



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2020년01월31일

(11) 등록번호 10-2071860

(24) 등록일자 2020년01월23일

(51) 국제특허분류(Int. Cl.)
G10L 19/16 (2013.01) G10L 25/60 (2013.01)

(52) CPC특허분류
G10L 19/167 (2013.01)
G10L 25/60 (2013.01)

(21) 출원번호 10-2016-7036002(분할)

(22) 출원일자(국제) 2014년01월15일

심사청구일자 2019년01월14일

(85) 번역문제출일자 2016년12월22일

(65) 공개번호 10-2017-0001716

(43) 공개일자 2017년01월04일

(62) 원출원 특허 10-2015-7019693

원출원일자(국제) 2014년01월15일

심사청구일자 2015년07월20일

(86) 국제출원번호 PCT/US2014/011681

(87) 국제공개번호 WO 2014/113471

국제공개일자 2014년07월24일

(30) 우선권주장
61/754,882 2013년01월21일 미국(US)

(뒷면에 계속)

(56) 선행기술조사문헌

KR1020060045675 A*

KR1020110022058 A*

Cossette S G Et al: "New Techniques for Audio Metadata Use and Distribution", Proceeding of the audio engineering convention society, AES 117th

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

돌비 레버러토리즈 라이쎄싱 코오폰레이션

미합중국, 캘리포니아 94103, 샌프란시스코, 마켓 스트리트 1275

돌비 인터네셔널 에이비

네덜란드 1101 씨엔 암스트레담 주이두스트 헤리 커베르그백 1-35 3이 아폴로 빌딩

(72) 발명자

리드밀러, 제프리

미국 캘리포니아 94103 샌프란시스코 포트레로 애비뉴 100 돌비 레버러토리즈 인크. 내

노크로스, 스캇 그레고리

미국 캘리포니아 94103 샌프란시스코 포트레로 애비뉴 100 돌비 레버러토리즈 인크. 내

로텐, 칼 조나스

스웨덴 스톡홀름 에스-113 30 개블레가탄 12에이 돌비 스웨덴 에이비 내

(74) 대리인

장훈

전체 청구항 수 : 총 15 항

심사관 : 옥윤철

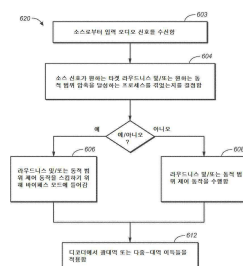
(54) 발명의 명칭 **상이한 재생 디바이스들에 걸친 라우드니스 및 동적 범위의 최적화**

(57) 요약

실시예들은 비트스트림에서 오디오 데이터와 연관된 메타데이터를 수신하며, 제 1 그룹의 오디오 재생 디바이스들에 대한 라우드니스 파라미터가 상기 비트스트림에서 이용 가능한지 여부를 결정하기 위해 상기 메타데이터를 분석하기 위한 방법 및 시스템에 관한 것이다. 상기 파라미터들이 상기 제 1 그룹을 위해 존재한다고 결정하는

(뒷면에 계속)

대표도 - 도16



것에 응답하여, 상기 시스템은 오디오를 렌더링하기 위해 상기 파라미터들 및 오디오 데이터를 사용한다. 상기 라우드니스 파라미터들이 상기 제 1 그룹을 위해 존재하지 않는다고 결정하는 것에 응답하여, 상기 시스템은 상기 제 1 그룹의 하나 이상의 특성들을 분석하며, 상기 하나 이상의 특성들에 기초하여 상기 파라미터를 결정한다.

(30) 우선권주장

61/809,250 2013년04월05일 미국(US)

61/824,010 2013년05월16일 미국(US)

명세서

청구범위

청구항 1

인코딩된 오디오 비트스트림의 하나 이상의 프레임들을 디코딩하기 위한 오디오 처리 유닛으로서, 상기 인코딩된 오디오 비트스트림은 오디오 데이터와 메타데이터 컨테이너를 포함하며, 상기 메타데이터 컨테이너는 하나 이상의 메타데이터 페이로드들을 포함하는, 상기 오디오 처리 유닛에 있어서:

상기 인코딩된 오디오 비트스트림의 하나 이상의 프레임들을 저장하기 위한 입력 버퍼;

상기 오디오 데이터를 디코딩하기 위해 버퍼 메모리에 연결된 오디오 디코더; 및

상기 오디오 디코더와 일체로 되거나 상기 오디오 디코더와 연결되어 상기 오디오 데이터를 파싱하도록 구성된 파서를 포함하며;

상기 메타데이터 컨테이너는 상기 메타데이터 컨테이너의 시작을 식별하는 싱크워드(syncword)로 시작하고, 상기 하나 이상의 메타데이터 페이로드들은 복수의 동적 범위 압축(DRC) 프로파일들로부터 선택된 동적 범위 압축 프로파일을 지정하는 파라미터를 포함하고, 상기 복수의 DRC 프로파일들 각각은 연관된 시간 상수들에 대한 고유 압축 곡선(unique curve with associated time constants)에 대응하고, 상기 하나 이상의 메타데이터 페이로드들에는 상기 하나 이상의 메타데이터 페이로드들을 해독(decrypt)하거나, 인증(authenticate)하거나, 유효화(validate)하는데 사용될 수 있는 보호 데이터가 후속하는, 오디오 처리 유닛.

청구항 2

제 1 항에 있어서,

상기 시간 상수들은 저속 및 고속 어택 시간 상수들(slow and fast attack time constants)과, 저속 및 고속 해제 시간 상수들(slow and fast release time constants)을 포함하는, 오디오 처리 유닛.

청구항 3

제 2 항에 있어서,

상기 고유 압축 곡선은 널-대역 범위(null-band range) 및 최대 부스트(maximum boost)에 의해 더 정의되는, 오디오 처리 유닛.

청구항 4

제 1 항에 있어서,

상기 파라미터는, 상기 오디오 데이터에 상대적으로 큰 압축(relatively heavy compression)이 적용되어야 함을 나타내는 휴대용 디바이스들 용 DRC 프로파일을 지정하는, 오디오 처리 유닛.

청구항 5

제 1 항에 있어서,

상기 메타데이터 컨테이너는, 보조데이터 필드(auxdata field), 추가 비트스트림 정보 필드(addbsi field) 및 그 조합으로 구성된 그룹으로부터 선택된 AC-3 또는 E-AC-3 예약 데이터 공간 내에 저장되는, 오디오 처리 유닛.

청구항 6

제 1 항에 있어서,

상기 하나 이상의 메타데이터 페이로드들은 오디오 프로그램의 측정된 라우드니스를 나타내는 데이터를 가진 프로그램 라우드니스 페이로드를 포함하는, 오디오 처리 유닛.

청구항 7

제 6 항에 있어서,

상기 프로그램 라우드니스 페이로드는 오디오 채널이 음성 다이얼로그(spoken dialogue)를 포함하는지 여부를 나타내는 필드를 포함하는, 오디오 처리 유닛.

청구항 8

제 6 항에 있어서,

상기 프로그램 라우드니스 페이로드는 상기 프로그램 라우드니스 페이로드에 포함된 라우드니스 데이터를 생성하는 데에 사용된 라우드니스 측정 방법을 나타내는 필드를 포함하는, 오디오 처리 유닛.

청구항 9

제 6 항에 있어서,

상기 오디오 처리 유닛은 상기 프로그램 라우드니스 페이로드를 이용하여 적응적 라우드니스 처리를 수행하도록 구성되는, 오디오 처리 유닛.

청구항 10

제 1 항에 있어서,

상기 인코딩된 오디오 비트스트림은 AC-3 비트스트림 또는 E-AC-3 비트스트림인, 오디오 처리 유닛.

청구항 11

제 1 항에 있어서,

각각의 메타데이터 페이로드는 각각의 메타데이터 페이로드의 시작부에 위치된 고유한 페이로드 식별자를 포함하는, 오디오 처리 유닛.

청구항 12

인코딩된 오디오 비트스트림의 적어도 한 프레임을 기록한 비-일시적 컴퓨터 판독가능한 기록 매체로서, 상기 적어도 한 프레임은:

오디오 데이터; 및

메타데이터 컨테이너를 포함하며, 상기 메타데이터 컨테이너는 하나 이상의 메타데이터 페이로드들 및 상기 하나 이상의 메타데이터 페이로드들을 해독(decrypt)하거나, 인증(authenticate)하거나, 유효화(validate)하는데 사용될 수 있는 보호 데이터를 포함하고;

상기 메타데이터 컨테이너는 상기 메타데이터 컨테이너의 시작을 식별하는 싱크워드로 시작하고, 상기 하나 이상의 메타데이터 페이로드들은 복수의 동적 범위 압축(DRC) 프로파일들로부터 선택된 동적 범위 압축 프로파일을 지정하는 파라미터를 포함하고, 상기 복수의 DRC 프로파일들 각각은 시간 상수들에 대한 고유 압축 곡선(unique compression curve with time constants)에 대응하고, 상기 하나 이상의 메타데이터 페이로드들에는 상기 보호 데이터가 후속하는, 비-일시적 컴퓨터 판독가능한 기록 매체.

청구항 13

제 12 항에 있어서,

상기 시간 상수들은 저속 및 고속 어택 시간 상수들과, 저속 및 고속 해제 시간 상수들을 포함하는, 비-일시적 컴퓨터 판독가능한 기록 매체.

청구항 14

제 12 항에 있어서,

상기 고유 압축 곡선은 널-대역 범위(null-band range) 및 최대 부스트(maximum boost)에 의해 더 정의되는,

비-일시적 컴퓨터 판독가능한 기록 매체.

청구항 15

제 12 항에 있어서,

상기 파라미터는, 상기 오디오 데이터에 상대적으로 큰 압축(relatively heavy compression)이 적용되어야 함을 나타내는 휴대용 디바이스들 용 DRC 프로파일을 지정하는, 비-일시적 컴퓨터 판독가능한 기록 매체.

발명의 설명

기술 분야

[0001] 관련 출원들에 대한 상호 참조

[0002] 본 출원은, 모두 참조로서 여기에 통합되는, 2013년 1월 21일에 출원된 미국 가출원 번호 제61/754,882호; 2013년 4월 5일에 출원된 미국 가출원 번호 제61/809,250호; 및 2013년 5월 16일에 출원된 미국 가출원 번호 제61/824,010호에 대한 우선권을 주장한다.

[0003] 하나 이상의 실시예들은 일반적으로 오디오 신호 프로세싱에 관한 것이며, 보다 구체적으로 재생 환경들 및 디바이스들에 기초한 오디오 콘텐츠의 라우드니스(loudness) 및 동적 범위 특성들을 나타내는 메타데이터를 갖는 오디오 데이터 비트스트림들을 프로세싱하는 것에 관한 것이다.

배경 기술

[0004] 배경 부분에서 논의된 주제는 단지 배경 부분에서 언급되고 있다는 이유만으로 종래 기술인 것으로 추정되어서는 안 된다. 유사하게, 배경 부분에 언급되거나 또는 배경 부분의 주제와 연관된 문제점은 종래 기술에서 이전에 인식되어 온 것으로 추정되어서는 안 된다. 배경 부분에서의 주제는 단지 상이한 접근법들을 나타내는 것일 뿐이며, 이것은 본질적으로 및 그 자체로 또한 발명들이 될 수 있다.

[0005] 오디오 신호의 동적 범위는 신호에 포함된 사운드의 최대 및 최소 가능한 값들 사이의 비이며, 보통 데시벨(베이스-10) 값으로서 측정된다. 많은 오디오 프로세싱 시스템들에서, 동적 범위 제어(또는 동적 범위 압축, DRC)는 넓은 동적 범위 소스 콘텐츠를, 전자 장비를 사용하여 보다 쉽게 저장되며 재생될 수 있는 보다 좁은 레코딩된 동적 범위에 맞추기 위해 큰 사운드들의 레벨을 감소시키고 및/또는 조용한 사운드들의 레벨을 증폭시키기 위해 사용된다. 오디오/비주얼(AV) 콘텐츠에 대해, DRC 메커니즘을 통한 압축을 위해 "널(null)" 포인트를 정의하도록 다이얼로그 기준 레벨이 사용될 수 있다. DRC는 다이얼로그 기준 레벨 아래의 콘텐츠를 부스팅(boost)하며 기준 레벨 위의 콘텐츠를 컷팅하도록 동작한다.

[0006] 알려진 오디오 인코딩 시스템에서, 오디오 신호와 연관된 메타데이터는 콘텐츠의 유형 및 의도된 사용에 기초하여 DRC 레벨을 설정하기 위해 사용된다. DRC 모드는 오디오 신호에 적용된 압축의 양을 설정하며 디코더의 출력 기준 레벨을 정의한다. 이러한 시스템들은 인코더로 프로그램되며 사용자에게 의해 선택되는 두 개의 DRC 레벨 설정들에 제한될 수 있다. 예를 들면, -31 dB(라인)의 Dialnorm(다이얼로그 정규화) 값이 AVR 또는 전체 동적 범위 가능 디바이스들 상에서 재생되는 콘텐츠를 위해 전통적으로 사용되며, -20 dB(RF)의 Dialnorm 값이 텔레비전 세트들 또는 유사한 디바이스들 상에서 재생된 콘텐츠를 위해 사용된다. 이러한 유형의 시스템은 단일 오디오 비트스트림이, 두 개의 상이한 세트들의 DRC 메타데이터들의 사용을 통해 두 개의 일반적이지만 매우 상이한 재생 시나리오들에서 사용되도록 허용한다. 그러나, 이러한 시스템들은 사전 설정된 Dialnorm 값들에 제한되며 디지털 미디어 및 인터넷-기반 스트리밍 기술의 출현을 통해 현재 가능한 매우 다양한 상이한 재생 디바이스들 및 청취 환경들에서의 재생에 대해 최적화되지 않았다.

[0007] 현재 메타데이터-기반 오디오 인코딩 시스템들에서, 오디오 데이터의 스트림은 오디오 콘텐츠(예로서, 오디오 콘텐츠의 하나 이상의 채널들) 및 상기 오디오 콘텐츠의 적어도 하나의 특성을 나타내는 메타데이터 양쪽 모두를 포함할 수 있다. 예를 들면, AC-3 비트스트림에서, 청취 환경에 전달된 프로그램의 사운드를 변경할 때 사용하는 데 특히 의도된 여러 개의 오디오 메타데이터 파라미터들이 있다. 상기 메타데이터 파라미터들 중 하나는 오디오 프로그램에서 발생한 다이얼로그의 평균 라우드니스 레벨(또는 콘텐츠의 평균 라우드니스)을 나타내는 Dialnorm 파라미터이며, 오디오 재생 신호 레벨을 결정하기 위해 사용된다.

[0008] 상이한 오디오 프로그램 세그먼트들(각각은 상이한 Dialnorm 파라미터를 가진다)의 시퀀스를 포함한 비트스트림의 재생 동안, AC-3 디코더는 세그먼트의 다이얼로그의 지각된 라우드니스가 일관된 레벨에 있도록 세그먼트의

재생 레벨 또는 라우드니스를 변경하는 라우드니스 프로세싱의 유형을 수행하도록 각각의 세그먼트의 Dialnorm 파라미터를 사용한다. 인코딩된 오디오 아이템들의 시퀀스에서 각각의 인코딩된 오디오 세그먼트(아이템)는 (일반적으로) 상이한 Dialnorm 파라미터를 가지며, 재생 동안 아이템들의 상이한 레벨들에 대한 상이한 이득 양들의 적용을 요구할 수 있지만, 디코더는 각각의 아이템에 대한 다이얼로그의 라우드니스 또는 재생 레벨이 동일하거나 또는 매우 유사하도록 아이템들의 각각의 레벨을 스케일링할 것이다.

[0009] 일부 실시예들에서, Dialnorm 파라미터는 사용자에게 의해 설정되며, 어떤 값도 사용자에게 의해 설정되지 않는다면 디폴트 Dialnorm 값이 있지만, 자동으로 생성되지 않는다. 예를 들면, 콘텐츠 생성기는 AC-3 인코더의 외부에 있는 디바이스로 라우드니스 측정들을 할 수 있으며 그 후 Dialnorm 값을 설정하기 위해 그 결과(오디오 프로그램의 음성 다이얼로그의 라우드니스를 나타내는)를 인코더에 전송하여 Dialnorm 값을 설정하게 한다. 따라서, Dialnorm 파라미터를 정확하게 설정하기 위해서는 콘텐츠 생성기를 의존해야한다.

[0010] AC-3 비트스트림에서 Dialnorm 파라미터가 부정확할 수 있는 여러 개의 상이한 이유들이 있다. 첫 번째로, 각각의 AC-3 인코더는 Dialnorm 값이 콘텐츠 생성기에 의해 설정되지 않는다면 비트스트림의 생성 동안 사용되는 디폴트 Dialnorm 값을 가진다. 이러한 디폴트 값은 오디오의 실제 다이얼로그 라우드니스 레벨과 상당히 상이할 수 있다. 두 번째로, 콘텐츠 생성기가 라우드니스를 측정하여 그에 따라 Dialnorm 값을 설정하더라도, 추천된 라우드니스 측정 방법을 따르지 않는 라우드니스 측정 알고리즘 또는 미터가 사용되었을 수 있어, 부정확한 Dialnorm 값을 야기할 수 있다. 세 번째로, AC-3 비트스트림이 측정된 Dialnorm 값으로 생성되어 콘텐츠 생성기에 의해 정확하게 설정되더라도, 그것은 비트스트림의 송신 및/또는 저장 동안 중간 모듈에 의해 부정확한 값으로 변경되었을 수 있다. 예를 들면, 텔레비전 방송 애플리케이션들에서 AC-3 비트스트림들이 부정확한 Dialnorm 메타데이터 정보를 사용하여 디코딩되고, 변경되며 그 후 재-인코딩되는 것은 드물지 않다. 따라서, AC-3 비트스트림에 포함된 Dialnorm 값은 부정확하거나 또는 오류가 있을 수 있으며 그러므로 청취 경험의 품질에 대한 부정적인 영향을 줄 수 있다.

