, 디지털 스틸 카메라, 디지털 비디오 카메라, 또는 디터폰(Dictaphone)에서 실현될 수 있다.
- [0021] 이하에서 기술되는 예시적인 실시예는 멀티미디어 샘플의 시간-주파수 변환의 양자화로부터 얻어지는 양자화된 스펙트럼 빈의 인코딩 및 디코딩에 관한 것이다.
- [0022] 본 발명은 이미 전송된 양자화된 스펙트럼 빈, 예컨대 시퀀스에서 현재 양자화된 스펙트럼 빈(BIN)에 선행하는 선행 양자화된 스펙트럼 빈이 현재 양자화된 스펙트럼 빈(BIN)의 산술 인코딩 및 디코딩을 위해, 각각, 사용될 확률 밀도 함수(PDF)를 결정하는 데 사용되는 방식에 기초하고 있다.
- [0023] 기술된 예시적인 산술 인코딩 또는 산술 디코딩 방법 및 장치 실시예는, 각각, 몇개의 비균등 양자화 단계 또는 수단을 포함한다. 모든 단계 또는 수단은, 각각, 모두가 최고 코딩 효율을 제공하지만, 각각의 단계 또는 수단은, 각각, 단독으로 이미 본 발명의 개념을 실현하고 인코딩/디코딩 지연 시간 및/또는 메모리 요구사항에 관한 이점을 제공한다. 따라서, 상세한 설명은 각각 기술된 단계들 또는 수단들 중 하나만을 실현하는 예시적인 실시예를 기술하는 것은 물론, 기술된 단계들 또는 수단들 중 2개 이상의 조합을 실현하는 예시적인 실시예를 기술하는 것으로도 해석되어야 한다.
- [0024] 방법의 예시적인 실시예에 포함될 수 있지만 포함될 필요는 없는 제1 단계는 어느 일반 변환 모드가 사용될 것인지가 결정되는 전환 단계이다. 예를 들어, USAC 무잡음 코딩 방식(Noiseless Coding Scheme)에서, 일반 변환 모드는 FD(Frequency Domain, 주파수 영역) 모드 또는 wLPT(weighted Linear Prediction Transform, 가중된 선형 예측 변환) 모드일 수 있다. 각각의 일반 모드는 PDF를 결정하기 위해 상이한 이웃, 즉 각각 이미 인코딩될 또는 디코딩될 스펙트럼 빈의 상이한 집단을 사용할 수 있다.