[0011] 뿐만 아니라, Dialnorm 파라미터는 대응하는 오디오 데이터의 라우드니스 프로세싱 상태(예로서, 어떤 유형(들)의 라우드니스 프로세싱이 오디오 데이터에 대해 수행되어 왔는지)를 나타내지 않는다. 부가적으로, 돌비 디지털(DD) 및 돌비 디지털 플러스(DD+) 시스템들에서의 시스템들과 같은, 현재 배치된 라우드니스 및 DRC 시스템들은 소비자의 거실 또는 영화관에서 AV 콘텐츠를 렌더링하도록 설계되었다. 다른 환경들 및 청취 장비(예로서, 이동 디바이스)에서의 재생에 이러한 콘텐츠를 맞추기 위해서는, 상기 청취 환경에 AV 콘텐츠를 맞추기 위해 재생 디바이스에서 후-처리가 '맹목적으로' 적용되어야 한다. 다시 말해서, 후-처리(또는 디코더)는 수신된 콘텐츠의 라우드니스 레벨이 특정한 레벨(예로서, -31 또는 -20 dB)에 있다고 추정하며 상기 후-처리는 특정한 디바이스에 적합한 미리 결정된 고정 타겟 레벨로 상기 레벨을 설정한다. 상기 추정된 라우드니스 레벨 또는 미리 결정된 타겟 레벨이 부정확하다면, 후-처리는 그것의 의도된 효과와는 반대의 효과를 가질 수 있으며; 즉, 후-처리는 사용자에게 덜 바람직하게 출력 오디오를 만들 수 있다.

[0012] 개시된 실시예들은 AC-3 비트스트림, E-AC-3 비트스트림, 또는 돌비 E 비트스트림과 함께 사용하는데 제한되지는 않지만, 편리함을 위해, 이러한 비트스트림들은 라우드니스 프로세싱 상태 메타데이터를 포함하는 시스템과 함께 논의될 것이다. 돌비, 돌비 디지털, 돌비 디지털 플러스, 및 돌비 E는 돌비 래버러토리스 라이선싱 코퍼레이션(Dolby Laboratories Licensing Corporation)의 상표들이다. 돌비 래버러토리스는 돌비 디지털 및 돌비 디지털 플러스로서 각각 알려진 AC-3 및 E-AC-3의 독점 구현들을 제공한다.

발명의 내용

해결하려는 과제

[0013] 본 발명은 상이한 재생 디바이스들에 걸쳐 라우드니스 및 동적 범위를 최적화하기 위한 개선된 시스템 및 방법을 제공한다.

과제의 해결 수단

[0014] 실시예들은 오디오 데이터와 연관된 메타데이터를 포함하는 비트스트림을 수신함으로써, 상기 오디오 데이터를 디코딩하며, 제 1 그룹의 오디오 재생 디바이스들에 대한 라우드니스 파라미터가 상기 비트스트림에서 이용 가능한지 여부를 결정하기 위해 상기 비트스트림내의 메타데이터를 분석하기 위한 방법에 관한 것이다. 상기 제 1 그룹에 대한 상기 파라미터들이 존재한다는 결정에 응답하여, 프로세싱 구성요소는 오디오를 렌더링하기 위해 상기 파라미터들 및 오디오 데이터를 사용한다. 상기 제 1 그룹에 대한 상기 라우드니스 파라미터들이 존재하지

않는다는 결정에 응답하여, 상기 프로세싱 구성요소는 상기 제 1 그룹의 하나 이상의 특성들을 분석하며 상기 하나 이상의 특성들에 기초하여 상기 파라미터를 결정한다. 상기 방법은 재생을 위해 상기 오디오를 렌더링하는 다운스트림 모듈로 상기 파라미터 및 오디오 데이터를 송신함으로써 오디오를 렌더링하기 위해 상기 파라미터 및 오디오 데이터를 추가로 사용할 수 있다. 상기 파라미터 및 오디오 데이터는 또한 상기 파라미터 및 오디오 데이터에 기초하여 상기 오디오 데이터를 렌더링함으로써 오디오를 렌더링하기 위해 사용될 수 있다.

[0015] 실시예에서, 상기 방법은 또한 상기 수신된 오디오 스트림을 렌더링할 출력 디바이스를 결정하는 단계, 및 상기 출력 디바이스가 상기 제 1 그룹의 오디오 재생 디바이스들에 속하는지 여부를 결정하는 단계를 포함하며, 여기에서 상기 제 1 그룹의 오디오 재생 디바이스들에 대한 라우드니스 파라미터가 이용 가능하지 여부를 결정하기 위해 상기 스트림내의 메타데이터를 분석하는 단계는 상기 출력 디바이스가 상기 제 1 그룹의 오디오 재생 디바이스들에 속하는 것을 결정하는 단계 후 실행된다. 일 실시예에서, 상기 출력 디바이스가 상기 제 1 그룹의 오디오 재생 디바이스들에 속하는 것을 결정하는 단계는: 상기 출력 디바이스의 아이덴티티를 나타내거나 또는 상기 출력 디바이스를 포함하는 디바이스들의 그룹의 아이덴티티를 나타내는 표시를 상기 출력 디바이스에 접속된 모듈로부터 수신하는 단계, 및 상기 수신된 표시에 기초하여 상기 출력 디바이스가 상기 제 1 그룹의 오디오 재생 디바이스들에 속하는 것을 결정하는 단계를 포함한다.

[0016] 실시예들은 또한 상기 인코딩 방법 실시예들에서 설명된 동작들을 수행하는 프로세싱 구성요소들을 포함하는 장치 또는 시스템에 관한 것이다.

[0017] 실시예들은 또한 오디오 데이터 및 상기 오디오 데이터와 연관된 메타데이터를 수신함으로써 상기 오디오 데이터를 디코딩하고, 제 1 그룹의 오디오 디바이스들에 대한 라우드니스 파라미터들과 연관된 라우드니스 정보가 상기 스트림에서 이용 가능하지 여부를 결정하기 위해 상기 비트스트림내의 메타데이터를 분석하며, 상기 제 1 그룹에 대한 라우드니스 정보가 존재하는 것에 대한 결정에 응답하여, 상기 스트림으로부터 라우드니스 정보를 결정하고, 오디오를 렌더링하는데 사용하기 위해 상기 오디오 데이터 및 라우드니스 정보를 송신하거나 또는, 상기 제 1 그룹에 대한 라우드니스 정보가 존재하지 않는다면, 출력 프로파일과 연관된 라우드니스 정보를 결정하며, 오디오를 렌더링하는데 사용하기 위해 상기 출력 프로파일에 대한 상기 결정된 라우드니스 정보를 송신하는 방법에 관한 것이다. 일 실시예에서, 출력 프로파일과 연관된 라우드니스 정보를 결정하는 단계는 상기 출력 프로파일의 특성들을 분석하는 단계, 상기 특성들에 기초하여 상기 파라미터들을 결정하는 단계를 포함하며, 결정된 라우드니스 정보를 송신하는 단계는 상기 결정된 파라미터들을 송신하는 단계를 포함한다. 상기 라우드니스 정보는 출력 프로파일에 대한 라우드니스 파라미터들 또는 그것의 특성들을 포함할 수 있다. 실시예에서, 상기 방법은 송신될 저 비트 레이트 인코딩 스트림을 결정하는 단계를 더 포함할 수 있으며, 여기에서 상기 라우드니스 정보는 하나 이상의 출력 프로파일들에 대한 특성들을 포함한다.

[0018] 실시예들은 또한 상기 디코딩 방법 실시예들에서 설명된 동작들을 수행하는 프로세싱 구성요소들을 포함하는 장치 또는 시스템에 관한 것이다.

발명의 효과

[0019] 본 발명에 따른 시스템은 데이터 레이트를 감소시키기 위해 원래 이득들의 파라미터화를 통한 인코더로부터의 제어 하에서, 인코더에서 라우드니스 제어 및 동적 범위 요건들에 기초하여 적절한 이득들을 발생시키거나 또는 디코더에서 이득들을 발생시키며, 또한 다른 메타데이터(내부 또는 외부) 파라미터들이 라우드니스 및 동적 범위 이득들 및/또는 프로파일들을 적절히 제어하기 위해 사용되도록 허용한다.

도면의 간단한 설명

[0020] 다음의 도면들에서, 유사한 참조 번호들은 유사한 요소들을 나타내기 위해 사용된다. 다음의 도면들은 다양한 예들을 묘사하지만, 여기에 설명된 구현들은 도면들에 묘사된 예들에 제한되지 않는다.

도 1은 일부 실시예들에서, 라우드니스 및 동적 범위의 최적화를 수행하도록 구성된 오디오 프로세싱 시스템의 실시예의 블록도.

도 2는 일부 실시예들에서, 도 1의 시스템에서의 사용을 위한 인코더의 블록도.

도 3은 일부 실시예들에서, 도 1의 시스템에서의 사용을 위한 디코더의 블록도.

도 4는 세그먼트들로 분할된 AC-3 프레임의 다이어그램.

- 도 5는 세그먼트들로 분할된 AC-3 프레임의 동기화 정보(SI) 세그먼트의 다이어그램.
- 도 6은 세그먼트들로 분할된 AC-3 프레임의 비트스트림 정보(BSI) 세그먼트의 다이어그램.
- 도 7은 세그먼트들로 분할된 E-AC-3 프레임의 다이어그램.
- 도 8은 일부 실시예들에서, 인코딩된 비트스트림의 특정한 프레임들 및 메타데이터의 포맷을 예시한 표.
- 도 9는 일부 실시예들에서, 라우드니스 프로세싱 상태 메타데이터의 포맷을 예시한 표.
- 도 10은 일부 실시예들에서, 라우드니스 및 동적 범위의 최적화를 수행하도록 구성될 수 있는 도 1의 오디오 프로세싱 시스템의 보다 상세한 블록도.
- 도 11은 예시적인 사용 케이스에서 다양한 재생 디바이스들 및 배경 청취 환경들을 위한 상이한 동적 범위 요건들을 예시한 표.
- 도 12는 실시예에서, 동적 범위 최적화 시스템의 블록도.
- 도 13은 실시예에서, 다양한 상이한 재생 디바이스 클래스들에 대한 상이한 프로파일들 사이에서의 인터페이스를 예시한 블록도.
- 도 14는 실시예에서, 복수의 정의된 프로파일들에 대한 장기 라우드니스와 단기 동적 범위 사이에서의 상관관계를 예시하는 표.
- 도 15는 실시예에서, 상이한 유형들의 오디오 콘텐츠에 대한 라우드니스 프로파일들의 예들을 예시한 도면.
- 도 16은 실시예에서, 재생 디바이스들 및 애플리케이션들에 걸쳐 라우드니스 및 동적 범위를 최적화하는 방법을 예시하는 흐름도.

발명을 실시하기 위한 구체적인 내용

[0021] 정의들 및 명명법

[0022] 청구항들을 포함하여, 본 개시 전체에 걸쳐, 신호 또는 데이터"에 대한" 동작을 수행하는 표현(예로서, 신호 또는 데이터를 필터링하고, 스케일링하고, 변환하거나 또는 그것에 이득을 적용하는)은 신호 또는 데이터에 대해, 또는 신호 또는 데이터의 프로세싱된 버전에 대해(예로서, 동작의 수행 이전에 예비 필터링 또는 전처리를 받은 신호의 버전에 대해) 직접 동작을 수행하는 것을 나타내기 위해 광범위한 의미로 사용된다. 표현("시스템")은 디바이스, 시스템, 또는 서브시스템을 나타내기 위해 광범위한 의미로 사용된다. 예를 들면, 디코더를 구현하는 서브시스템은 디코더 시스템이라 할 수 있으며, 이러한 서브시스템을 포함한 시스템(예로서, 다수의 입력들에 응답하여 X개의 출력 신호들을 발생시키는 시스템, 여기에서 서브시스템은 M의 입력들을 발생시키며 다른 X-M 입력들이 외부 소스로부터 수신된다)은 또한 디코더 시스템이라 할 수 있다. 용어("프로세서")는 데이터(예로서, 오디오, 또는 비디오 또는 다른 이미지 데이터)에 대한 동작들을 수행하도록 프로그램가능한 또는 다른 방식으로 구성 가능한(예로서, 소프트웨어 또는 펌웨어로) 시스템 또는 디바이스를 나타내기 위해 광범위한 의미로 사용된다. 프로세서들의 예들은 필드-프로그램 가능한 게이트 어레이(또는 다른 구성 가능한 집적 회로 또는 칩 셋), 오디오 또는 다른 사운드 데이터에 대한 파이프라인 프로세싱을 수행하도록 프로그램되거나 및/또는 다른 방식으로 구성된 디지털 신호 프로세서, 프로그램 가능한 범용 프로세서 또는 컴퓨터, 및 프로그램 가능한 마이크로프로세서 칩 또는 칩 셋을 포함한다.

[0023] 표현들 "오디오 프로세서" 및 "오디오 프로세싱 유닛"은 상호 교환 가능하게 사용되며, 광범위한 의미로는 오디오 데이터를 프로세싱하도록 구성된 시스템을 나타내기 위해 사용된다. 오디오 프로세싱 유닛들의 예들은, 이에 제한되지 않지만, 인코더들(예로서, 트랜스코더들), 디코더들, 코덱들, 전-처리 시스템들, 후-처리 시스템들, 및 비트스트림 프로세싱 시스템들(때때로 비트스트림 프로세싱 툴이라 함)을 포함한다. 표현 "프로세싱 상태 메타데이터"(예로서, 표현 "라우드니스 프로세싱 상태 메타데이터")은 대응하는 오디오 데이터(또한 프로세싱 상태 메타데이터를 포함하는 오디오 데이터 스트림의 오디오 콘텐츠)로부터의 별개의 상이한 데이터를 나타낸다. 프로세싱 상태 메타데이터는 오디오 데이터와 연관되고, 대응하는 오디오 데이터의 라우드니스 프로세싱 상태(예로서, 어떤 유형(들)의 프로세싱이 오디오 데이터에 대해 이미 수행되었는지)를 나타내며, 선택적으로 또한 오디오 데이터의 적어도 하나의 특징 또는 특성을 나타낸다. 일부 실시예에서, 오디오 데이터와 프로세싱 상태 메타데이터의 연관은 시간-동기식이다. 따라서, 현재(가장 최근에 수신된 또는 업데이트된) 프로세싱 상태 메타데이터는 대응하는 오디오 데이터가 표시된 유형(들)의 오디오 데이터 프로세싱의 결과들을 동시에 포함한다는

것을 나타낸다. 몇몇 경우들에서, 프로세싱 상태 메타데이터는 프로세싱 이력 및/또는 표시된 유형들의 프로세싱에서 사용되고 및/또는 그로부터 얻어지는 파라미터들의 일부 또는 모두를 포함할 수 있다. 부가적으로, 프로세싱 상태 메타데이터는 오디오 데이터로부터 계산되거나 또는 그로부터 추출되어 온 대응하는 오디오 데이터의 적어도 하나의 특징 또는 특성을 포함할 수 있다. 프로세싱 상태 메타데이터는 또한 대응하는 오디오 데이터의 어떤 프로세싱에도 관련되지 않거나 그로부터 얻어지지 않는 다른 메타데이터를 포함할 수 있다. 예를 들면, 제 3 자 데이터, 추적 정보, 식별자들, 독점(proprietary) 또는 표준 정보, 사용자 주석 데이터, 사용자 선호 데이터 등이 다른 오디오 프로세싱 유닛들로 전하기 위해 특정의 오디오 프로세싱 유닛에 의해 추가될 수 있다.

[0024] 표현 "라우드니스 프로세싱 상태 메타데이터" (또는 "LPSM")은 대응하는 오디오 데이터의 라우드니스 프로세싱 상태(예로서, 어떤 유형(들)의 라우드니스 프로세싱이 오디오 데이터에 대해 수행되었는지)를 나타내고 또한 선택적으로 대응하는 오디오 데이터의 적어도 하나의 특징 또는 특성(예로서, 라우드니스)을 나타내는 프로세싱 상태 메타데이터를 표시한다. 라우드니스 프로세싱 상태 메타데이터는 라우드니스 프로세싱 상태 메타데이터가 아닌(즉, 그것이 단독으로 고려될 때) 데이터(예로서, 다른 메타데이터)를 포함할 수 있다. 용어 "결합하다" 또는 "결합된"은 직접 또는 간접 접속 중 하나를 의미하기 위해 사용된다.