- [0025] 그 후에, 현재 스펙트럼 빈(BIN)의 컨텍스트가 컨텍스트 발생 모듈(COCL)에서 결정될 수 있다. 결정된 컨텍스트로부터, 컨텍스트를 분류함으로써 컨텍스트 클래스가 결정되고, 여기서, 분류 이전에, 컨텍스트가 바람직하게는 컨텍스트의 스펙트럼 빈의 비균등 양자화(NUQ1)(반드시 그럴 필요는 없음)에 의해 처리된다. 분류는 컨텍스트의 분산(VES)을 추정하는 것 및 분산을 적어도 하나의 임계값과 비교하는 것을 포함할 수 있다. 또는, 분산 추정치가 컨텍스트로부터 직접 구해진다. 분산 추정치는 이어서 바람직하게는 비선형(반드시 그럴 필요는 없음)인 추가의 양자화(NUQ2)를 제어하는 데 사용된다.
- [0026] 도 1에 예시적으로 나타낸 인코딩 프로세스에서, 현재 양자화된 스펙트럼 빈(BIN)을 인코딩하는 데 적합한 확률 밀도 함수(PDF)가 결정된다. 이를 위해, 디코더측에서도 이미 알고 있는 정보만이 사용될 수 있다. 즉, 인코딩된 또는 디코딩된 선행 양자화된 스펙트럼 빈만이 사용될 수 있다. 이것은 컨텍스트 분류기 블록(COCL)에서 행해진다. 그곳에서, 선택된 선행 스펙트럼 빈은 실제 컨텍스트 클래스를 결정하는 데 사용되는 이웃(NBH)을 정의한다. 컨텍스트 클래스는 컨텍스트 클래스 번호에 의해 표시될 수 있다. 컨텍스트 클래스 번호는 매핑(MAP)을 통해, 예컨대 탐색 테이블 또는 해시 테이블을 통해 PDF 메모리(MEM1)로부터 대응하는 PDF를 검색하는 데 사용된다. 컨텍스트 클래스의 결정은 선택된 모드에 따라 상이한 이웃을 사용할 수 있게 하는 일반 모드 스위치(GMS)에 의존할 수 있다. 앞서 언급한 바와 같이, USAC의 경우, 2개의 일반 모드(FD 모드 및 wLPT 모드)가 있을 수 있다. 일반 모드 스위치(GMS)가 인코더측에서 실현되는 경우, 모드 변경 신호 또는 현재 일반 신호가, 디코더에서도 알도록, 비트스트림에 포함되어야만 한다. 예를 들어, ISO/IEC JTC1/SC29/WG11 N10847(2009년 7월, 영국 런던)에 의해 제안된 USAC(Unified Speech and Audio Coding)에 대한 참조 모델에서, 일반 모드의 전송을 위해 제안된 WD 테이블 4.4 core_mode 및 테이블 4.5 core_mode0/1이 있다.
- [0027] 산술 인코더(AEC)에 의한 현재 양자화된 스펙트럼 빈(BIN)의 인코딩에 적합한 PDF의 결정 후에, 현재 양자화된 스펙트럼 빈(BIN)이 이웃 메모리(MEM2)에 입력되는데, 즉 현재 빈(BIN)이 선행 빈으로 된다. 이웃 메모리(MEM2)에 포함된 선행 스펙트럼 빈은 블록(COCL)에서 그 다음 스펙트럼 빈(BIN)을 코딩하는 데 사용될 수 있다. 현재 스펙트럼 빈(BIN)을 기억하는 동안, 또는 그 이전 또는 그 이후에, 상기 현재 빈(BIN)이 산술 인코더(AEC)에 의해 산술 인코딩된다. 산술 인코딩(AEC)의 출력은 비트 버퍼(BUF)에 저장되거나, 비트스트림에 직접 기입된다.
- [0028] 비트스트림 또는 버퍼(BUF)의 내용은, 예를 들어, 케이블 또는 위성을 통해 전송되거나 방송될 수 있다. 또는, 산술 인코딩된 스펙트럼 빈이 DVD, 하드 디스크, 블루레이 디스크 등과 같은 저장 매체에 기록될 수 있다. PDF-메모리(MEM1) 및 이웃 메모리(MEM2)는 단일 물리 메모리에 실현될 수 있다.
- [0029] 리셋 스위치(RS)는 때때로 선행 스펙트럼을 알지 못한 상태에서 인코딩 및 디코딩이 시작될 수 있는 전용 프레임(전용 프레임은 디코딩 진입점이라고 함)에서 인코딩 또는 디코딩을 재시작할 수 있게 할 수 있다. 리셋 스위치(RS)가 인코더측에서 실현되는 경우, 리셋 신호가, 디코더에서도 알도록, 비트스트림에 포함되어야만 한다. 예를 들어, ISO/IEC JTC1/SC29/WG11 N10847(2009년 7월, 영국 런던)에 의해 제안된 USAC(Unified Speech and Audio Coding)에 대한 참조 모델에서, WD 테이블 4.10 및 테이블 4.14에 arith_reset_flag가 있다.
- [0030] 대응하는 이웃 기반 디코딩 방식이 도 2에 예시적으로 나타내어져 있다. 이는 인코딩 방식과 유사한 블록을 포함하고 있다. 산술 디코딩에 사용될 PDF의 결정은, 인코더 및 디코더 둘다에서, 결정된 PDF가 동일한 것을 보장하기 위해 인코딩 방식에서와 동일하다. 산술 디코딩은 비트 버퍼(BUF)로부터 비트를 받거나 비트스트림을 직접 받으며, 결정된 PDF를 사용하여 현재 양자화된 스펙트럼 빈(BIN)을 디코딩한다. 그 후에, 디코딩된 양자화된 스펙트럼 빈이 컨텍스트 클래스 번호 결정 블록(COCL)의 이웃 메모리(MEM2)에 입력되고, 그 다음 스펙트럼 빈을 디코딩하는 데 사용될 수 있다.
- [0031] 도 3은 컨텍스트 클래스를 결정하는 컨텍스트 분류기(COCL)의 제1 실시예를 상세히 예시적으로 나타낸 것이다.
- [0032] 현재 양자화된 스펙트럼 빈(BIN)을 스펙트럼 메모리(MEM2)에 저장하기 전에, 이는 블록(NUQ1)에서 비균등 양자화될 수 있다. 이것은 2가지 이점을 가지는데, 첫째, 보통 16비트 부호 있는 정수 값인 양자화된 빈의 보다 효율적인 저장을 가능하게 한다. 둘째, 각각의 양자화된 빈이 가질 수 있는 값의 수가 감소된다. 이것은 블록(CLASS)에서의 컨텍스트 클래스 결정 프로세스에서 가능한 컨텍스트 클래스의 수를 엄청나게 감소시킬 수 있다. 게다가, 컨텍스트 클래스 결정에서와 같이, 양자화된 빈의 부호가 무시될 수 있고, 절대값의 계산이 비균등 양자화 블록(NUQ1)에 포함될 수 있다. 블록(NUQ1)에 의해 수행될 수 있는 예시적인 비균등 양자화가 표 1에 나타내어져 있다. 이 예에서, 비균등 양자화 후에, 각각의 빈에 대해 3개의 상이한 값이 가능하다. 그러나, 일반적으로, 비균등 양자화에 대한 유일한 제약조건은 비균등 양자화가 빈이 취할 수 있는 값의 수를 감소시킨다는

것이다.