[0025] 시스템들 및 방법들은 상이한 타겟 라우드니스 값들을 요구하거나 또는 사용하며 상이한 동적 범위 능력들을 가진 다양한 디바이스들에 걸쳐 오디오의 라우드니스 및 동적 범위를 비-파괴적으로 정규화하는 오디오 인코더/디코더에 대해 설명된다. 일부 실시예들에 따른 방법들 및 기능 구성요소들은 하나 이상의 디바이스 프로파일들에 대해 오디오 콘텐츠에 대한 정보를 인코더에서 디코더로 전송한다. 디바이스 프로파일은 하나 이상의 디바이스들에 대해 원하는 타겟 라우드니스 및 동적 범위를 특정한다. 시스템은 확장 가능하며, 따라서 상이한 "공칭" 라우드니스 타겟들을 가진 새로운 디바이스 프로파일들이 지원될 수 있다.

[0026] 실시예에서, 상기 시스템은 데이터 레이트를 감소시키기 위해 원래 이득들의 파라미터화를 통한 인코더로부터의 제어 하에서, 인코더에서 라우드니스 제어 및 동적 범위 요건들에 기초하여 적절한 이득들을 발생시키거나 또는 디코더에서 상기 이득들을 발생시킨다. 상기 동적 범위 시스템은 라우드니스 제어를 구현하기 위한 두 개의 메커니즘들을 포함한다: 오디오가 어떻게 재생될 지에 대해서 콘텐츠 생성기 제어를 제공하는 예술적 동적 범위 프로파일, 및 오버로딩이 다양한 재생 프로파일들에 대해 발생하지 않음을 보장하기 위한 별개의 보호 메커니즘. 상기 시스템은 또한 다른 메타데이터(내부 또는 외부) 파라미터들이 라우드니스 및 동적 범위 이득들 및/또는 프로파일들을 적절히 제어하기 위해 사용될 수 있도록 구성된다. 상기 디코더는 디코더-측 라우드니스 및 동적 범위 설정들/프로세싱을 레버리징할 n-채널 보조 입력을 지원하도록 구성된다.

[0027] 일부 실시예들에서, 라우드니스 프로세싱 상태 메타데이터(LPSM)는 또한 다른 세그먼트들(오디오 데이터 세그먼트들)에 오디오 데이터를 포함하는 오디오 비트스트림의 메타데이터 세그먼트들의 하나 이상의 예약(reserved) 필드들(또는 슬롯들)에 내장된다. 예를 들면, 비트스트림의 각각의 프레임의 적어도 하나의 세그먼트는 LPSM을 포함하며, 프레임의 적어도 하나의 다른 세그먼트는 대응하는 오디오 데이터(즉, 그것의 라우드니스 프로세싱 상태 및 라우드니스가 LPSM에 의해 나타내어지는 오디오 데이터)를 포함한다. 일부 실시예들에서, LPSM의 데이터 볼륨은 오디오 데이터를 운반하도록 할당된 비트 레이트에 영향을 미치지 않고 운반되기에 충분히 작을 수 있다.

[0028] 오디오 데이터 프로세싱 체인에서 라우드니스 프로세싱 상태 메타데이터를 전달하는 것은 둘 이상의 오디오 프로세싱 유닛들이 프로세싱 체인(또는 콘텐츠 라이프사이클) 전체에 걸쳐 서로 협력하여 작동하도록 요구할 때 특히 유용하다. 오디오 비트스트림에 라우드니스 프로세싱 상태 메타데이터를 포함하지 않으면, 품질, 레벨 및 공간적 열화와 같은 미디어 프로세싱 문제점들이 예를 들면, 둘 이상의 오디오 코덱들이 체인에서 이용되며 단일-단 볼륨 레벨링이 미디어 소비 디바이스(또는 비트스트림의 오디오 콘텐츠의 렌더링 포인트)에 대해 비트스트림의 여정 동안 1회 이상 적용될 때 발생할 수 있다.

[0029] 라우드니스 및 동적 범위 메타데이터 프로세싱 시스템

[0030] 도 1은 특정한 메타데이터 프로세싱(예로서, 전-처리 및 후-처리) 구성요소들을 사용한 일부 실시예들에서, 라우드니스 및 동적 범위의 최적화를 수행하도록 구성될 수 있는 오디오 프로세싱 시스템의 실시예의 블록도이다. 도 1은 예시적인 오디오 프로세싱 체인(오디오 데이터 프로세싱 시스템)을 예시하며, 여기에서 시스템의 요소들 중 하나 이상은 본 발명의 실시예에 따라 구성될 수 있다. 도 1의 시스템(10)은 도시된 바와 같이 함께 결합된, 다음의 요소들을 포함한다: 전-처리 유닛(12), 인코더(14), 신호 분석 및 메타데이터 교정 유닛(16), 트랜스코더(18), 디코더(20), 및 후-처리 유닛(24). 도시된 시스템상에서의 변화로서, 요소들 중 하나 이상이 생략되거나, 또는 추가의 오디오 데이터 프로세싱 유닛들이 포함된다. 예를 들면, 일 실시예에서, 후-처리 유닛(22)은

별개의 유닛 대신에 디코더(20)의 부분이다.

- [0031] 몇몇 구현들에서, 도 1의 전-처리 유닛은 입력(11)으로서 오디오 콘텐츠를 포함한 PCM(시간-도메인) 샘플들을 받아들이고, 및 프로세싱된 PCM 샘플들을 출력하도록 구성된다. 인코더(14)는 입력으로서 상기 PCM 샘플들을 받아들이고, 상기 오디오 콘텐츠를 나타내는 인코딩된(예로서, 압축된) 오디오 비트스트림을 출력하도록 구성될 수 있다. 상기 오디오 콘텐츠를 나타내는 비트스트림의 데이터는 때때로 여기에서 "오디오 데이터"로서 불린다. 일 실시예에서, 상기 인코더로부터 출력된 오디오 비트스트림은 오디오 데이터뿐만 아니라 라우드니스 프로세싱 상태 메타데이터(및 선택적으로 또한 다른 메타데이터)를 포함한다.
- [0032] 신호 분석 및 메타데이터 교정 유닛(16)은 입력으로서 하나 이상의 인코딩된 오디오 비트스트림들을 받아들이고 신호 분석을 수행함으로써, 각각의 인코딩된 오디오 비트스트림에서의 프로세싱 상태 메타데이터가 정확한지를 결정(예로서, 검증)할 수 있다. 일부 실시예들에서, 검증은 도 2에 도시된 요소(102)와 같은, 상태 검증기 구성 요소에 의해 수행될 수 있으며, 하나의 이러한 검증 기술이 상태 검증기(102)의 맥락에서 이하에 설명된다. 일부 실시예들에서, 유닛(16)은 인코더에 포함되며 검증은 유닛(16) 또는 검증기(102)에 의해 행해진다. 신호 분석 및 메타데이터 교정 유닛이 포함된 메타데이터가 유효하지 않음을 발견한다면, 메타데이터 교정 유닛(16)은 정확한 값(들)을 결정하기 위해 신호 분석을 수행하며 부정확한 값(들)을 상기 결정된 정확한 값(들)으로 대체한다. 따라서, 신호 분석 및 메타데이터 교정 유닛으로부터 출력된 각각의 인코딩된 오디오 비트스트림은 인코딩된 오디오 데이터뿐만 아니라 교정된 프로세싱 상태 메타데이터를 포함할 수 있다. 신호 분석 및 메타데이터 교정 유닛(16)은 전-처리 유닛(12), 인코더(14), 트랜스코더(18), 디코더(20), 또는 후-처리 유닛(22)의 부분이 될 수 있다. 대안적으로, 신호 분석 및 메타데이터 교정 유닛(16)은 오디오 프로세싱 체인에서 별개의 유닛 또는 또 다른 유닛의 부분이 될 수 있다.
- [0033] 트랜스코더(18)는 입력으로서 인코딩된 오디오 비트스트림들을 받아들이고, 이에 응답하여(예로서, 입력 스트림을 디코딩하고 상이한 인코딩 포맷으로 디코딩된 스트림을 재-인코딩함으로써) 수정된(예로서, 상이하게 인코딩된) 오디오 비트스트림들을 출력할 수 있다. 상기 트랜스코더로부터 출력된 오디오 비트스트림은 인코딩된 오디오 데이터뿐만 아니라 라우드니스 프로세싱 상태 메타데이터(및 선택적으로 또한 다른 메타데이터)를 포함한다. 상기 메타데이터는 비트스트림에 포함될 수 있다.
- [0034] 도 1의 디코더(20)는 입력으로서 인코딩된(예로서, 압축된) 오디오 비트스트림들을 받아들이고, (이에 응답하여) 디코딩된 PCM 오디오 샘플들의 스트림들을 출력할 수 있다. 일 실시예에서, 디코더의 출력은 다음 중 어떤 것이거나 또는 이를 포함한다: 오디오 샘플들의 스트림, 및 입력 인코딩된 비트스트림으로부터 추출된 라우드니스 프로세싱 상태 메타데이터(및 선택적으로 또한 다른 메타데이터)의 대응 스트림; 오디오 샘플들의 스트림, 및 입력 인코딩된 비트스트림으로부터 추출된 라우드니스 프로세싱 상태 메타데이터(및 선택적으로 또한 다른 메타데이터)로부터 결정된 제어 비트들의 대응 스트림; 또는 프로세싱 상태 메타데이터의 대응 스트림 또는 프로세싱 상태 메타데이터로부터 결정된 제어 비트들 없는, 오디오 샘플들의 스트림. 이러한 마지막 경우에, 디코더는, 상기 추출된 메타데이터 또는 그로부터 결정된 제어 비트들을 출력하지 않더라도, 입력 인코딩된 비트스트림으로부터 라우드니스 프로세싱 상태 메타데이터(및/또는 다른 메타데이터)를 추출하고 상기 추출된 메타데이터에 대한 적어도 하나의 동작(예로서, 검증)을 수행할 수 있다.
- [0035] 본 발명의 실시예에 따라 도 1의 후-처리 유닛을 구성함으로써, 후-처리 유닛(22)은 디코딩된 PCM 오디오 샘플들의 스트림을 받아들이고, 샘플들과 함께 수신된 라우드니스 프로세싱 상태 메타데이터(및 선택적으로 또한 다른 메타데이터), 또는 상기 샘플들과 함께 수신된 제어 비트들(라우드니스 프로세싱 상태 메타데이터 및 선택적으로 또한 다른 메타데이터로부터 디코더에 의해 결정)을 사용하여 그것에 대한 후 처리(예로서, 오디오 콘텐츠의 볼륨 레벨링)를 수행하도록 구성된다. 상기 후-처리 유닛(22)은 선택적으로 또한 하나 이상의 스피커들에 의해 재생을 위한 상기 후-처리된 오디오 콘텐츠를 렌더링하도록 구성된다. 이들 스피커들은 컴퓨터들, 텔레비전들, 스테레오 시스템들(홈 또는 시네마), 이동 전화들, 및 다른 휴대용 재생 디바이스들과 같은, 재생 장비의 다양한 상이한 청취 디바이스들 또는 아이템들 중 어떠한 것으로도 구현될 수 있다. 상기 스피커들은 어떠한 적절한 크기 및 출력 정격도 될 수 있으며, 독립된 드라이버들, 스피커 엔클로저들, 서라운드-사운드 시스템들, 사운드바들, 헤드폰들, 이어버드들 등의 형태로 제공될 수 있다.
- [0036] 일부 실시예들은 오디오 프로세싱 유닛들(예로서, 인코더들, 디코더들, 트랜스코더들, 및 전- 및 후-처리 유닛들)이 오디오 프로세싱 유닛들에 의해 각각 수신된 라우드니스 프로세싱 상태 메타데이터에 의해 나타내어진 바와 같이 미디어 데이터의 동시 발생 상태에 따라 오디오 데이터에 적용되도록 그것들 각각의 프로세싱을 적응시키는 강화된 오디오 프로세싱 체인을 제공한다. 시스템(100)의 임의의 오디오 프로세싱 유닛(예로서, 도 1의 인

코더 또는 트랜스코더)으로의 오디오 데이터 입력(11)은 오디오 데이터(예로서, 인코딩된 오디오 데이터)뿐만 아니라 라우드니스 프로세싱 상태 메타데이터(및 선택적으로 또한 다른 메타데이터)를 포함할 수 있다. 이러한 메타데이터는 일부 실시예들에 따라 또 다른 요소 또는 또 다른 소스에 의해 입력 오디오에 포함될 수 있다. 상기 입력 오디오를 수신하는(메타데이터와 함께) 프로세싱 유닛은 메타데이터에 대한 또는 메타데이터에 응답하여(예로서, 입력 오디오의 적응적 프로세싱) 적어도 하나의 동작(예로서, 검중)을 수행하고, 또한 선택적으로 그것의 출력 오디오에 메타데이터, 메타데이터의 프로세싱된 버전, 또는 메타데이터로부터 결정된 제어 비트들을 포함하도록 구성될 수 있다.

[0037]

오디오 프로세싱 유닛(또는 오디오 프로세서)의 실시예는 오디오 데이터에 대응하는 라우드니스 프로세싱 상태 메타데이터에 의해 나타내어지는 바와 같이 오디오 데이터의 상태에 기초하여 오디오 데이터의 적응적 프로세싱을 수행하도록 구성된다. 일부 실시예들에서, 메타데이터가 라우드니스 프로세싱 또는 그것과 유사한 프로세싱이 오디오 데이터에 대해 아직 수행되지 않았다는 것을 나타내면, 적응적 프로세싱은 라우드니스 프로세싱이지만(또는 이를 포함하지만), 메타데이터가 이러한 라우드니스 프로세싱 또는 그것과 유사한 프로세싱이 이미 오디오 데이터에 대해 수행되었다는 것을 나타내면, 라우드니스 프로세싱이 아니다(이를 포함하지 않는다). 일부 실시예들에서, 적응적 프로세싱은 오디오 프로세싱 유닛이 라우드니스 프로세싱 상태 메타데이터에 의해 나타내어지는 바와 같이 오디오 데이터의 상태에 기초하여 오디오 데이터의 다른 적응적 프로세싱을 수행하는 것을 보장하기 위해 메타데이터 검중(예로서, 메타데이터 검중 서브-유닛에서 수행)이거나 또는 이를 포함한다. 일부 실시예들에서, 검중은 오디오 데이터와 연관된(예로서, 그와 함께 비트스트림에 포함된) 라우드니스 프로세싱 상태 메타데이터의 신뢰성을 결정한다. 예를 들면, 메타데이터가 신뢰 가능한 것으로 검중된다면, 이전 수행된 오디오 프로세싱의 유형으로부터의 결과들은 재사용될 수 있으며 동일한 유형의 오디오 프로세싱의 부가적인 수행은 회피될 수 있다. 다른 한편으로, 메타데이터가 변경되었다고(또는 그 외 신뢰 가능하지 않다고) 밝혀진다면, (신뢰 가능하지 않은 메타데이터로 나타내어진 바와 같이) 이전 수행된 것으로 알려진 미디어 프로세싱의 유형은 오디오 프로세싱 유닛에 의해 반복될 수 있으며, 및/또는 다른 프로세싱이 메타데이터 및/또는 오디오 데이터에 대해 오디오 프로세싱 유닛에 의해 수행될 수 있다. 상기 오디오 프로세싱 유닛은 또한 유닛이 프로세싱 상태 메타데이터가 유효하다고(예로서, 추출된 암호 값 및 기준 암호 값의 매칭에 기초하여) 결정한다면, (예로서, 미디어 비트스트림에 존재하는) 라우드니스 프로세싱 상태 메타데이터가 유효한 강화된 미디어 프로세싱 체인에서 아래쪽으로 다른 오디오 프로세싱 유닛들에 시그널링하도록 구성될 수 있다.

[0038]

도 1의 실시예를 위해, 전-처리 구성요소(12)는 인코더(14)의 부분일 수 있으며, 후-처리 구성요소(22)는 디코더(22)의 부분일 수 있다. 대안적으로, 전-처리 구성요소(12)는 인코더(14)로부터 분리되는 기능 구성요소로 구현될 수 있다. 유사하게, 후-처리 구성요소(22)는 디코더(20)로부터 분리되는 기능 구성요소로 구현될 수 있다.

[0039]

도 2는 도 1의 시스템(10)과 함께 사용될 수 있는 인코더(100)의 블록도이다. 인코더(100)의 구성요소들 또는 요소들 중 어떠한 것도 하드웨어, 소프트웨어, 또는 하드웨어 및 소프트웨어의 조합으로, 하나 이상의 프로세서들 및/또는 하나 이상의 회로들(예로서, ASIC들, FPGA들, 또는 다른 집적 회로들)로서 구현될 수 있다. 인코더(100)는 도시된 바와 같이 접속된, 프레임 버퍼(110), 파서(111), 디코더(101), 오디오 상태 검중기(102), 라우드니스 프로세싱 스테이지(103), 오디오 스트림 선택 스테이지(104), 인코더(105), 스테리오/포맷터 스테이지(107), 메타데이터 생성 스테이지(106), 다이얼로그 라우드니스 측정 서브시스템(108), 및 프레임 버퍼(109)를 포함한다. 선택적으로 또한, 인코더(100)는 다른 프로세싱 요소들(도시되지 않음)을 포함한다. (트랜스코더인) 인코더(100)는 입력 오디오 비트스트림(예를 들면, AC-3 비트스트림, E-AC-3 비트스트림, 또는 돌비 E 비트스트림 중 하나일 수 있음)을 입력 비트스트림에 포함된 라우드니스 프로세싱 상태 메타데이터를 사용하여 적응적 및 자동화된 라우드니스 프로세싱을 수행함으로써 포함하는 인코딩된 출력 오디오 비트스트림(예를 들면, AC-3 비트스트림, E-AC-3 비트스트림, 또는 돌비 E 비트스트림 중 또 다른 것일 수 있음)으로 변환하도록 구성된다. 예를 들면, 인코더(100)는 입력 돌비 E 비트스트림(통상적으로 방송되는 오디오 프로그램들을 수신하는 소비자 디바이스들에서 사용되지 않지만 제작 및 방송 설비들에서 사용되는 포맷)에서 AC-3 또는 E-AC-3 포맷으로 인코딩된 출력 오디오 비트스트림(소비자 디바이스들에 방송하기에 적합함)으로 변환하도록 구성될 수 있다.

[0040]

도 2의 시스템은 또한 인코딩된 오디오 전달 서브시스템(150)(인코더(100)로부터 출력된 인코딩된 비트스트림들을 저장 및/또는 전달함) 및 디코더(152)를 포함한다. 인코더(100)로부터 출력된 인코딩된 오디오 비트스트림은(예로서, DVD 또는 블루레이 디스크의 형태로) 서브시스템(150)에 의해 저장될 수 있거나, 또는 (송신 링크 또는 네트워크를 구현할 수 있는) 서브시스템(150)에 의해 송신되거나, 또는 서브시스템(150)에 의해 저장 및 송신 양쪽 모두가 될 수 있다. 디코더(152)는 (인코더(100)에 의해 생성된) 서브시스템(150)을 통해 수신하는 인코딩된 오디오 비트스트림을 디코딩하도록 구성되고, 비트스트림의 각각의 프레임으로부터 라우드니스 프로세싱

상태 메타데이터(LPSM)를 추출하고, 디코딩된 오디오 데이터를 생성하는 것을 포함한다. 일 실시예에서, 디코더(152)는 LPSM을 사용하여 디코딩된 오디오 데이터에 대한 적응적 라우드니스 프로세싱을 수행하도록 및/또는 LPSM을 사용하여 디코딩된 오디오 데이터에 대한 적응적 라우드니스 프로세싱을 수행하도록 구성된 전-처리기로 상기 디코딩된 오디오 데이터 및 LPSM을 포워딩하도록 구성된다. 선택적으로, 디코더(152)는 버퍼를 포함하며, 이것은 서브시스템(150)으로부터 수신된 인코딩된 오디오 비트스트림을 (예로서, 비-일시적 방식으로) 저장한다.

[0041] 인코더(100) 및 디코더(152)의 다양한 구현들이 여기에 설명된 상이한 실시예들을 수행하도록 구성된다. 프레임 버퍼(110)는 인코딩된 입력 오디오 비트스트림을 수신하기 위해 결합된 버퍼 메모리이다. 동작 시, 버퍼(110)는 인코딩된 오디오 비트스트림의 적어도 하나의 프레임을 (예로서, 비-일시적 방식으로) 저장하며, 인코딩된 오디오 비트스트림의 프레임들의 시퀀스는 버퍼(110)로부터 파서(110)로 어서팅(assert)된다. 파서(111)는 인코딩된 입력 오디오의 각각의 프레임으로부터 라우드니스 프로세싱 상태 메타데이터(LPSM) 및 다른 메타데이터를 추출하고, 적어도 상기 LPSM을 오디오 상태 검증기(102), 라우드니스 프로세싱 스테이지(103), 스테이지(106) 및 서브시스템(108)에 어서팅하고, 인코딩된 입력 오디오로부터 오디오 데이터를 추출하고, 오디오 데이터를 디코더(101)에 어서팅하도록 결합되며 구성된다. 인코더(100)의 디코더(101)는 디코딩된 오디오 데이터를 생성하기 위해 오디오 데이터를 디코딩하고, 디코딩된 오디오 데이터를 라우드니스 프로세싱 스테이지(103), 오디오 스트림 선택 스테이지(104), 서브시스템(108), 및 선택적으로 또한 상태 검증기(102)에 어서팅하도록 구성된다.

[0042] 상태 검증기(102)는 어서팅된 LPSM(및 선택적으로 다른 메타데이터)을 인증 및 검증하도록 구성된다. 일부 실시예들에서, 상기 LPSM은 (예로서, 본 발명의 실시예에 따르면) 입력 비트스트림에 포함되는 데이터 블록이다(또는 그것에 포함된다). 상기 블록은 상기 LPSM(및 선택적으로 또한 다른 메타데이터)을 프로세싱하기 위한 암호 해시(해시-기반 메시지 인증 코드 또는 "HMAC") 및/또는 (디코더(101)에서 검증기(102)로 제공된) 기본 오디오 데이터를 포함할 수 있다. 상기 데이터 블록은 이들 실시예들에서 디지털로 서명될 수 있으며, 따라서 다운스트림 오디오 프로세싱 유닛은 비교적 쉽게 프로세싱 상태 메타데이터를 인증하고 검증할 수 있다.

[0043] 예를 들면, HMAC은 다이제스트(digest)를 생성하기 위해 사용되며, 본 발명의 비트스트림에 포함된 보호 값(들)은 상기 다이제스트를 포함할 수 있다. 상기 다이제스트는 AC-3 프레임에 대해 다음과 같이 생성될 수 있다: (1) AC-3 데이터 및 LPSM이 인코딩된 후, 프레임 데이터 바이트들(연쇄된 프레임_데이터#1 및 프레임_데이터 #2) 및 LPSM 데이터 바이트들이 해싱-함수 HMAC에 대한 입력으로서 사용된다. 보조데이터 필드 내부에 존재할 수 있는 다른 데이터는 다이제스트를 산출하기 위해 고려되지 않는다. 이러한 다른 데이터는 AC-3 데이터에도 LPSM 데이터에도 속하지 않는 바이트들일 수 있다. 상기 LPSM에 포함된 보호 비트들은 HMAC 다이제스트를 산출하기 위해 고려되지 않을 수 있다. (2) 다이제스트가 산출된 후, 보호 비트들을 위해 예약된 필드에서 비트스트림으로 기록된다. (3) 완전한 AC-3 프레임의 생성의 마지막 단계는 CRC-검사의 산출이다. 이것은 프레임의 제일 끝에 기록되며 LPSM 비트들을 포함하여, 이러한 프레임에 속하는 모든 데이터는 LPSM 비트들을 포함하는 것으로 고려된다.

[0044] 이에 제한되지 않지만 하나 이상의 비-HMAC 암호 방법들 중 임의의 것을 포함한 다른 암호 방법들이 LPSM 및/또는 기본 오디오 데이터의 안전한 송신 및 수신을 보장하기 위해 LPSM의 검증(예로서, 검증기(102)에서)을 위해 사용될 수 있다. 예를 들면, (이러한 암호 방법을 사용하는) 검증은 비트스트림에 포함된 라우드니스 프로세싱 상태 메타데이터 및 비트스트림에 포함된 대응하는 오디오 데이터가 (메타데이터에 의해 나타내어진 바와 같이) 특정 라우드니스 프로세싱을 받았고(및/또는 그로부터의 결과이고) 이러한 특정 라우드니스 프로세싱의 수행 후 변경되지 않았는지를 결정하기 위해 오디오 비트스트림의 전형(embodiment)을 수신하는 각각의 오디오 프로세싱 유닛에서 수행될 수 있다.

[0045] 상태 검증기(102)는, 검증 동작의 결과들을 나타내기 위해, 제어 데이터를 오디오 스트림 선택 스테이지(104), 메타데이터 발생기(106), 및 다이얼로그 라우드니스 측정 서브시스템(108)에 어서팅한다. 제어 데이터에 응답하여, 스테이지(104)는 다음 중 하나를 선택할 수 있다(및 인코더(105)로 통과할 수 있다): (1) 라우드니스 프로세싱 스테이지(103)의 적응적으로 프로세싱된 출력(예로서, LPSM이 디코더(101)로부터 출력된 오디오 데이터가 특정 유형의 라우드니스 프로세싱을 받지 않음을 나타내고, 검증기(102)로부터의 제어 비트들이 LPSM이 유효하다는 것을 나타낼 때); 또는 (2) 디코더(101)로부터 출력된 오디오 데이터(예로서, LPSM이 디코더(101)로부터 출력된 오디오 데이터가 스테이지(103)에 의해 수행될 특정 유형의 라우드니스 프로세싱을 이미 받았음을 나타내고 검증기(102)로부터의 제어 비트들이 LPSM이 유효하다는 것을 나타낼 때). 실시예에서, 라우드니스 프로세싱 스테이지(103)는 라우드니스를 특정 타겟 및 라우드니스 범위로 교정한다.

- [0046] 인코더(100)의 스테이지(103)는 디코더(101)에 의해 추출된 LPSM에 의해 나타내어진 하나 이상의 오디오 데이터 특성들에 기초하여, 디코더(101)로부터 출력된 디코딩된 오디오 데이터에 대한 적응적 라우드니스 프로세싱을 수행하도록 구성된다. 스테이지(103)는 적응적 변환-도메인 실시간 라우드니스 및 동적 범위 제어 프로세서일 수 있다. 스테이지(103)는 사용자 입력(예로서, 사용자 타겟 라우드니스/동적 범위 값들 또는 Dialnorm 값들), 또는 다른 메타데이터 입력(예로서, 하나 이상의 유형들의 제 3 자 데이터, 추적 정보, 식별자들, 독점 또는 표준 정보, 사용자 주식 데이터, 사용자 선호 데이터 등), 및/또는 (예로서, 핑거프린팅 프로세스로부터의) 다른 입력을 수신하고, 디코더(101)로부터 출력된 디코딩된 오디오 데이터를 프로세싱하기 위해 이러한 입력을 사용할 수 있다.
- [0047] 다이얼로그 라우드니스 측정 서브시스템(108)은 검증기(102)로부터의 제어 비트들이 LPSM이 유효하지 않음을 표시할 때, 예를 들면, 디코더(101)에 의해 추출된 LPSM(및/또는 다른 메타데이터)을 사용하여, 다이얼로그(또는 다른 스피치)를 나타내는 (디코더(101)로부터의) 디코딩된 오디오의 세그먼트들의 라우드니스를 결정하도록 동작할 수 있다. 다이얼로그 라우드니스 측정 서브시스템(108)의 동작은 검증기(102)로부터의 제어 비트들이 LPSM이 유효함을 나타낼 때 LPSM이 (디코더(101)로부터의) 디코딩된 오디오의 다이얼로그(또는 다른 스피치) 세그먼트들의 이전 결정된 라우드니스를 나타낼 경우 디스에이블될 수 있다.
- [0048] 유용한 툴들(예로서, 돌비 LM100 라우드니스 미터)이 오디오 콘텐츠에서 다이얼로그의 레벨을 편리하고 쉽게 측정하기 위해 존재한다. APU(예로서, 인코더(100)의 스테이지(108))의 일부 실시예들은 오디오 비트스트림(예로서, 인코더(100)의 디코더(101)로부터 스테이지(108)로 어서팅된 디코딩된 AC-3 비트스트림)의 오디오 콘텐츠의 평균 다이얼로그 라우드니스를 측정하기 위해 이러한 툴을 포함하도록(또는 그것의 기능을 수행하도록) 구현된다. 스테이지(108)가 오디오 데이터의 실제 평균 다이얼로그 라우드니스를 측정하도록 실행된다면, 측정은 대개 스피치를 포함하는 오디오 콘텐츠의 세그먼트들을 분리하는 단계를 포함할 수 있다. 대개 스피치인 오디오 세그먼트들은 그 후 라우드니스 측정 알고리즘에 따라 프로세싱된다. AC-3 비트스트림으로부터 디코딩된 오디오 데이터에 대해, 이러한 알고리즘은 (국제 표준 ITU-R BS.1770에 따라) 표준 K-가중 라우드니스 측정될 수 있다. 대안적으로, 다른 라우드니스 측정들이 사용될 수 있다(예로서, 라우드니스의 음향 심리학적 모델들에 기초한 것들).
- [0049] 스피치 세그먼트들의 분리는 오디오 데이터의 평균 다이얼로그 라우드니스를 측정하는데 필수적인 것은 아니다. 그러나, 그것은 측정의 정확도를 개선하며 청취자의 관점으로부터 보다 만족스러운 결과들을 제공한다. 모든 오디오 콘텐츠가 다이얼로그(스피치)를 포함하는 것은 아니기 때문에, 스피치가 존재한다면, 전체 오디오 콘텐츠의 라우드니스 측정은 오디오의 다이얼로그 레벨의 충분한 근사를 제공할 수 있다.
- [0050] 메타데이터 발생기(106)는 인코더(100)로부터 출력될 인코딩된 비트스트림에서 스테이지(107)에 의해 포함될 메타데이터를 생성한다. 메타데이터 발생기(106)는 인코더(101)에 의해 추출된 LPSM(및/또는 다른 메타데이터)을 스테이지(107)로 통과시키거나(예로서, 검증기(102)로부터의 제어 비트들이 LPSM 및/또는 다른 메타데이터가 유효하다고 표시할 때) 또는 새로운 LPSM(및/또는 다른 메타데이터)을 생성하고, 새로운 메타데이터를 스테이지(107)로 어서팅하거나(예를 들면, 검증기(102)로부터의 제어 비트들이 디코더(101)에 의해 추출된 LPSM 및/또는 다른 메타데이터가 유효하지 않다고 표시할 때) 또는 스테이지(107)로 디코더(101)에 의해 추출된 메타데이터 및 새롭게 생성된 메타데이터의 조합을 어서팅할 수 있다. 메타데이터 발생기(106)는 서브시스템(108)에 의해 생성된 라우드니스 데이터, 및 서브시스템(108)에 의해 수행된 라우드니스 프로세싱의 유형을 나타내는 적어도 하나의 값을 포함할 수 있으며, LPSM에서, 인코더(100)로부터 출력될 인코딩된 비트스트림에 포함하기 위해 스테이지(107)로 어서팅한다. 메타데이터 발생기(106)는 인코딩된 비트스트림에 포함될 LPSM(및 선택적으로 또한 다른 메타데이터) 및/또는 인코딩된 비트스트림에 포함될 기본 오디오 데이터의 복호화(decryption), 인증, 또는 검증 중 적어도 하나에 유용한 보호 비트들(해시-기반 메시지 인증 코드 또는 "HMAC")으로 이루어지거나 또는 이를 포함할 수 있음)을 생성할 수 있다. 메타데이터 발생기(106)는 인코딩된 비트스트림에 포함하기 위해 스테이지(107)로 이러한 보호 비트들을 제공할 수 있다.
- [0051] 일 실시예에서, 다이얼로그 라우드니스 측정 서브시스템(108)은 디코더(101)로부터 출력된 오디오 데이터를 프로세싱하고 그에 응답하여 라우드니스 값들(예로서, 게이팅된 및 게이팅되지 않은 다이얼로그 라우드니스 값들) 및 동적 범위 값들을 생성한다. 이들 값들에 응답하여, 메타데이터 발생기(106)는 (스터퍼/포맷터(107)에 의해) 인코더(100)로부터 출력될 인코딩된 비트스트림에 포함하기 위해 라우드니스 프로세싱 상태 메타데이터(LPSM)를 생성할 수 있다. 실시예에서, 라우드니스는 ITU-R BS.1770-1 및 ITU-R BS.1770-2 표준들, 또는 다른 유사한 라우드니스 측정 표준들에 의해 특정된 기술들에 기초하여 산출될 수 있다. 게이팅된 라우드니스는 다이얼로그-게이팅 라우드니스 또는 상대적-게이팅 라우드니스, 또는 이들 게이팅 라우드니스 유형들의 조합일 수 있으며, 시

시스템은 애플리케이션 요건들 및 시스템 제약들에 의존하여 적절한 게이팅 블록들을 이용할 수 있다.