표 1

절대값의 계산을 포함한 예시적인 비균등 양자화 단계

양자화된 스펙트럼 빈의 절대값	0	1	2	3	4	5	6	7	8	>8
비균등 양자화	0	1	2							

[0034] 비균등 양자화된/매핑된 스펙트럼 빈이 스펙트럼 메모리(MEM2)에 저장된다. 선택된 일반 모드 선택(GMS)에 따라, 코딩될 각각의 빈에 대한 컨텍스트 클래스 결정(CLASS)을 위해, 스펙트럼 빈의 선택된 이웃(NBH)이 선택된다.

[0035] 도 5의 (a)는 인코딩될 또는 디코딩될 스펙트럼 빈(BIN)의 제1 예시적인 이웃(NBH)을 예시적으로 나타낸 것이다.

[0036] 이 예에서, 실제 또는 현재 스펙트럼(프레임)의 스펙트럼 빈 및 하나의 선행 스펙트럼(프레임)의 스펙트럼 빈만이 이웃(NBH)을 정의한다. 물론, 2개 이상의 선행 스펙트럼으로부터의 스펙트럼 빈을 이웃의 일부로서 사용하는 것이 가능하며, 그 결과 복잡도가 더 높아지지만, 또한 결국은 더 높은 코딩 효율을 제공할 수 있다. 실제 스펙트럼으로부터, 단지 이미 전송된 빈만이 이웃(NBH)을 정의하는 데 사용될 수 있다는 것에 유의해야 하는데, 그 이유는 디코더에서도 액세스가능해야만 하기 때문이다. 여기서는 물론 이하의 예에서, 스펙트럼 빈에 대한 하위 주파수로부터 상위 주파수의 전송 순서가 가정된다.

[0037] 선택된 이웃(NBH)은 이어서 컨텍스트 클래스 결정 블록(COCL)에서 입력으로서 사용된다. 이하에서, 먼저 컨텍스트 클래스 결정을 뒷받침하는 일반적 생각 및 간단화된 버전이 설명되고, 이어서 특수한 실현이 기술된다.

[0038] 컨텍스트 클래스 결정을 뒷받침하는 일반적 생각은 코딩될 빈의 분산의 신뢰성있는 추정을 가능하게 하는 것이다. 이러한 예측된 분산은, 다시 코딩될 빈의 PDF의 추정치를 얻는 데 사용될 수 있다. 분산 추정을 위해, 이웃에 있는 빈의 부호를 평가할 필요가 없다. 따라서, 부호가 스펙트럼 메모리(MEM2)에 저장하기 전에 양자화 단계에서 이미 무시될 수 있다. 아주 간단한 컨텍스트 클래스 결정은 다음과 같을 수 있다: 스펙트럼 빈(BIN)의 이웃(NBH)은 도 5의 (a)에서와 같을 수 있고 7개의 스펙트럼 빈으로 이루어져 있다. 예시적으로, 표에 나타난 비균등 양자화가 사용되는 경우, 각각의 빈은 3개의 값을 가질 수 있다. 이 결과 $3^7 = 2187$ 개의 가능한 컨텍스트 클래스가 얻어진다.

[0039] 이 가능한 컨텍스트 클래스의 수를 추가로 감소시키기 위해, 이웃(NBH)에 있는 각각의 빈의 상대 위치가 무시될 수 있다. 따라서, 각각 값 0, 1 또는 2를 갖는 빈의 수만이 카운트되고, 여기서 물론 0-빈의 수, 1-빈의 수 및 2-빈의 수의 합은 이웃에 있는 빈의 총 수와 같다. 각각이 3개의 상이한 값 중 하나를 가질 수 있는 n개의 빈을 포함하는 이웃(NBH)에, $0.5 * (n^2 + 3 * n + 2)$ 개의 컨텍스트 클래스가 있다. 예를 들어, 7개 빈의 이웃에는, 36개의 가능한 컨텍스트 클래스가 있고, 6개 빈의 이웃에는 28개의 가능한 컨텍스트 클래스가 있다.

[0040] 보다 복잡하지만 여전히 꽤 간단한 컨텍스트 클래스 결정은 연구에서 보여주듯이 동일한 주파수에서의 선행 스펙트럼의 스펙트럼 빈(도 5의 (a), 도 5의 (b), 도 6의 (a), 도 6의 (b), 도 7의 (a), 도 7의 (b), 도 7의 (c), 도 8 및 도 9에서 점선 원으로 나타낸 스펙트럼 빈)이 특히 중요하다는 것을 고려하고 있다. 이웃에 있는 다른 빈(각각의 도면에서 수평 줄무늬 원으로 나타내어져 있음)에 대해, 상대 위치는 덜 중요하다. 따라서, 선행 스펙트럼에서 동일한 주파수에 있는 빈이 컨텍스트 클래스 결정을 위해 명시적으로 사용되는 반면, 나머지 6개 빈에 대해, 0-빈의 수, 1-빈의 수 및 2-빈의 수만이 카운트된다. 이 결과 $3 * 28 = 84$ 개의 가능한 컨텍스트 클래스가 얻어진다. 실험은 이러한 컨텍스트 분류가 FD 모드에 대해 아주 효율적이라는 것을 보여주었다.

[0041] 컨텍스트 클래스 결정이 제2 비균등 양자화(NUQ2)를 제어하는 분산 추정(VES)에 의해 확장될 수 있다. 이것은 컨텍스트 클래스 발생(COCL)이 코딩될 빈의 예측된 분산의 더 높은 동적 범위에 더 잘 적용할 수 있게 한다. 확장된 컨텍스트 클래스 결정의 대응하는 블록도가 도 4에 예시적으로 도시되어 있다.

[0042] 도 4에 도시된 예에서, 비균등 양자화는 2개의 단계로 분리되어 있고, 선행 단계는 보다 미세한 양자화(블록 NUQ1)를 제공하고, 후속 단계는 보다 대략적인 양자화(블록 NUQ2)를 제공한다. 이것은 양자화가, 예컨대, 이웃의 분산에 적응하는 것을 가능하게 한다. 이웃의 분산이 분산 추정 블록(VES)에서 추정되고, 여기서 분산 추정은 블록(NUQ1)에서 이웃(NBH)에서의 빈의 상기 선행하는 보다 미세한 양자화에 기초하고 있다. 분산의 추정이

정확할 필요는 없고 아주 대략적일 수 있다. 예를 들어, USAC 응용 프로그램이 상기 보다 미세한 양자화 이후에 이웃(NBH)에 있는 빈의 절대값의 합이 분산 임계값을 만족시키거나 초과하는지 여부를 결정하는 것으로 충분한데, 즉 높은 분산과 낮은 분산 사이의 전환으로 충분하다.

[0043] 2-단계 비균등 양자화는 표 2에 나타낸 바와 같을 수 있다. 이 예에서, 낮은 분산 모드는 표 2에 나타낸 1-단계 양자화에 대응한다.

표 2

[0044] 양자화된 스펙트럼 빈의 절대값	0	1	2	3	4	5	6	7	8	>8
보다 미세한 양자화 단계 1(6개의 값)	0	1	2	3		4		5		
보다 대략적인 양자화 단계 2(낮은 분산)(3개의 값)	0	1		2						
보다 대략적인 양자화 단계 2(높은 분산)(3개의 값)	0			1			2			

[0045] 표 2는 예시적인 2-단계 비균등 양자화를 나타낸 것이고, 제2 또는 후속 단계는 분산이 높은 것으로 추정되는지 낮은 것으로 추정되는지에 따라 상이하게 양자화한다.

[0046] 블록(CLASS)에서의 최종적인 컨텍스트 클래스 결정은 도 3의 간략화된 버전에서와 동일하다. 분산 모드에 따라 상이한 컨텍스트 클래스 결정을 사용하는 것이 가능하다. 또한, 3개 이상의 분산 모드를 사용하는 것이 가능하며, 이 결과 물론 컨텍스트 클래스의 수의 증가 및 복잡도의 증가가 있게 된다.

[0047] 스펙트럼에서의 제1 빈에 대해, 도 5의 (a) 또는 도 5의 (b)에 도시된 것과 같은 이웃이 적용가능하지 않은데, 그 이유는 제1 빈에 대해 보다 낮은 주파수 빈이 전혀 존재하지 않거나 모두 존재하는 것은 아니기 때문이다. 이들 특별한 경우 각각에 대해, 자신의 이웃이 정의될 수 있다. 추가의 실시예에서, 비존재 빈이 소정의 값으로 채워진다. 도 5의 (a)에 주어진 예시적인 이웃에 대해, 스펙트럼에서의 전송될 제1 빈에 대한 정의된 이웃이 도 6의 (a) 및 도 6의 (b)에 도시되어 있다. 작상은 스펙트럼의 나머지에 대해서와 동일한 컨텍스트 클래스 결정 함수를 사용하는 것을 가능하게 하기 위해 보다 높은 주파수 빈으로 이웃을 확장하는 것이다. 이것은 또한 동일한 컨텍스트 클래스 및 마침내 동일한 PDF가 사용될 수 있다는 것을 의미한다. 이웃의 크기가 단지 감소되는 경우(물론, 이것도 역시 옵션임), 이것은 가능하지 않을 것이다.

[0048] 리셋은 보통 새로운 스펙트럼이 코딩되기 전에 일어난다. 이미 언급한 바와 같이, 이것은 디코딩을 위한 전용 시작점을 가능하게 하기 위해 필요하다. 예를 들어, 디코딩 프로세스가 특정의 프레임/스펙트럼으로부터 시작하는 경우, 실제로 디코딩 프로세스는 원하는 시작 스펙트럼까지 선행 프레임을 성공적으로 디코딩하기 위해 마지막 리셋의 지점으로부터 시작해야만 한다. 이것은 리셋이 많이 일어날수록 디코딩을 위한 진입점이 많이 존재한다는 것을 의미한다. 그러나, 리셋 이후의 스펙트럼에서 코딩 효율이 더 작다.