- [0052] 부가적으로, 선택적으로, 또는 대안적으로, 인코더(100)의 106 및/또는 108의 서브시스템들은 스테이지(107)로부터 출력될 인코딩된 비트스트림에 포함하기 위해 오디오 데이터의 적어도 하나의 특성을 나타내는 메타데이터를 생성하기 위해 오디오 데이터의 부가적인 분석을 수행할 수 있다. 인코더(105)는 선택 스테이지(104)로부터 출력된 오디오 데이터를 (예로서, 그에 대한 압축을 수행함으로써) 인코딩하고, 인코딩된 오디오를 스테이지(107)로부터 출력될 인코딩된 비트스트림에 포함하기 위해 스테이지(107)로 어서팅한다.
- [0053] 스테이지(107)는 스테이지(107)로부터 출력될 인코딩된 비트스트림을 생성하기 위해 발생기(106)로부터의 메타데이터(LPSM을 포함) 및 인코더(105)로부터의 인코딩된 오디오를 다중화하며, 따라서 인코딩된 비트스트림은 실시예에 의해 특정된 바와 같은 포맷을 가진다. 프레임 버퍼(109)는 스테이지(107)로부터 출력된 인코딩된 오디오 비트스트림의 적어도 하나의 프레임을 (예로서, 비-일시적 방식으로) 저장하는 버퍼 메모리이며, 인코딩된 오디오 비트스트림의 프레임들의 시퀀스는 그 후 인코더(100)로부터의 출력으로서 버퍼(109)로부터 전달 시스템(150)으로 어서팅된다.
- [0054] 메타데이터 발생기(106)에 의해 생성되고 스테이지(107)에 의해 인코딩된 비트스트림에 포함된 LPSM은 대응하는 오디오 데이터의 라우드니스 프로세싱 상태(예로서, 어떤 유형(들)의 라우드니스 프로세싱이 오디오 데이터에 대해 수행되었는지) 및 대응하는 오디오 데이터의 라우드니스(예로서, 측정된 다이얼로그 라우드니스, 게이팅된 및/또는 게이팅되지 않은 라우드니스, 및/또는 동적 범위를)를 나타낸다. 여기에서, 오디오 데이터에 대해 수행된 라우드니스 및/또는 레벨 측정들의 "게이팅"은 임계치를 초과하는 계산된 값(들)이 최종 측정에 포함되는 라우드니스 임계치 또는 특정 레벨(예로서, 최종 측정 값들에서 -60 dBFS 아래의 단기 라우드니스 값들을 무시)을 나타낸다. 절대 값에 대한 게이팅은 고정 레벨 또는 라우드니스를 나타내는 반면, 상대적 값에 대한 게이팅은 현재 "게이팅되지 않은" 측정 값에 의존하는 값을 나타낸다.
- [0055] 인코더(100)의 몇몇 구현에서, 메모리(109)에 버퍼링된(및 전달 시스템(150)에 출력된) 인코딩된 비트스트림은 AC-3 비트스트림 또는 E-AC-3 비트스트림이며, 오디오 데이터 세그먼트들(예로서, 도 4에 도시된 프레임의 ABO 내지 AB5 세그먼트들) 및 메타데이터 세그먼트들을 포함하고, 여기에서 오디오 데이터 세그먼트들은 오디오 데이터를 나타내며, 메타데이터 세그먼트들의 적어도 일부의 각각은 라우드니스 프로세싱 상태 메타데이터(LPSM)를 포함한다. 스테이지(107)는 다음의 포맷에서 비트스트림에 LPSM을 삽입한다. LPSM을 포함하는 메타데이터 세그먼트들의 각각은 비트스트림의 프레임의 비트스트림 정보("BSI") 세그먼트의 "addbsi" 필드에, 또는 비트스트림의 프레임의 끝에서 보조데이터 필드(예로서, 도 4에 도시된 AUX 세그먼트)에 포함된다.
- [0056] 비트스트림의 프레임은 그 각각이 LPSM을 포함하는, 하나 또는 두 개의 메타데이터 세그먼트들을 포함하며, 프레임이 두 개의 메타데이터 세그먼트들을 포함한다면, 하나는 프레임의 addbsi 필드에 존재하고 다른 하나는 프레임의 AUX 필드에 존재한다. LPSM을 포함한 각각의 메타데이터 세그먼트는 다음의 포맷을 가진 LPSM 페이로드(또는 컨테이너) 세그먼트를 포함한다: 헤더(예로서, LPSM 페이로드의 시작을 식별하는 syncword, 이어서 적어도 하나의 식별 값, 예로서 LPSM 포맷 버전, 길이, 기간, 카운트, 및 이하의 표 2에 표시된 서브스트림 연관 값들을 포함); 및 헤더 후, 대응하는 오디오 데이터가 다이얼로그를 표시하는지 또는 다이얼로그를 표시하는 않는지를 표시한 적어도 하나의 다이얼로그 표시 값(예로서, 표 2의 파라미터 "다이얼로그 채널(들)")(예로서, 대응하는 오디오 데이터의 어떤 채널들이 다이얼로그를 표시하는지); 대응하는 오디오 데이터가 표시된 세트의 라우드니스 규정들을 따르는지를 표시한 적어도 하나의 라우드니스 규정 준수 값(예로서, 표 2의 파라미터 "라우드니스 규정 유형"); 대응하는 오디오 데이터에 대해 수행된 라우드니스 프로세싱의 적어도 하나의 유형을 표시한 적어도 하나의 라우드니스 프로세싱 값(예로서, 표 2의 파라미터들 "다이얼로그 게이팅 라우드니스 교정 플래그", "라우드니스 교정 유형") 중 하나 이상); 및 대응하는 오디오 데이터의 적어도 하나의 라우드니스 (예로서, 피크 또는 평균 라우드니스) 특성을 표시한 적어도 하나의 라우드니스 값(예로서, 표 2의 파라미터들 "ITU 상대적 게이팅 라우드니스", "ITU 스피치 게이팅 라우드니스", "ITU (EBU 3341) 단기 3s 라우드니스", 및 "실제 피크" 중 하나 이상).
- [0057] 몇몇 구현들에서, 스테이지(107)에 의해 비트스트림의 프레임의 "addbsi" 필드 또는 보조데이터 필드에 삽입된 메타데이터 세그먼트들의 각각은 다음의 포맷을 가진다: 코어 헤더(예로서, 메타데이터 세그먼트의 시작을 식별하는 syncword, 이어서 식별 값들, 예로서 코어 요소 버전, 길이, 및 이하의 표 1에 표시된 기간, 연장된 요소 카운트, 및 서브스트림 연관 값들을 포함); 및 상기 코어 헤더 후, 라우드니스 프로세싱 상태 메타데이터 또는 대응하는 오디오 데이터의 적어도 하나의 복호화, 인증, 또는 검증의 적어도 하나에 유용한 적어도 하나의 보호 값(예로서, 표 1의 HMAC 다이제스트 및 오디오 핑거프린트 값들); 및 또한 상기 코어 헤더 후, 메타데이터 세그

먼트가 LPSM을 포함한다면, 이어지는 메타데이터를 LPSM 페이로드로서 식별하며 LPSM 페이로드의 크기를 표시하는 LPSM 페이로드 식별("ID") 및 LPSM 페이로드 크기 값들.

[0058] LPSM 페이로드(또는 컨테이너) 세그먼트(예로서, 상기-특정된 포맷을 가짐)는 LPSM 페이로드 ID 및 LPSM 페이로드 크기 값들을 따른다.

[0059] 일부 실시예들에서, 프레임의 보조데이터 필드(또는 "addbsi" 필드)에서의 메타데이터 세그먼트들의 각각은 3개의 레벨들의 구조를 가진다: 보조데이터(또는 addbsi) 필드가 메타데이터를 포함하는지를 표시한 플래그, 어떤 유형(들)의 메타데이터가 존재하는지를 표시한 적어도 하나의 ID 값, 및 선택적으로 또한 얼마나 많은 비트들의 (예로서, 각각의 유형의) 메타데이터가 존재하는지(메타데이터가 존재한다면)를 표시한 값을 포함한 고 레벨 구조. 존재할 수 있는 메타데이터의 한가지 유형은 LPSM이며, 존재할 수 있는 메타데이터의 또 다른 유형은 미디어 리서치 메타데이터(예로서, nilsen 미디어 리서치 메타데이터)이다; 각각의 식별된 유형의 메타데이터를 위한 코어 요소(예로서, 상기 언급된 바와 같이, 메타데이터의 각각의 식별된 유형에 대한, 코어 헤더, 보호 값들, 및 LPSM 페이로드 ID 및 LPSM 페이로드 크기 값들)를 포함한 중간 레벨 구조; 및 하나의 코어 요소에 대한 각각의 페이로드(예로서, 하나가 코어 요소에 의해 존재하는 것으로서 식별된다면, LPSM 페이로드, 및/또는 하나가 코어 요소에 의해 존재하는 것으로서 식별된다면, 또 다른 유형의 메타데이터 페이로드)를 포함한, 저 레벨 구조.

[0060] 이러한 3 레벨 구조에서의 데이터 값들은 내포될 수 있다. 예를 들면, 코어 요소에 의해 식별된 LPSM 페이로드 및/또는 또 다른 메타데이터 페이로드에 대한 보호 값(들)은 코어 요소에 의해 식별된 각각의 페이로드 후(및 따라서 코어 요소의 코어 헤더 후) 포함될 수 있다. 일 예에서, 코어 헤더는 LPSM 페이로드 및 또 다른 메타데이터 페이로드를 식별할 수 있고, 제 1 페이로드(예로서, LPSM 페이로드)에 대한 페이로드 ID 및 페이로드 크기 값들은 코어 헤더를 따를 수 있고, 제 1 페이로드 자체는 ID 및 크기 값들을 따를 수 있고, 제 2 페이로드에 대한 페이로드 ID 및 페이로드 크기 값은 제 1 페이로드를 따를 수 있고, 제 2 페이로드 자체는 이들 ID 및 크기 값들을 따를 수 있으며, 양쪽 페이로드들에 대한(또는 코어 요소 값들 및 양쪽 페이로드들 모두에 대한) 보호 비트들은 마지막 페이로드를 따를 수 있다.

[0061] 일부 실시예들에서, 디코더(101)가 암호 해시를 갖는 본 발명의 실시예에 따라 생성된 오디오 비트스트림을 수신한다면, 디코더는 비트스트림으로부터 결정된 데이터 블록으로부터 암호 해시를 파싱 및 검색하도록 구성되며, 상기 블록은 라우드니스 프로세싱 상태 메타데이터(LPSM)를 포함한다. 검증기(102)는 수신된 비트스트림 및/또는 연관된 메타데이터를 검증하기 위해 암호 해시를 사용할 수 있다. 예를 들면, 검증기(102)는 기준 암호 해시와 데이터 블록으로부터 검색된 암호 해시 사이에서의 매칭에 기초하여 LPSM을 유효한 것으로 발견하며, 그 후 대응하는 오디오 데이터에 대한 프로세서(103)의 동작을 불능시킬 수 있으며 선택 스테이지(104)로 하여금 (변경되지 않은) 오디오 데이터를 통과시키게 할 수 있다. 부가적으로, 선택적으로, 또는 대안적으로, 다른 유형들의 암호 기술들이 암호 해시에 기초한 방법을 대신하여 사용될 수 있다.

[0062] 도 2의 인코더(100)는 후/전-처리 유닛이 (요소들(105, 106, 및 107)에서) 인코딩될 오디오 데이터에 대한 일 유형의 라우드니스 프로세싱을 수행했는지를 (디코더(101)에 의해 추출된 LPSM에 응답하여) 결정할 수 있으며, 따라서 이전 수행된 라우드니스 프로세싱에서 사용되며 및/또는 그로부터 도출된 특정 파라미터들을 포함하는 라우드니스 프로세싱 상태 메타데이터를 (발생기(106)에서) 생성할 수 있다. 몇몇 구현들에서, 인코더(100)는 인코더가 오디오 콘텐츠에 대해 수행되는 프로세싱의 유형들을 알고 있는 한 오디오 콘텐츠에 대한 프로세싱 이력을 나타내는 프로세싱 상태 메타데이터를 생성할 수 있다(및 그로부터 출력된 인코딩된 비트스트림에 포함할 수 있다).

[0063] 도 3은 도 1의 시스템(10)과 함께 사용될 수 있는 디코더의 블록도이다. 디코더(200) 및 후-처리기(300)의 구성 요소들 또는 요소들 중 어떠한 것도 하드웨어, 소프트웨어, 또는 하드웨어 및 소프트웨어의 조합으로 하나 이상의 프로세스들 및/또는 하나 이상의 회로들(예로서, ASIC들, FPGA들, 또는 다른 집적 회로들)로서 구현될 수 있다. 디코더(200)는 도시된 바와 같이 접속된, 프레임 버퍼(201), 파서(205), 오디오 디코더(202), 오디오 상태 검증 스테이지(검증기)(203), 및 제어 비트 발생 스테이지(204)를 포함한다. 디코더(200)는 다른 프로세싱 요소들(도시되지 않음)을 포함할 수 있다. 프레임 버퍼(201)(버퍼 메모리)는 디코더(200)에 의해 수신된 인코딩된 오디오 비트스트림의 적어도 하나의 프레임을 (예로서, 비-일시적 방식으로) 저장한다. 인코딩된 오디오 비트스트림의 프레임들의 시퀀스는 버퍼(201)로부터 파서(205)로 어서팅된다. 파서(205)는 인코딩된 입력 오디오의 각각의 프레임으로부터 라우드니스 프로세싱 상태 메타데이터(LPSM) 및 다른 메타데이터를 추출하고, 적어도 상기 LPSM을 오디오 상태 검증기(203) 및 스테이지(204)에 어서팅하고, 출력으로서 상기 LPSM을 (예로서, 후-처리기

(300)에) 어서팅하고, 인코딩된 입력 오디오로부터 오디오 데이터를 추출하고, 추출된 오디오 데이터를 디코더(202)에 어서팅하도록 결합되고 구성된다. 디코더(200)에 입력된 인코딩된 오디오 비트스트림은 AC-3 비트스트림, E-AC-3 비트스트림, 또는 돌비 E 비트스트림 중 하나일 수 있다.

[0064] 도 3의 시스템은 또한 후-처리기(300)를 포함한다. 후-처리기(300)는 프레임 버퍼(301) 및 버퍼(301)에 결합된 적어도 하나의 프로세싱 요소를 포함한 다른 프로세싱 요소들(도시되지 않음)을 포함한다. 프레임 버퍼(301)는 후-처리기(300)에 의해 디코더(200)로부터 수신된 디코딩된 오디오 비트스트림의 적어도 하나의 프레임을 (예로서, 비-일시적 방식으로) 저장한다. 후-처리기(300)의 프로세싱 요소들은 디코더(202)로부터 출력된 메타데이터(LPSM 값들을 포함) 및/또는 디코더(200)의 스테이지(204)로부터 출력된 제어 비트들을 사용하여, 버퍼(301)로부터 출력된 디코딩된 오디오 비트스트림의 프레임들의 시퀀스를 수신하며 적응적으로 프로세싱하도록 결합되고 구성된다. 일 실시예에서, 후-처리기(300)는 (예로서, LPSM에 의해 표시된, 라우드니스 프로세싱 상태, 및/또는 하나 이상의 오디오 데이터 특성들에 기초하여) LPSM 값들을 사용하여 디코딩된 오디오 데이터에 대한 적응적 라우드니스 프로세싱을 수행하도록 구성된다. 디코더(200) 및 후-처리기(300)의 다양한 구현들은 여기에 설명된 실시예들에 따른 방법들의 상이한 실시예들을 수행하도록 구성된다.

[0065] 디코더(200)의 오디오 디코더(202)는 디코딩된 오디오 데이터를 생성하기 위해 파서(205)에 의해 추출된 오디오 데이터를 디코딩하고, 출력으로서 상기 디코딩된 오디오 데이터를 (예로서, 후-처리기(300)에) 어서팅하도록 구성된다. 상태 검증기(203)는 어서팅된 LPSM(및 선택적으로 다른 메타데이터)을 인증 및 검증하도록 구성된다. 일부 실시예들에서, 상기 LPSM은 (예로서, 본 발명의 실시예에 따라) 입력 비트스트림에 포함된 데이터 블록이다(또는 그것에 포함된다). 상기 블록은 LPSM(및 선택적으로 또한 다른 메타데이터) 및/또는 (파서(205) 및/또는 디코더(202)로부터 검증기(203)로 제공된) 기본 오디오 데이터를 프로세싱하기 위한 암호 해시(해시-기반 메시지 인증 코드 또는 "HMAC")를 포함할 수 있다. 데이터 블록은 이들 실시예들에서 디지털로 서명될 수 있으며, 따라서 다운스트림 오디오 프로세싱 유닛은 프로세싱 상태 메타데이터를 비교적 쉽게 인증 및 검증할 수 있다.

[0066] 이에 제한되지 않지만, 하나 이상의 비-HMAC 암호 방법들 중 임의의 것을 포함한 다른 암호 방법들이 LPSM 및/또는 기본 오디오 데이터의 안전한 송신 및 수신을 보장하기 위해 (예로서, 검증기(203)에서) LPSM의 검증을 위해 사용될 수 있다. 예를 들면, (이러한 암호 방법을 사용한) 검증은 비트스트림에 포함된 라우드니스 프로세싱 상태 메타데이터 및 대응하는 오디오 데이터가 (메타데이터에 의해 표시된 바와 같이) 특정 라우드니스 프로세싱을 받는지(및/또는 그로부터 결과인지) 및 이러한 특정 라우드니스 프로세싱의 수행 후 변경되지 않았는지를 결정하기 위해 본 발명의 오디오 비트스트림의 전형을 수신하는 각각의 오디오 프로세싱에서 수행될 수 있다.

[0067] 상태 검증기(203)는, 검증 동작의 결과들을 표시하기 위해, 제어 데이터를 제어 비트 발생기(204)에 어서팅하고, 및/또는 (예로서, 후-처리기(300)에) 출력으로서 제어 데이터를 어서팅한다. 제어 데이터(및 선택적으로 또한 입력 비트스트림으로부터 추출된 다른 메타데이터)에 응답하여, 스테이지(204)는 다음 중 하나를 생성할 수 있다(및 후-처리기(300)에 어서팅할 수 있다): 디코더(202)로부터 출력된 디코딩된 오디오 데이터가 특정 유형의 라우드니스 프로세싱을 받았는지를 표시하는 제어 비트들(LPSM이 디코더(202)로부터 출력된 오디오 데이터가 특정 유형의 라우드니스 프로세싱을 받았음을 표시하며, 검증기(203)로부터의 제어 비트들이 LPSM이 유효하다고 표시할 때); 또는 디코더(202)로부터 출력된 디코딩된 오디오 데이터가 특정 유형의 라우드니스 프로세싱을 받아야 한다고 표시하는 제어 비트들(예로서, LPSM이 디코더(202)로부터 출력된 오디오 데이터가 특정 유형의 라우드니스 프로세싱을 받지 않았음을 표시할 때, 또는 LPSM이 디코더(202)로부터 출력된 오디오 데이터가 특정 유형의 라우드니스 프로세싱을 받았음을 표시하지만 검증기(203)로부터의 제어 비트들은 LPSM이 유효하지 않다고 표시할 때).