[0049] 리셋이 일어난 후에, 이웃 정의에 이용가능한 선행 스펙트럼이 없다. 이것은 실제 스펙트럼의 선행 스펙트럼 빈만이 이웃에서 이용될 수 있다는 것을 의미한다. 그러나, 일반적인 절차가 변경되지 않을 수 있고, 동일한 "도구"가 사용될 수 있다. 다시, 제1 빈이 이전의 섹션에서 이미 설명한 바와 같이 상이하게 처리되어야만 한다.

[0050] 도 8에서, 예시적인 리셋 이웃 정의가 도시되어 있다. 이 정의는 USAC의 FD 모드에서의 리셋의 경우에 사용될 수 있다.

[0051] 도 8의 예(마지막 3개의 가능한 양자화된 값 또는 양자화 단계 1 이후의 값이 사용되는 경우 6개의 값에 의한 테이블의 양자화를 사용함)에 도시된 바와 같은 부가적인 컨텍스트 클래스의 수는 다음과 같다: 첫번째 빈에 대한 처리가 1개의 컨텍스트 클래스를 추가하고, 두번째 빈은 6개의 컨텍스트 클래스(양자화 단계 1이 사용된 후의 값)를 추가하며, 세번째 빈은 6개의 컨텍스트 클래스를 추가하고, 네번째 빈은 10개의 컨텍스트 클래스를 추가한다. 부가하여 2개의(낮은 및 높은) 분산 모드를 고려하는 경우, 이 컨텍스트 클래스의 수는 (이용가능한 정보가 없는 첫번째 빈에 대해서만) 거의 배로 된다(두번째 빈에 대해서는, 양자화 단계 1이 사용된 후의 빈에 대한 값이 두배로 되지 않음).

[0052] 이 예에서, 이 결과 리셋을 처리하기 위한 $1 + 6 + 2 \times 6 + 2 \times 10 = 39$ 개의 부가적인 컨텍스트 클래스가 얻어진다.

[0053] 매핑 블록(MAP)은 블록(COCL)에 의해 결정된 컨텍스트 분류, 예컨대 결정된 컨텍스트 클래스 번호를 받고, PDF 메모리(MEM1)로부터 대응하는 PDF를 선택한다. 이 단계에서, 2개 이상의 컨텍스트 클래스에 대해 단일 PDF를 사용함으로써 필요한 메모리 크기의 양을 추가로 감소시키는 것이 가능하다. 즉, 유사한 PDF를 갖는 컨텍스트

클래스는 결합 PDF를 사용할 수 있다. 이들 PDF는 충분히 큰 대표적인 데이터 집합을 사용하여 훈련 단계에서 사전 정의될 수 있다. 이 훈련은 유사한 PDF에 대응하는 컨텍스트 클래스가 식별되고 대응하는 PDF가 병합되는 최적화 단계를 포함할 수 있다. 데이터의 통계에 따라, 이 결과 메모리에 저장되어야만 하는 꽤 적은 수의 PDF가 얻어질 수 있다. USAC에 대한 예시적인 실험 버전에서, 822개의 컨텍스트 클래스로부터 64개의 PDF로의 매핑이 성공적으로 적용되었다.

- [0054] 이 매핑 함수(MAP)의 실현은, 컨텍스트 클래스의 수가 그다지 크지 않은 경우, 간단한 테이블 탐색일 수 있다. 수가 더 커지는 경우, 효율성을 위해 해시 테이블 검색이 적용될 수 있다.
- [0055] 앞서 언급한 바와 같이, 일반 모드 스위치(GMS)는 주파수 영역 모드(FD)와 가중된 선형 예측 변환 모드(wLPT) 사이의 전환을 가능하게 한다. 모드에 따라, 상이한 이웃이 사용될 수 있다. 도 5의 (a), 도 6의 (a), 도 6의 (b) 및 도 8에 나타난 예시적인 이웃은 실험에서 FD 모드에 대해 충분히 큰 것으로 밝혀졌다. 그러나, wLPT 모드에 대해서는, 도 5의 (b), 도 7의 (a), 도 7의 (b), 도 7의 (c) 및 도 9에 예시적으로 나타난 보다 큰 이웃이 유익한 것으로 밝혀졌다.
- [0056] 즉, wLPT 모드에서의 예시적인 리셋 처리가 도 9에 나타내어져 있다. 스펙트럼에서의 가장 낮은, 두번째로 가장 낮은, 세번째로 가장 낮은 및 네번째로 가장 낮은 빈에 대한 wLPT 모드에서의 예시적인 이웃이 도 7의 (a), 도 7의 (b), 도 7의 (c) 및 도 7의 (d)에, 각각, 나타내어져 있다. 그리고, 스펙트럼에서의 모든 다른 빈에 대한 wLPT 모드에서의 예시적인 이웃이 도 5의 (b)에 나타내어져 있다.
- [0057] 도 5의 (b)에 나타내어진 예시적인 이웃으로부터 얻어지는 컨텍스트 클래스의 수는 $3 \times 91 = 273$ 개의 컨텍스트 클래스이다. 인자 3은 현재 인코딩될 또는 현재 디코딩될 것과 동일한 주파수에서의 하나의 빈의 특별한 처리로부터 얻어진다. 이상에 주어진 식으로부터, 이웃에 있는 나머지 12개 빈에 대해 값 2, 1 또는 0을 갖는 $0.5 * ((12 * 12) + 3 * 12 + 2) = 91$ 개 조합의 빈의 수가 있다. 이웃의 분산이 임계값을 만족하거나 초과하는지에 따라 컨텍스트 클래스를 구분하는 실시예에서, 273개의 컨텍스트 클래스가 두배로 된다.
- [0058] 도 9에 나타난 예시적인 리셋 처리는 또한 다수의 컨텍스트 클래스를 추가할 수 있다.
- [0059] 실험에서 양호한 결과를 산출한 테스트된 예시적인 실시예에서, 이하의 표 3에 분류되어 있는 822개의 가능한 컨텍스트 클래스가 있다.

표 3

[0060] MPEG USAC CE 제안의 분류된 가능한 컨텍스트 클래스

모드	낮은 분산 모드	높은 분산 모드
FD 모드	84	84
리셋 후의 FD 모드		39
wLPT 모드	273	273
리셋 후의 wLPT 모드		69

- [0061] 테스트된 예시적인 실시예에서, 이들 822개의 가능한 컨텍스트 클래스는 64개의 PDF에 매핑된다. 이 매핑은, 앞서 기술한 바와 같이, 훈련 단계에서 결정된다.
- [0062] 얻어지는 64개의 PDF는 ROM 테이블에, 예컨대, 고정 소수점 산술 코더의 경우 16 비트 정확도로 저장되어야만 한다. 여기에 제안된 방식의 다른 이점은 나타나 있다: 배경 기술 부분에서 언급한 USAC 표준화의 현재 작업 중인 초안 버전에서, 쿼드러플(quadruple)(4개의 스펙트럼 빈을 포함하는 벡터)이 단일 코드워드를 사용하여 결합 코딩된다. 이 결과 벡터에서의 각각의 성분의 동적 범위가 아주 작더라도 아주 큰 코드북이 얻어진다(예컨대, 각각의 성분은 값 $[-4, \dots, 3] \rightarrow 8^4 = 4096$ 개의 가능한 상이한 벡터를 가질 수 있음). 그러나, 스칼라의 코딩은 아주 작은 코드북을 사용하여 각각의 빈에 대한 높은 동적 범위를 가능하게 한다. 테스트된 예시적인 실시예에서 사용된 코드북은 -15부터 +15까지의 빈에 대한 동적 범위를 제공하는 32개의 항목 및 Esc-코드워드(이 경우, 빈의 값은 이 범위 밖에 있음)를 가진다. 이것은 64×32 개의 16 비트 값만이 ROM 테이블에 저장되어야만 한다는 것을 의미한다.
- [0063] 이상에서, 선형 스펙트럼 계수를 사용하여 현재 스펙트럼 계수를 산술 인코딩하는 방법이 기술되어 있으며, 상기 선형 스펙트럼 계수는 이미 인코딩되어 있고, 상기 선형 스펙트럼 계수 및 상기 현재 스펙트럼 계수 둘다 비디오, 오디오 또는 음성 신호 샘플 값의 시간-주파수 변환을 양자화하는 것으로부터 얻어지는 하나 이상의 양

자화된 스펙트럼에 포함되어 있다. 일 실시예에서, 상기 방법은 선행 스펙트럼 계수를 처리하는 단계, 처리된 선행 스펙트럼 계수를 사용하여 적어도 2개의 상이한 컨텍스트 클래스(context class) 중 하나인 컨텍스트 클래스를 결정하는 단계, 결정된 컨텍스트 클래스 및 적어도 2개의 상이한 컨텍스트 클래스로부터 적어도 2개의 상이한 확률 밀도 함수로의 매핑을 사용하여 확률 밀도 함수를 결정하는 단계, 및 결정된 확률 밀도 함수에 기초하여 현재 스펙트럼 계수를 산술 인코딩하는 단계를 포함하고, 선행 스펙트럼 계수를 처리하는 단계는 선행 스펙트럼 계수를 비균등 양자화하는 단계를 포함한다.

[0064] 다른 예시적인 실시예에서, 이미 인코딩된 선행 스펙트럼 계수를 사용하여 현재 스펙트럼 계수를 산술 인코딩하는 장치는 처리 수단, 컨텍스트 클래스를 결정하는 제1 수단, 적어도 2개의 상이한 확률 밀도 함수를 저장하는 메모리, 확률 밀도를 검색하는 제2 수단, 및 산술 인코더를 포함한다.