[0068] 대안적으로, 디코더(200)는 입력 비트스트림으로부터 디코더(202)에 의해 추출된 LPSM(및 임의의 다른 메타데이터)을 후-처리기(300)에 어서팅하며, 후-처리기(300)는 LPSM을 사용하여 디코딩된 오디오 데이터에 대한 라우드니스 프로세싱을 수행하거나, 또는 LPSM의 검증을 수행하며 그 후 검증이 LPSM이 유효하다고 표시한다면 LPSM을 사용하여 디코딩된 오디오 데이터에 대한 라우드니스 프로세싱을 수행한다.

[0069] 일부 실시예들에서, 디코더(201)가 암호 해시를 갖는 본 발명의 실시예에 따라 생성된 오디오 비트스트림을 수신한다면, 디코더는 비트스트림으로부터 결정된 데이터 블록으로부터 암호 해시를 파싱 및 검색하도록 구성되며, 상기 블록은 라우드니스 프로세싱 상태 메타데이터(LPSM)를 포함한다. 검증기(203)는 수신된 비트스트림 및/또는 연관된 메타데이터를 검증하기 위해 암호 해시를 사용할 수 있다. 예를 들면, 검증기(203)가 기준 암호 해시와 데이터 블록으로부터 검색된 암호 해시 사이에서의 매칭에 기초하여 LPSM을 유효한 것으로 발견한다면, 비트스트림의 (변경되지 않은) 오디오 데이터를 통과시키기 위해 다운스트림 오디오 프로세싱 유닛(예로

서, 볼륨 레벨링 유닛이거나 또는 이를 포함할 수 있는, 후-처리기(300))으로 시그널링할 수 있다. 부가적으로, 선택적으로, 또는 대안적으로, 다른 유형들의 암호 기술들이 암호 해시에 기초한 방법을 대신하여 사용될 수 있다.

[0070] 디코더(100)의 몇몇 구현들에서, 수신된(및 메모리(201)에 버퍼링된) 인코딩된 비트스트림은 AC-3 비트스트림 또는 E-AC-3 비트스트림이며, 오디오 데이터 세그먼트들(예로서, 도 4에 도시된 프레임의 AB0 내지 AB5 세그먼트들) 및 메타데이터 세그먼트들을 포함하고, 여기에서 오디오 데이터 세그먼트들은 오디오 데이터를 표시하며, 메타데이터 세그먼트들 중 적어도 일부의 각각은 라우드니스 프로세싱 상태 메타데이터(LPSM)를 포함한다. 디코더 스테이지(202)는 다음의 포맷을 가진 LPSM을 비트스트림으로부터 추출하도록 구성된다. LPSM을 포함하는 메타데이터 세그먼트들의 각각은 비트스트림의 프레임의 비트스트림 정보("BSI") 세그먼트의 "addbsi" 필드에, 또는 비트스트림의 프레임의 끝에서 보조데이터 필드(예로서, 도 4에 도시된 AUX 세그먼트)에 포함된다. 비트스트림의 프레임은 그 각각이 LPSM을 포함하는 하나 또는 두 개의 메타데이터 세그먼트들을 포함할 수 있으며, 프레임이 두 개의 메타데이터 세그먼트들을 포함한다면, 하나는 프레임의 addbsi 필드에 존재하며 다른 하나는 프레임의 AUX 필드에 존재한다. LPSM을 포함한 각각의 메타데이터 세그먼트는 다음의 포맷을 가진 LPSM 페이로드(또는 컨테이너) 세그먼트를 포함한다: 헤더(예로서, LPSM 페이로드의 시작을 식별하는 syncword, 이어서 식별 값들, 예로서 이하의 표 2에 표시된 LPSM 포맷 버전, 길이, 기간, 카운트, 및 서브스트림 연관 값들을 포함); 및 상기 헤더 후, 대응하는 오디오 데이터가 다이얼로그를 표시하는지 또는 다이얼로그를 표시하지 않는지(예로서, 대응하는 오디오 데이터의 어떤 채널들이 다이얼로그를 표시하는지)를 표시한 적어도 하나의 다이얼로그 표시 값(예로서, 표 2의 파라미터 "다이얼로그 채널(들)"); 대응하는 오디오 데이터가 표시된 세트의 라우드니스 규정들을 따르는지를 표시한 적어도 하나의 라우드니스 규정 준수 값(예로서, 표 2의 파라미터 "라우드니스 규정 유형"); 대응하는 오디오 데이터에 대해 수행된 라우드니스 프로세싱의 적어도 하나의 유형을 표시한 적어도 하나의 라우드니스 프로세싱 값(예로서, 표 2의 파라미터들 "다이얼로그 게이팅 라우드니스 교정 플러그", "라우드니스 교정 유형" 중 하나 이상); 및 대응하는 오디오 데이터의 적어도 하나의 라우드니스(예로서, 피크 또는 평균 라우드니스) 특성을 표시한 적어도 하나의 라우드니스 값(예로서, 표 2의 파라미터들 "ITU 상대적 게이팅 라우드니스", "ITU 스피치 게이팅 라우드니스", "ITU (EBU 3341) 단기 3s 라우드니스", 및 "실제 피크" 중 하나 이상).

[0071] 몇몇 구현들에서, 디코더 스테이지(202)는 비트스트림의 프레임의 "addbsi" 필드 또는 보조데이터 필드로부터, 다음의 포맷을 가진 각각의 메타데이터 세그먼트를 추출하도록 구성된다: 코어 헤더(예로서, 메타데이터 세그먼트의 시작을 식별하는 syncword, 이어서 적어도 하나의 식별 값, 예로서 이하의 표 1에 표시된 코어 요소 버전, 길이, 및 기간, 연장된 요소 카운트, 및 서브스트림 연관 값들을 포함); 및 상기 코어 헤더 후, 라우드니스 프로세싱 상태 메타데이터 또는 대응하는 오디오 데이터 중 적어도 하나의 복호화, 인증, 또는 검증 중 적어도 하나에 유용한 적어도 하나의 보호 값(예로서, 표 1의 HMAC 다이제스트 및 오디오 핑거프린트 값들); 및 또한 상기 코어 헤더 후, 메타데이터 세그먼트가 LPSM을 포함한다면, 다음의 메타데이터를 LPSM 페이로드로 식별하며 LPSM 페이로드의 크기를 표시하는 LPSM 페이로드 식별("ID") 및 LPSM 페이로드 크기. (예로서, 상기 특정된 포맷을 가진) 상기 LPSM 페이로드(또는 컨테이너) 세그먼트는 LPSM 페이로드 ID 및 LPSM 페이로드 크기 값들을 따른다.

[0072] 보다 일반적으로, 실시예에 의해 생성된 인코딩된 오디오 비트스트림은 코어(의무적) 또는 확장(선택적 요소들)으로서 메타데이터 요소들 및 서브-요소들을 라벨링하기 위한 메커니즘을 제공하는 구조를 가진다. 이것은 (그 메타데이터를 포함한) 비트스트림의 데이터 레이트로 하여금 다수의 애플리케이션들에 걸쳐 스케일링하도록 허용한다. 비트스트림 선택스의 코어(의무적) 요소들은 또한 오디오 콘텐츠와 연관된 확장(선택적) 요소들이 존재하며(밴드-내) 및/또는 원격 위치에(밴드 외) 있음을 시그널링할 수 있어야 한다.

[0073] 일부 실시예에서, 코어 요소(들)는 비트스트림의 매 프레임에 존재하도록 요구된다. 코어 요소들의 몇몇 서브-요소들은 선택적이며 임의의 조합으로 존재할 수 있다. 확장된 요소들은 (비트레이트 오버헤드를 제한하기 위해) 매 프레임에 존재하도록 요구되지 않는다. 따라서, 확장 요소들은 몇몇 프레임들에 존재할 수 있으며 다른 것들에는 존재하지 않을 수 있다. 확장 요소의 몇몇 서브-요소들은 선택적이며 임의의 조합으로 존재할 수 있지만, 확장 요소의 몇몇 서브-요소들은 의무적일 수도 있다(즉, 확장 요소가 비트스트림의 프레임에 존재한다면).

[0074] 일부 실시예들에서, 오디오 데이터 세그먼트 및 메타데이터 세그먼트들의 시퀀스를 포함한 인코딩된 오디오 비트스트림이 (예로서, 본 발명을 구현한 오디오 프로세싱 유닛에 의해) 생성된다. 오디오 데이터 세그먼트들은 오디오 데이터를 나타내며, 메타데이터 세그먼트들 중 적어도 일부의 각각은 라우드니스 프로세싱 상태 메타데

이터(LPSM)를 포함하며 오디오 데이터 세그먼트들은 메타데이터 세그먼트들과 시간-분할 다중화된다. 이 클래스에서의 일부 실시예들에서, 메타데이터 세그먼트들의 각각은 여기에 설명될 포맷을 가진다. 하나의 포맷에서, 인코딩된 비트스트림은 AC-3 비트스트림 또는 E-AC-3 비트스트림이며, LPSM을 포함하는 메타데이터 세그먼트들의 각각은 비트스트림의 프레임의 비트스트림 정보("BSI") 세그먼트의 "addbsi" 필드(도 6에 도시)에서, 또는 비트스트림의 프레임의 보조데이터 필드에서 (예로서, 인코더(100)의 스테이지(107)에 의해) 추가적인 비트 스트림 정보로서 포함된다. 프레임들의 각각은 도 8의 표 1에 도시된 포맷을 가진 프레임의 addbsi 필드에 코어 요소를 포함한다.

[0075] 하나의 포맷에서, LPSM을 포함하는 addbsi(또는 보조데이터) 필드들의 각각은 코어 헤더(및 선택적으로 또한 추가적인 코어 요소들), 및 상기 코어 헤더(또는 코어 헤더 및 다른 코어 요소들) 후, 다음의 LPSM 값들(파라미터들)을 포함한다: 코어 요소 값들(예로서, 표 1에 특정된 바와 같이)을 따르는 (LPSM으로서 메타데이터를 식별하는) 페이로드 ID; 페이로드 ID를 따르는 (LPSM 페이로드의 크기를 표시한) 페이로드 크기; 및 도 9의 표 2에 표시된 바와 같은 포맷을 가진 (페이로드 ID 및 페이로드 크기 값을 따르는) LPSM 데이터.

[0076] 인코딩된 비트스트림의 제 2 포맷에서, 비트스트림은 AC-3 비트스트림 또는 E-AC-3 비트스트림이며, LPSM을 포함하는 메타데이터 세그먼트들의 각각은 (예로서, 인코더(100)의 스테이지(107)에 의해) 다음 중 하나에 포함된다: 비트스트림의 프레임의 비트스트림 정보("BSI") 세그먼트의 "addbsi" 필드(도 6에 도시); 또는 비트스트림의 프레임의 끝에서 보조데이터 필드(예로서, 도 4에 도시된 AUX 세그먼트). 프레임은 그 각각이 LPSM을 포함하는 하나 또는 두 개의 메타데이터 세그먼트들을 포함할 수 있으며, 프레임이 두 개의 메타데이터 세그먼트들을 포함한다면, 하나는 프레임의 addbsi 필드에 및 다른 하나는 프레임의 AUX 필드에 존재한다. LPSM을 포함한 각각의 메타데이터 세그먼트는 상기 표 1 및 표 2를 참조하여 상기한 특정된 포맷을 가진다(즉, 그것은 표 1에 특정된 코어 요소들, 이어서 상기 특정된 페이로드 ID(LPSM으로서 메타데이터를 식별) 및 페이로드 크기 값들, 이어서 페이로드(표 2에 표시된 바와 같은 포맷을 가진 LPSM 데이터)를 포함한다).

[0077] 또 다른 것에서, 인코딩된 비트스트림은 돌비 E 비트스트림이며, LPSM을 포함하는 메타데이터 세그먼트들의 각각은 돌비 E 보호 밴드 간격의 첫 N 샘플 위치들이다. LPSM을 포함하는 이러한 메타데이터 세그먼트를 포함하는 돌비 E 비트스트림은, 예로서 SMPTE 337M 프리앰블의 Pd 워드에서 시그널링된 LPSM 페이로드 길이를 나타내는 값을 포함한다(SMPTE 337M Pa 워드 반복 레이트는 연관된 비디오 프레임 레이트와 동일한 채로 유지될 수 있다).

[0078] 인코딩된 비트스트림이 E-AC-3 비트스트림인 포맷에서, LPSM을 포함하는 메타데이터 세그먼트들의 각각은 (예로서, 인코더(100)의 스테이지(107)에 의해) 비트스트림의 프레임의 비트스트림 정보("BSI") 세그먼트의 "addbsi" 필드에서 추가적인 비트스트림 정보로서 포함된다. 이러한 포맷에서 LPSM을 갖는 E-AC-3 비트스트림을 인코딩하는 추가적인 양상들은 다음과 같이 설명된다: (1) E-AC-3 비트스트림의 생성 동안, (LPSM 값들을 비트스트림에 삽입하는) E-AC-3 인코더는 발생된 매 프레임(syncframe)에 대해 "활성"인 동안, 비트스트림은 프레임의 addbsi 필드에서 운반된 (LPSM을 포함한) 메타데이터 블록을 포함해야 한다. 메타데이터 블록을 운반하기 위해 요구된 비트들은 인코더 비트레이트(프레임 길이)를 증가시키지 않아야 한다; (2) (LPSM을 포함한) 매 메타데이터 블록은 다음의 정보를 포함해야 한다: 라우드니스_교정_유형_플래그: 여기에서 '1'은 대응하는 오디오 데이터의 라우드니스가 인코더로부터 업스트림에서 교정되었음을 표시하며, '0'은 라우드니스가 인코더에 내장된 라우드니스 교정기에 의해 교정되었음을 표시한다(예로서, 도 2의 인코더(100)의 라우드니스 프로세서(103)); 스피치_채널: (이전 0.5 초에 걸쳐) 어떤 소스 채널(들)이 스피치를 포함하는지를 표시한다. 어떤 스피치도 검출되지 않았다면, 이것은 이와 같이 표시될 것이다; 스피치_라우드니스: (이전 0.5 초에 걸쳐) 스피치를 포함하는 각각의 대응하는 오디오 채널의 통합된 스피치 라우드니스를 표시한다; ITU_라우드니스: 각각의 대응하는 오디오 채널의 통합된 ITU BS.1770-2 라우드니스를 표시한다; 이득: (가역성을 보여주기 위해) 디코더에서의 역전(reversal)을 위한 라우드니스 복합 이득(들).

[0079] (LPSM 값들을 비트스트림에 삽입하는) E-AC-3 인코더가 "활성"이고 "신뢰" 플래그를 가진 AC-3 프레임을 수신하는 동안, 인코더의 라우드니스 제어기(예로서, 도 2의 인코더(100)의 라우드니스 프로세서(103))는 바이패싱된다. "신뢰된" 소스 Dialnorm 및 DRC 값들은 E-AC-3 인코더 구성요소(예로서, 인코더(100)의 스테이지(107))로 (예로서, 인코더(100)의 발생기(106)에 의해) 통과된다. LPSM 블록 생성은 계속되며 라우드니스_교정_유형_플래그는 '1'로 설정된다. 라우드니스 제어기 바이패스 시퀀스는 '신뢰된' 플래그가 나타나는 디코딩된 AC-3 프레임의 시작에 동기화된다. 라우드니스 제어기 바이패스 시퀀스는 다음과 같이 구현된다: 레벨러_양 제어는 10 오디오 블록 기간들(즉, 53.3 밀리초)에 걸쳐 9의 값으로부터 0의 값으로 감소되며 레벨러_백_엔드_미터 제어는 바이패스 모드에 위치된다(이 동작은 끊임없는 전이가 되어야 한다). 용어인 레벨러의 '신뢰된' 바이패스는 소스

비트스트림의 Dialnorm 값이 또한 인코더의 출력에서 재-이용됨을 의미한다. (예로서, '신뢰된' 소스 비트스트림이 -30의 Dialnorm 값을 가진다면, 인코더의 출력은 아웃바운드 Dialnorm 값에 대해 -30을 사용해야 한다).

[0080] (LPSM 값들을 비트스트림에 삽입하는) E-AC-3 인코더가 "활성"이고 '신뢰된' 플래그 없이 AC-3 프레임을 수신하는 동안, 인코더에 내장된 라우드니스 제어기(예로서, 도 2의 인코더(100)의 라우드니스 프로세서(103))는 활성화이다. LPSM 블록 생성은 계속되며 라우드니스_교정_유형_플래그는 '0'으로 설정된다. 라우드니스 제어기 활성화 시퀀스는 '신뢰' 플래그가 사라지는 디코딩된 AC-3 프레임의 시작에 동기화된다. 라우드니스 제어기 활성화 시퀀스는 다음과 같이 구현된다: 레벨러_양 제어는 1 오디오 블록 기간(즉, 5.3 밀리초)에 걸쳐 0의 값으로부터 9의 값으로 증가되며 레벨러_백_엔드_미터 제어는 '활성' 모드로 위치된다(이 동작은 끊임없는 전이가 되며 백_엔드-미터 통합 리셋을 포함한다); 및 인코딩 동안, 그래픽 사용자 인터페이스(GUI)는 다음의 파라미터들을 사용자에게 표시한다: "입력 오디오 프로그램: [신뢰됨/신뢰되지 않음]" - 이러한 파라미터의 상태는 입력 신호 내에서 "신뢰" 플래그의 존재에 기초한다; 및 "실시간 라우드니스 교정: [인에이블/디스에이블]" - 이 파라미터의 상태는 인코더에 내장된 이러한 라우드니스 제어기가 활성화인지 여부에 기초한다.