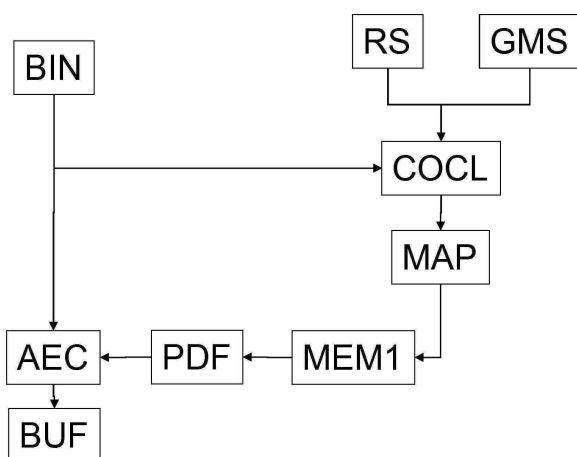
[0065] 이어서, 처리 수단은 이미 인코딩된 선행 스펙트럼 계수를 비균등 양자화함으로써 처리하도록 적응되고, 상기 제1 수단은 처리 결과를 사용하여 적어도 2개의 상이한 컨텍스트 클래스 중 하나인 컨텍스트 클래스를 결정하도록 적응된다. 메모리는 적어도 2개의 상이한 확률 밀도 함수 및 적어도 2개의 상이한 컨텍스트 클래스로부터 적어도 2개의 상이한 확률 밀도 함수로의 매핑 - 이 매핑은 결정된 컨텍스트 클래스에 대응하는 확률 밀도 함수를 검색할 수 있게 함 - 을 저장한다. 제2 수단은 결정된 컨텍스트 클래스에 대응하는 확률 밀도를, 메모리로부터, 검색하도록 적응되고, 산술 인코더는 검색된 확률 밀도 함수에 기초하여 현재 스펙트럼 계수를 산술 인코딩하도록 적응된다.

[0066] 이미 디코딩된 선행 스펙트럼 계수를 사용하여 현재 스펙트럼 계수를 산술 디코딩하는 장치의 대응하는 다른 예시적인 실시예가 있으며, 이 장치는 처리 수단, 컨텍스트 클래스를 결정하는 제1 수단, 적어도 2개의 상이한 확률 밀도 함수를 저장하는 메모리, 확률 밀도를 검색하는 제2 수단, 및 산술 디코더를 포함한다.

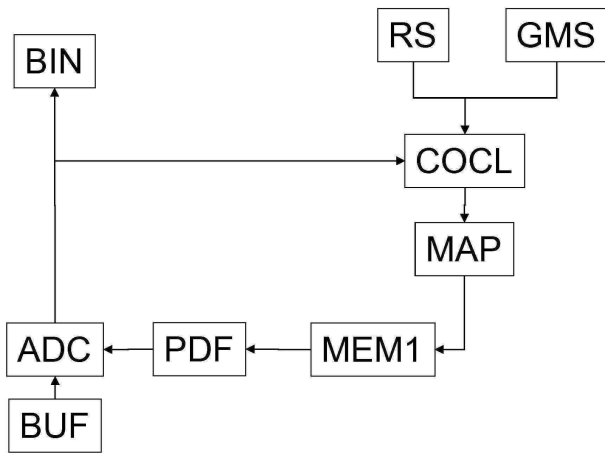
[0067] 이어서, 처리 수단은 이미 디코딩된 선행 스펙트럼 계수를 비균등 양자화함으로써 처리하도록 적응되고, 상기 제1 수단은 처리 결과를 사용하여 적어도 2개의 상이한 컨텍스트 클래스 중 하나인 컨텍스트 클래스를 결정하도록 적응된다. 메모리는 적어도 2개의 상이한 확률 밀도 함수 및 적어도 2개의 상이한 컨텍스트 클래스로부터 적어도 2개의 상이한 확률 밀도 함수로의 매핑 - 이 매핑은 결정된 컨텍스트 클래스에 대응하는 확률 밀도 함수를 검색할 수 있게 함 - 을 저장한다. 제2 수단은 결정된 컨텍스트 클래스에 대응하는 확률 밀도를, 메모리로부터, 검색하도록 적응되고, 산술 디코더는 검색된 확률 밀도 함수에 기초하여 현재 스펙트럼 계수를 산술 디코딩하도록 적응된다.

도면

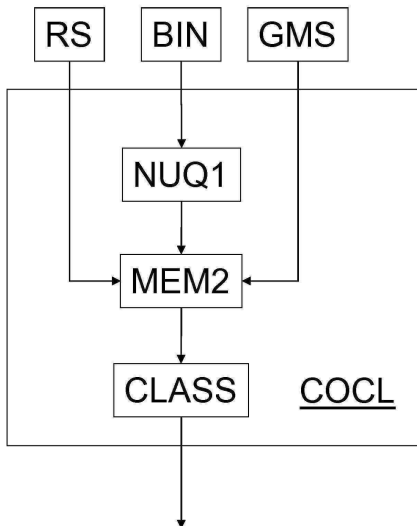
도면1



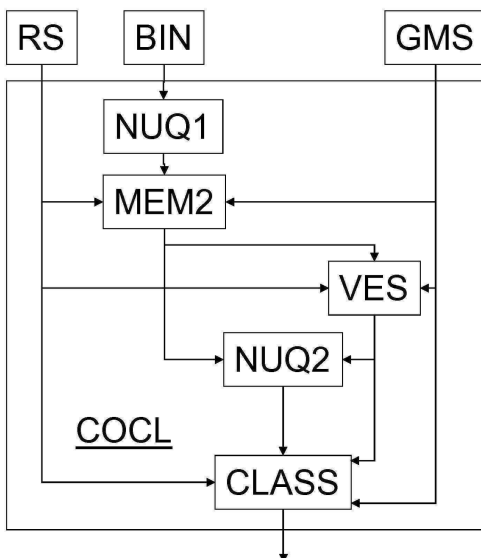
도면2



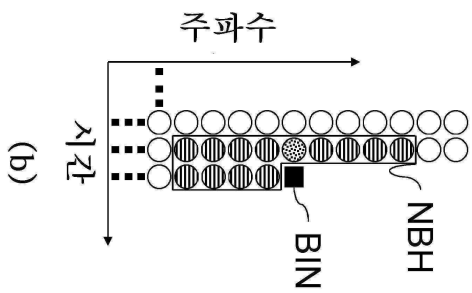
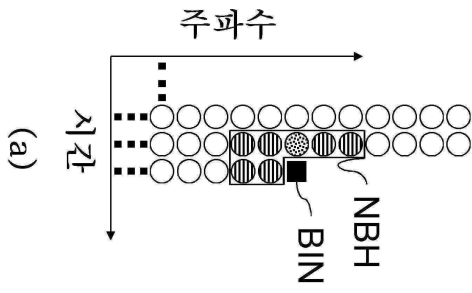
도면3



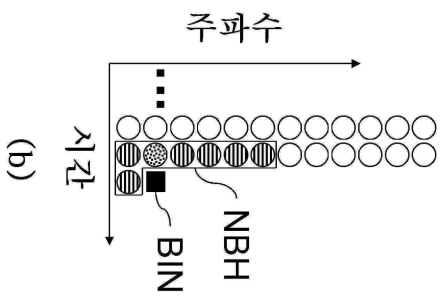
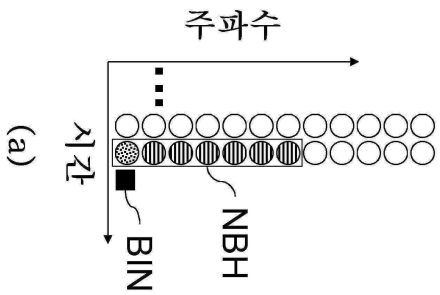
도면4



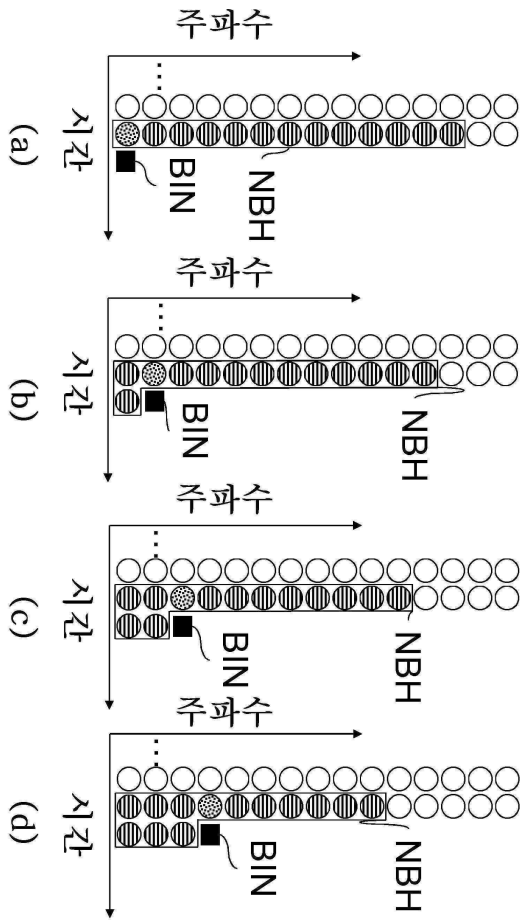
도면5



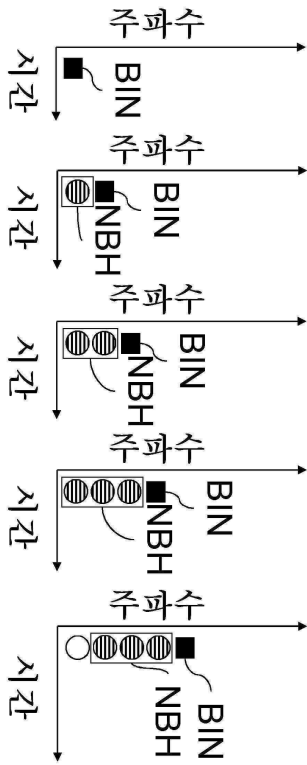
도면6



도면7



도면8



도면9