[0081] 비트스트림의 각각의 프레임의 비트스트림 정보("BSI") 세그먼트의 "addbsi" 필드에 포함된 (설명된 포맷의) LPSM을 가진 AC-3 또는 E-AC-3 비트스트림을 디코딩할 때, 디코더는 (addbsi 필드에서) LPSM 블록 데이터를 파싱하며 추출된 LPSM 값들을 그래픽 사용자 인터페이스(GUI)로 전달한다. 추출된 LPSM 값들의 세트는 매 프레임마다 리프레싱된다.

[0082] 또 다른 포맷에서, 인코딩된 비트스트림은 AC-3 비트스트림 또는 E-AC-3 비트스트림이며, LPSM을 포함하는 메타데이터 세그먼트들의 각각은 비트스트림의 프레임의 비트스트림 정보("BSI") 세그먼트의 "addbsi" 필드(도 6에 도시된)에서(또는 Aux 세그먼트에서) 부가적인 비트 스트림 정보로서 포함된다(예로서, 인코더(100)의 스테이지(107)에 의해). (표 1 및 표 2를 참조하여 상기 설명된 포맷에 대한 변형인) 이 포맷에서, LPSM을 포함하는 addbsi(또는 Aux) 필드들의 각각은 다음의 LPSM 값들을 포함한다: 표 1에 특정된 코어 요소들, 이어서 (LPSM으로서 메타데이터를 식별하는) 페이로드 ID 및 페이로드 크기 값들, 이어서 (상기 표 2에 표시된 요소들과 유사한) 다음의 포맷을 가진 페이로드(LPSM 데이터): LPSM 페이로드의 버전: LPSM 페이로드의 버전을 표시하는 2-비트 필드; dialchan: 대응하는 오디오 데이터의 좌, 우 및/또는 중심 채널들이 음성 다이얼로그를 포함하는지 여부를 표시하는 3-비트 필드. dialchan 필드의 비트 할당은 다음과 같을 수 있다: 좌 채널에서 다이얼로그의 존재를 표시하는 비트 0은 dialchan 필드의 최상위 비트에 저장되며; 중심 채널에서 다이얼로그의 존재를 표시하는 비트 2는 dialchan 필드의 최하위 비트에 저장된다. dialchan 필드의 각각의 비트는 대응하는 채널이 프로그램의 이전 0.5 초 동안 음성 다이얼로그를 포함한다면 '1'로 설정된다; loudregtyp: 프로그램 라우드니스가 어떤 라우드니스 규정 표준을 따르는지를 표시하는 3-비트 필드. "loudregtyp" 필드를 '000'으로 설정하는 것은 LPSM이 라우드니스 규정 준수를 표시하지 않음을 표시한다. 예를 들면, 이러한 필드의 하나의 값(예로서, 000)은 라우드니스 규정 표준의 준수가 표시되지 않음을 표시할 수 있고, 이러한 필드의 또 다른 값(예로서, 001)은 프로그램의 오디오 데이터가 ATSC A/85 표준을 준수함을 표시할 수 있으며, 이러한 필드의 또 다른 값(예로서, 010)은 프로그램의 오디오 데이터가 EBU R128 표준을 준수함을 표시할 수 있다. 상기 예에서, 필드가 '000'외의 임의의 값으로 설정된다면, 페이로드에서 loudcorrldialgat 및 loudcorrtyp 필드들이 뒤따라야 한다; loudcorrldialgat: 다이얼로그-게이팅 라우드니스 교정이 적용되었는지를 표시하는 1-비트 필드. 프로그램의 라우드니스가 다이얼로그 게이팅을 사용하여 교정되었다면, loudcorrldialgat 필드의 값은 '1'로 설정된다: 그렇지 않다면 그것은 '0'으로 설정된다; loudcorrtyp: 프로그램에 적용된 라우드니스 교정의 유형을 표시하는 1-비트 필드. 프로그램의 라우드니스가 무한 예견(infinite look-ahead)(파일-기반) 라우드니스 교정 프로세스로 교정되었다면, loudcorrtyp 필드의 값은 '0'으로 설정된다. 프로그램의 라우드니스가 실시간 라우드니스 측정 및 동적 범위 제어의 조합을 사용하여 교정되었다면, 이 필드의 값은 '1'로 설정된다; loudrelgate: 상대적 게이팅 라우드니스 데이터(ITU)가 존재하는지 여부를 표시한 1-비트 필드. loudrelgate 필드가 '1'로 설정된다면, 7-비트 ituloudrelgat 필드가 페이로드에서 이어져야 한다; loudrelgat: 상대적 게이팅 프로그램 라우드니스(ITU)를 표시하는 7-비트 필드. 이 필드는 적용되는 Dialnorm 및 동적 범위 압축으로 인한 어떠한 이득 조정들도 없이 ITU-R BS.1770-2에 따라 측정된, 오디오 프로그램의 통합 라우드니스를 표시한다. 0 내지 127의 값들이 0.5 LKFS 단차들로, -58 LKFS 내지 +5.5 LKFS로서 해석된다; loudspchgate: 스피치-게이팅 라우드니스 데이터(ITU)가 존재하는지 여부를 표시하는 1-비트 필드. loudspchgate 필드가 '1'로 설정된다면, 7-비트 loudspchgat 필드가 페이로드에서 이어져야 한다; loudspchgat: 스피치-게이팅 프로그램 라우드니스를 표시하는 7-비트 필드. 이 필드는 적용되는 Dialnorm 및 동적 범위 압축으로 인한 어떠한 이득 조정들도 없이 및 ITU-R BS.1770-3의 식(2)에 따라 측정된, 전체 대응하는 오디오 프로그램의 통합 라우드니스를 표시한다. 0 내지 127의 값들은 0.5 LKFS 단차들로, -58 내지 +5.5 LKFS로서 해석된다; loudstrm3se: 단기(3 초) 라우드니스 데이터가 존재하

는지 여부를 표시하는 1-비트 필드. 필드가 '1'로 설정된다면, 7-비트 loudstrm3s 필드가 페이로드에서 이어져야 한다; loudstrm3s: 적용되는 Dialnorm 및 동적 범위 압축으로 인한 어떠한 이득 조정들도 없이, ITU-R BS.1771-1에 따라 측정되는 대응하는 오디오 프로그램의 이전 3초의 게이팅되지 않은 라우드니스를 표시하는 7-비트 필드. 0 내지 256의 값들은 0.5 LKFS 단차들로 -116 LKFS 내지 +11.5 LKFS로서 해석된다; truepk: 실제 피크 라우드니스 데이터가 존재하는지 여부를 표시하는 1-비트 필드. truepk 필드가 '1'로 설정된다면, 8-비트 truepk 필드가 페이로드에서 이어져야 한다; 및 truepk: 적용되는 Dialnorm 및 동적 범위 압축으로 인한 어떠한 이득 조정들도 없이 ITU-R BS.1770-3의 부록 2에 따라 측정되는, 프로그램의 실제 피크 샘플 값을 표시하는 8-비트 필드. 0 내지 256의 값들이 0.5 LKFS 단차들로 -116 LKFS 내지 +11.5 LKFS로서 해석된다.

[0083] 일부 실시예들에서, AC-3 비트스트림 또는 E-AC-3 비트스트림의 프레임의 보조데이터 필드(또는 "addbsi" 필드)에서의 메타데이터 세그먼트의 코어 요소는 코어 헤더(선택적으로 식별 값들, 예로서 코어 요소 버전을 포함), 및 상기 코어 헤더 후: 핑거프린트 데이터가 메타데이터 세그먼트의 메타데이터를 위해 포함되는지(또는 다른 보호 값들이 포함되는지)를 표시한 값들, (메타데이터 세그먼트의 메타데이터에 대응하는 오디오 데이터와 관련된) 외부 데이터가 존재하는지 여부를 표시한 값들, 코어 요소에 의해 식별된 메타데이터의 각각의 유형(예로서, LPSM, 및/또는 LPSM 이외의 유형의 메타데이터)에 대한 페이로드 ID 및 페이로드 크기 값들, 및 코어 요소에 의해 식별된 메타데이터의 적어도 한 유형에 대한 보호 값들을 포함한다. 메타데이터 세그먼트의 메타데이터 페이로드(들)는 코어 헤더를 따르며, (몇몇 경우들에서) 코어 요소의 값들 내에 내포된다.

[0084] 최적화된 라우드니스 및 동적 범위 시스템

[0085] 상기 설명된 안전한 메타데이터 코딩 및 전송 기법은 도 1에 예시된 바와 같이, 상이한 재생 디바이스들, 애플리케이션들, 및 청취 환경들에 걸쳐 라우드니스 및 동적 범위를 최적화하기 위한 확대 가능하며 확장 가능한 시스템과 함께 사용된다. 실시예에서, 시스템(10)은 상이한 타겟 라우드니스 값들을 요구하며 상이한 동적 범위 능력들을 가진 다양한 디바이스들에 걸쳐 입력 오디오(11)의 라우드니스 레벨들 및 동적 범위를 정규화하도록 구성된다. 라우드니스 레벨들 및 동적 범위를 정규화하기 위해, 시스템(10)은 오디오 콘텐츠를 가진 상이한 디바이스 프로파일들을 포함하며 정규화는 이들 프로파일들에 기초하여 행해진다. 상기 프로파일들은 오디오 프로세싱 체인들에서 오디오 프로세싱 유닛들 중 하나에 의해 포함될 수 있으며 상기 포함된 프로파일들은 타겟 디바이스에 대한 원하는 타겟 라우드니스 및 동적 범위를 결정하기 위해 상기 오디오 프로세싱 체인에서 다운스트림 프로세싱 유닛에 의해 사용될 수 있다. 부가적인 프로세싱 구성요소들이 ((이에 제한되지 않지만) 널 대역 범위, 실제 피크 임계치, 라우드니스 범위, 고속/저속 시간 상수(계수들) 및 최대 부스트의 파라미터들을 포함하는) 디바이스 프로파일 관리, 이득 제어 및 광대역 및/또는 다중대역 이득 생성 함수들을 위한 정보를 제공하거나 또는 프로세싱할 수 있다.

[0086] 도 10은 일부 실시예들 하에서, 최적화된 라우드니스 및 동적 범위 제어를 제공하는 시스템에 대한 도 1의 시스템의 보다 상세한 다이어그램을 예시한다. 도 10의 시스템(321)에 대해, 인코더 스테이지는 디코더(312)로의 송신을 위해 적절한 디지털 포맷으로 오디오 입력(303)을 인코딩하는 코어 인코더 구성요소(304)를 포함한다. 상기 오디오는, 그 각각이 상이한 라우드니스 및/또는 동적 범위 타겟 설정들을 요구할 수 있는, 다양한 상이한 청취 환경들에서 재생될 수 있도록 프로세싱된다. 따라서, 도 10에 도시된 바와 같이, 전체 범위 스피커들(320), 소형 스피커들(322), 및 헤드폰들(324)을 포함한 다양한 상이한 드라이버 유형들을 통한 재생을 위해 디코더는 디지털-대-아날로그 변환기(316)에 의해 아날로그 포맷으로 변환되는 디지털 신호를 출력한다. 이들 드라이버들은 가능한 재생 드라이버들의 몇몇 예들만을 예시하며, 어떠한 적절한 크기 및 유형의 어떠한 트랜듀서 또는 드라이버도 사용될 수 있다. 또한, 도 10의 드라이버들/트랜듀서들(320 내지 324)은 임의의 대응하는 청취 환경에서의 사용을 위한 어떠한 적절한 재생 디바이스에서도 구현될 수 있다. 디바이스 유형들은, 예를 들면, AVR들, 텔레비전들, 스테레오 장비, 컴퓨터들, 이동 전화들, 태블릿 컴퓨터들, MP3 플레이어들 등을 포함할 수 있으며, 청취 환경들은, 예를 들면, 강당들, 가정들, 차들, 청취 부스들 등을 포함할 수 있다.

[0087] 재생 환경들 및 드라이버 유형들의 범위는 매우 작은 사설 콘텍스트들에서 매우 큰 공공 장소들로 변할 수 있기 때문에, 가능하며 최적의 재생 라우드니스 및 동적 범위 구성들의 폭은 콘텐츠 유형, 배경 잡음 레벨들 등에 의존하여 크게 달라질 수 있다. 예를 들면, 홈 시어터 환경에서, 광범위한 동적 범위 콘텐츠가 서라운드 사운드 장비를 통해 플레이될 수 있으며 보다 좁은 동적 범위 콘텐츠가 (평판 LED/LCD 유형과 같은) 통상의 텔레비전 시스템을 통해 플레이될 수 있는 반면, 매우 좁은 동적 범위 모드는 큰 레벨 변형들이 바람직하지 않은 특정한 청취 조건들을 위해 사용될 수 있다(예로서, 밤에 또는 심각한 음향 출력 전력 한계들을 가진 디바이스, 예로서 이동 전화/태블릿 내부 스피커들 또는 헤드폰 출력 상에서). 소형 컴퓨터 또는 독 스피커들, 또는 헤드폰들/이어버드들을 사용하는 것과 같은 휴대용 또는 이동 청취 콘텍스트들에서, 재생의 최적의 동적 범위는 환경에 의

존하여 달라질 수 있다. 예를 들면, 조용한 환경에서, 최적의 동적 범위는 시끄러운 환경과 비교하여 더 클 수 있다. 도 10의 적응적 오디오 프로세싱 시스템의 실시예들은 청취 디바이스 환경 및 재생 디바이스 유형과 같은, 파라미터들에 의존하여 오디오 콘텐츠를 보다 쉽게 이해할 수 있게 하기 위해 동적 범위를 변경할 것이다.

[0088] 도 11은 예시적인 사용 경우에 있어서 다양한 재생 디바이스들 및 배경 청취 환경들을 위한 상이한 동적 범위 조건들을 예시하는 표이다. 유사한 조건들이 라우드니스를 위해 얻어질 수 있다. 상이한 동적 범위 및 라우드니스 조건들은 최적화 시스템(321)에 의해 사용되는 상이한 프로파일들을 생성한다. 시스템(321)은 입력 오디오의 라우드니스 및 동적 범위를 분석 및 측정하는 라우드니스 및 동적 범위 측정 구성요소(302)를 포함한다. 실시예에서, 시스템은 전체 라우드니스 파라미터를 결정하기 위해 전체 프로그램 콘텐츠를 분석한다. 이러한 콘텍스트에서, 라우드니스는 프로그램의 장기 프로그램 라우드니스 또는 평균 라우드니스를 나타내며, 여기에서 프로그램은 영화, 텔레비전 쇼, 광고, 또는 유사한 프로그램 콘텐츠와 같은, 오디오 콘텐츠의 단일 유닛이다. 라우드니스는 오디오가 어떻게 재생될 것인지를 제어하기 위해 콘텐츠 생성기들에 의해 사용되는 예술적 동적 범위 프로파일의 표시를 제공하기 위해 사용된다. 라우드니스는 Dialnorm이 단일 프로그램(예로서, 영화, 쇼, 광고 등)의 평균 다이얼로그 라우드니스를 표현한다는 점에서 Dialnorm 메타데이터 값과 관련된다. 단기 동적 범위는 프로그램 라우드니스보다 훨씬 더 짧은 시간 기간에 걸쳐 신호들에서의 변형들을 수량화한다. 예를 들면, 단기 동적 범위는 초들(seconds)의 정도로 측정될 수 있는 반면, 프로그램 라우드니스는 분들 또는 심지어 시간들의 폭에 걸쳐 측정될 수 있다. 단기 동적 범위는 오버로딩이 다양한 재생 프로파일들 및 디바이스 유형들에 대해 발생하지 않음을 보장하기 위해 프로그램 라우드니스에 독립적인 보호 메커니즘을 제공한다. 실시예에서, 라우드니스(장기 프로그램 라우드니스) 타겟은 다이얼로그 라우드니스에 기초하며 단기 동적 범위는 상대적-게이팅된 및/또는 게이팅되지 않은 라우드니스에 기초한다. 이 경우에, 시스템에서 특정한 DRC 및 라우드니스 구성요소들은 콘텐츠 유형 및/또는 타겟 디바이스 유형들 및 특성들에 관한 콘텍스트-인식이다. 이러한 콘텍스트-인식 능력의 일부로서, 시스템은 디바이스가 AVR 유형 디바이스들, 텔레비전들, 컴퓨터들, 휴대용 디바이스들 등과 같은, 특정한 DRC 및 라우드니스 재생 조건들에 대해 최적화되는 특정한 그룹들의 디바이스들의 멤버인지 여부를 결정하기 위해 출력 디바이스의 하나 이상의 특성들을 분석하도록 구성된다.

[0089] 전-처리 구성요소는 복수의 상이한 프로파일들의 각각의 프로파일에 대한 고유한 메타데이터를 생성하기 위해 라우드니스, 피크들, 실제 피크들, 및 조용한 기간들을 결정하도록 상기 프로그램 콘텐츠를 분석한다. 실시예에서, 라우드니스는 다이얼로그-게이팅 라우드니스 및/또는 상대적-게이팅 라우드니스일 수 있다. 상이한 프로파일들이 다양한 DRC(동적 범위 제어) 및 타겟 라우드니스 모드들을 정의하며, 여기에서 상이한 이득 값들이 소스 오디오 콘텐츠의 특성들, 원하는 타겟 라우드니스 및 재생 디바이스 유형 및/또는 환경에 의존하여 인코더에서 생성된다. 상기 디코더는 (상기 언급된 프로파일들에 의해 인에이블된) 상이한 DRC 및 타겟 라우드니스 모드들을 제공할 수 있으며 오디오 신호의 압축이 없고 라우드니스 정규화가 없는 전체 동적 범위 목록을 허용하는 오프/디스에이블된 DRC 및 타겟 라우드니스, -31 LKFS의 타겟을 가진 라우드니스 정규화를 갖고 인코더에서 (구체적으로 이러한 재생 모드 및/또는 디바이스 프로파일을 위해) 생성된 이득 값들을 통해 적절한 동적 범위 압축을 제공하는, 홈 시어터 시스템들 상에서의 재생을 위한 -31 LKFS 라인 모드의 타겟을 가진 오프/디스에이블된 DRC 및 라우드니스 정규화; -24, -23 또는 -20 LKFS 중 하나의 타겟으로 라우드니스 정규화를 갖는 많은 양의 동적 범위 압축을 제공하는 TV 스피커들을 통한 재생을 위한 RF 모드, -14 LKFS의 타겟에서 라우드니스 정규화를 가진 압축을 제공하는, 컴퓨터들 또는 유사한 디바이스들에 걸친 재생을 위한 중간 모드, 및 -11 LKFS의 라우드니스 정규화 타겟을 가진 매우 많은 동적 범위 압축을 제공하는 휴대용 모드를 포함할 수 있다. -31, -23/-20, -14, 및 -11 LKFS의 타겟 라우드니스 값들은 일부 실시예들 하에서 시스템을 위해 정의될 수 있는 상이한 재생/디바이스 프로파일들의 예들이 되도록 의도되며, 임의의 다른 적절한 타겟 라우드니스 값들이 사용될 수 있고, 시스템은 이들 재생 모드들 및/또는 디바이스 프로파일에 특정한 적절한 이득 값들을 생성한다. 더욱이, 시스템은 확장 및 적응 가능하여, 상이한 재생 디바이스들 및 청취 환경들이 인코더 또는 그 밖의 것에서 새로운 프로파일을 정의함으로써 수용되어 인코더로 로딩될 수 있다. 이러한 방식으로, 새로운 고유의 재생/디바이스 프로파일들이 추가 응용을 위해 개선된 또는 상이한 재생 디바이스들을 지원하기 위해 생성될 수 있다.

[0090] 실시예에서, 이득 값들은 인코더(304), 디코더(312), 또는 트랜스코더(308)와 같은, 시스템(321)의 임의의 적절한 프로세싱 구성요소, 또는 인코더와 연관된 임의의 연관된 전-처리 구성요소 또는 디코더와 연관된 임의의 후-처리 구성요소에서 산출될 수 있다.

[0091] 도 13은 실시예 하에서, 다양한 상이한 재생 디바이스 클래스들을 위한 상이한 프로파일들 사이에서의 인터페이스를 예시한 블록도이다. 도 13에 도시된 바와 같이, 인코더(502)는 오디오 입력(501) 및 여러 개의 상이한 가

능한 프로파일들(506) 중 하나를 수신한다. 인코더는 타겟 재생 디바이스에 존재하거나 또는 그것과 연관된 디코더 구성요소에서 프로세싱되는 출력 비트스트림 파일을 생성하기 위해 선택된 프로파일과 오디오 데이터를 결합한다. 도 13의 예에서, 상이한 재생 디바이스들은 컴퓨터(510), 이동 전화(512), AVR(514), 및 텔레비전(516)일 수 있지만, 많은 다른 출력 디바이스들이 또한 가능하다. 디바이스들(510 내지 516)의 각각은 드라이버들(320 내지 324)과 같은 (드라이버들 및/또는 트랜듀서들을 포함한) 스피커들을 포함하거나 또는 그것에 결합된다. 재생 디바이스들 및 연관된 스피커들의 프로세싱, 출력 정격들, 및 크기들의 조합은 일반적으로 어떤 프로파일이 상기 특정한 타겟에 대해 가장 최적인지를 설명한다. 따라서, 상기 프로파일들(506)은 구체적으로 AVR들, TV들, 이동 스피커들, 이동 헤드폰들 등을 통한 재생을 위해 정의될 수 있다. 그것들은 또한 조용한 모드, 야간 모드, 옥외, 실내 등과 같은, 특정 동작 모드들 또는 상태들을 위해 정의될 수 있다. 도 13에 도시된 프로파일들은 단지 예시적인 모드들이며 특정 타겟들 및 환경들을 위한 맞춤 프로파일들을 포함한, 임의의 적절한 프로파일이 정의될 수 있다.

[0092] 도 13은 인코더(502)가 프로파일들(506)을 수신하며 라우드니스 및 DRC 프로세싱을 위한 적절한 파라미터들을 생성하는 실시예를 예시하지만, 프로파일 및 오디오 콘텐츠에 기초하여 생성된 파라미터들은 인코더, 디코더, 트랜코더, 전-처리, 후-처리 등과 같은 어떠한 적절한 오디오 프로세싱 유닛 상에서도 수행될 수 있다는 것이 주의되어야 한다. 예를 들면, 도 13의 각각의 출력 디바이스(510 내지 516)는 라우드니스 및 동적 범위의 적응화를 상기 디바이스 또는 타겟 출력 디바이스의 디바이스 유형에 매칭할 수 있게 하기 위해 인코더(502)로부터 전송된 파일(504)의 비트스트림에서 메타데이터를 프로세싱하는 디코더 구성요소를 갖거나 또는 그에 결합된다.

[0093] 실시예에서, 오디오 콘텐츠의 동적 범위 및 라우드니스는 각각의 가능한 재생 디바이스에 대해 최적화된다. 이것은 (신호 역학들, 샘플 피크들 및/또는 실제 피크들을 제어함으로써) 타겟 재생 모드들의 각각에 대한 오디오 경험을 최적화하기 위해 장기 라우드니스를 타겟으로 유지하며 단기 동적 범위를 제어함으로써 달성된다. 상이한 메타데이터 요소들이 장기 라우드니스 및 단기 동적 범위에 대해 정의된다. 도 10에 도시된 바와 같이, 구성요소(302)는 이들 별개의 DR 구성요소들의 양쪽 모두에 대한 관련 특성들을 도출하기 위해 전체 입력 오디오 신호(또는 적용 가능하다면, 스피커 구성요소와 같은 그것의 부분들)를 분석한다. 이것은 상이한 이득 값들이 예상적 이득들 대 클립(오버로드 보호) 이득 값들에 대해 정의될 수 있도록 한다.

[0094] 장기 라우드니스 및 단기 동적 범위에 대한 이들 이득 값들은 그 후 라우드니스 및 동적 범위 제어 이득 값들을 설명하는 파라미터들을 생성하기 위해 프로파일(305)에 매핑된다. 이들 파라미터들은 다중화기(306), 또는 트랜스코더(308)를 통해 디코더 스테이지로 송신되는 비트스트림의 생성을 위한 유사한 구성요소에서 인코더(304)로부터의 인코딩된 오디오 신호와 결합된다. 디코더 스테이지로 입력된 비트스트림은 역다중화기(310)에서 역다중화된다. 그것은 그 후 디코더(312)에서 디코딩된다. 이득 구성요소(314)는 그 후 적절한 재생 디바이스들 및 드라이버들 또는 트랜듀서들(320 내지 324)을 통한 재생을 위해 DACS 유닛(416)을 통해 프로세싱되는 디지털 오디오 데이터를 생성하기 위해 적절한 프로파일에 대응하는 이득들을 적용한다.

[0095] 도 14는 실시예 하에서, 복수의 정의된 프로파일들에 대한 장기 라우드니스 및 단기 동적 범위 사이에서의 상관관계를 예시하는 표이다. 도 14의 표 4에 도시된 바와 같이, 각각의 프로파일은 시스템의 디코더에서 또는 타겟 디바이스의 각각에서 적용된 동적 범위 압축(DRC)의 양을 디테이트하는 이득 값들의 세트를 포함한다. 프로파일들(1 내지 N)로 표시된 N개의 프로파일들의 각각은, 디코더 스테이지에서 적용된 대응하는 이득 값들을 디테이트함으로써, 특정한 장기 라우드니스 파라미터들(예로서, Dialnorm) 및 오버로드 압축 파라미터들을 설정한다. 상기 프로파일들에 대한 DRC 이득 값들은 인코더에 의해 수용되는 외부 소스에 의해 정의될 수 있거나, 또는 그것들은 외부 값들이 제공되지 않는다면 디폴트 이득 값들로서 인코더 내에서 내부적으로 생성될 수 있다.

[0096] 실시예에서, 각각의 프로파일에 대한 이득 값들은 각각의 가능한 디바이스 프로파일 및/또는 타겟 라우드니스에 대한 최종 DRC 이득들의 고속/저속 어택(attack) 및 고속/저속 해제를 실행하기 위해 필요한 시간 상수들뿐만 아니라 선택된 프로파일(즉, 전송 특성 또는 곡선)에 기초하여 정적 이득을 계산하기 위해, 피크, 실제 피크, 다이얼로그의 단기 라우드니스 또는 전체 단기 라우드니스 또는 양쪽 모두의 조합(하이브리드)과 같은, 오디오 신호의 특정한 특성들의 분석에 기초하여 계산되는 DRC 이득 워드들에서 구현된다. 상기 서술된 바와 같이, 이들 프로파일들은 인코더, 디코더에서 존재할 수 있거나 또는 외부적으로 생성되며 콘텐츠 생성기로부터 외부 메타데이터를 통해 인코더에 전달된다.

[0097] 실시예에서, 이득 값들은 오디오 콘텐츠의 전체 주파수들에 걸쳐 동일한 이득을 적용하는 광대역 이득일 수 있다. 대안적으로, 상기 이득은 상이한 이득 값들이 상이한 주파수들 또는 오디오 콘텐츠의 주파수 대역들에 적용

되도록 다중-대역 이득 값들로 구성될 수 있다. 다-채널 경우에, 각각의 프로파일은 단일 이득 값 대신에 상이한 주파수 대역들에 대한 이득들을 표시한 이득 값들의 행렬을 구성할 수 있다.

[0098] 도 10을 참조하면, 실시예에서, 청취 환경들 및/또는 재생 디바이스들의 능력들 및 구성들의 속성들 또는 특성들에 관한 정보가 디코더 스테이지에 의해, 피드백 링크(330)에 의해 인코더 스테이지에 제공된다. 프로파일 정보(332)는 또한 인코더(304)에 입력된다. 실시예에서, 디코더는 제 1 그룹의 오디오 재생 디바이스들에 대한 라우드니스 파라미터가 비트스트림에서 이용 가능하지 여부를 결정하기 위해 비트스트림내의 메타데이터를 분석한다. 그렇다면, 그것은 오디오를 렌더링할 때 사용하기 위해 다운스트림으로 파라미터들을 송신한다. 그렇지 않다면, 인코더는 파라미터들을 도출하기 위해 디바이스들의 특정한 특성들을 분석한다. 이들 파라미터들은 그 후 재생을 위해 다운스트림 렌더링 구성요소로 전송된다. 상기 인코더는 또한 수신된 오디오 스트림을 렌더링할 출력 디바이스(또는 출력 디바이스를 포함한 출력 디바이스들의 그룹)를 결정한다. 예를 들면, 출력 디바이스는 셀 전화인 것으로 또는 그룹 형 휴대용 디바이스들에 속하는 것으로 결정될 수 있다. 실시예에서, 디코더는 결정된 출력 디바이스 또는 출력 디바이스들의 그룹을 인코더에 표시하기 위해 피드백 링크(330)를 사용한다. 이러한 피드백에 대해, 출력 디바이스에 접속된 모듈(예로서, 헤드셋들에 접속되거나 또는 랩탑의 스피커들에 접속된 사운드카드에서의 모듈)은 출력 디바이스의 아이덴티티 또는 출력 디바이스를 포함하는 디바이스들의 그룹의 아이덴티티를 디코더에 표시할 수 있다. 디코더는 이러한 정보를 피드백 링크(330)를 통해 인코더에 송신한다. 실시예에서, 디코더는 라우드니스 및 DRC 파라미터들을 결정하기 위해 디코더를 수행한다. 실시예에서, 디코더는 라우드니스 및 DRC 파라미터들을 결정한다. 이 실시예에서, 피드백 링크(330)를 통해 정보를 송신하는 대신에, 디코더는 라우드니스 및 DRC 파라미터들을 결정하기 위해 결정된 디바이스 또는 출력 디바이스들의 그룹에 대한 정보를 사용한다. 또 다른 실시예에서, 또 다른 오디오 프로세싱 유닛은 라우드니스 및 DRC 파라미터들을 결정하며 디코더는 상기 정보를 디코더 대신에 상기 오디오 프로세싱 유닛에 송신한다.

[0099] 도 12는 실시예 하에서, 동적 범위 최적화 시스템의 블록도이다. 도 12에 도시된 바와 같이, 인코더(402)는 입력 오디오(401)를 수신한다. 상기 인코딩된 오디오는 선택된 압축 곡선(422) 및 Dialnorm 값(424)으로부터 생성된 파라미터들(404)과 다중화기(409)에서 결합된다. 결과적인 비트스트림은 디코더(406)에 의해 디코딩되는 오디오 신호들을 생성하는 역다중화기(411)에 송신된다. 상기 파라미터들 및 Dialnorm 값들은 디코더 출력의 증폭을 위해 증폭기(410)를 구동하는 이득 레벨들을 생성하기 위해 이득 산출 유닛(408)에 의해 사용된다. 도 12는 어떻게 동적 범위 제어가 파라미터화되고 비트스트림에 삽입되는지를 예시한다. 라우드니스가 또한 유사한 구성요소들을 사용하여 파라미터화되고 비트스트림에 삽입될 수 있다. 실시예에서, 출력 기준 레벨 제어(도시되지 않음)가 또한 디코더에 제공될 수 있다. 도면은 인코더에서 결정되어 삽입되는 것으로서 라우드니스 및 동적 범위 파라미터들을 예시하지만, 유사한 결정이 전-처리, 디코더, 및 후-처리와 같은 다른 오디오 프로세싱 유닛들에서 수행될 수 있다.

[0100] 도 15는 실시예 하에서, 상이한 유형들의 오디오 콘텐츠에 대한 라우드니스 프로파일들의 예들을 예시한다. 도 15에 도시된 바와 같이, 예시적인 곡선들(600 및 602)은 0 LKFS 주위에 중심을 둔 이득에 대한 입력 라우드니스(LKFS로)를 나타낸다. 상이한 유형들의 콘텐츠가 도 15에 도시된 바와 같이 상이한 곡선들을 보여주며, 여기에서 곡선(600)은 스피치를 나타내며 곡선(602)은 표준 필름 콘텐츠를 나타낼 수 있다. 도 15에 도시된 바와 같이, 스피치 콘텐츠는 필름 콘텐츠에 비해 보다 많은 양의 이득의 대상이 된다. 도 15는 특정한 유형들의 오디오 콘텐츠에 대한 대표적인 프로파일 곡선들의 예들이 되도록 의도되며, 다른 프로파일 곡선들이 또한 사용될 수 있다. 도 15에 도시된 바와 같은, 프로파일 특성들의 특정한 양상들이 최적화 시스템에 대한 관련 파라미터들을 도출하기 위해 사용된다. 실시예에서, 이들 파라미터들은: 널 대역폭, 컷 비(cut ratio), 부스트 비, 최대 부스트, FS 어택, FS 감쇠, 홀드오프(holdoff), 피크 제한, 및 타겟 레벨 라우드니스를 포함한다. 애플리케이션 요건들 및 시스템 제약들에 의존하여 이들 파라미터들 중 적어도 일부에 부가하여 또는 대안적으로 다른 파라미터들이 사용될 수 있다.

[0101] 도 16은 실시예 하에서, 재생 디바이스들 및 애플리케이션들에 걸쳐 라우드니스 및 동적 범위를 최적화하는 방법을 예시하는 흐름도이다. 도면은 인코더에서 수행되는 것으로서 라우드니스 및 동적 범위 최적화를 예시하지만, 유사한 최적화가 전-처리, 디코더, 및 후-처리와 같은 다른 오디오 프로세싱 유닛들에서 수행될 수 있다. 프로세스(620)에서 도시된 바와 같이, 방법은 인코더 스테이지가 소스로부터 입력 신호를 수신하는 것에서 시작한다(603). 상기 인코더 또는 전-처리 구성요소는 그 후 소스 신호가 타겟 라우드니스 및/또는 동적 범위를 달성하는 프로세스를 받았는지 여부를 결정한다(604). 상기 타겟 라우드니스는 장기 라우드니스에 대응하며 외부에서 또는 내부적으로 정의될 수 있다. 소스 신호가 타겟 라우드니스 및/또는 동적 범위를 달성하기 위한 프로세스를 받지 않았다면, 시스템은 적절한 라우드니스 및/또는 동적 범위 제어 동작을 수행한다(608); 그렇지

않고, 소스 신호가 이러한 라우드니스 및/또는 동적 범위 제어 동작을 받았다면, 시스템은 원래 프로세스가 적절한 장기 라우드니스 및/또는 동적 범위를 디테이트하도록 허용하기 위해 라우드니스 제어 및/또는 동적 범위 동작들을 스킵하도록 바이패스 모드에 들어간다(606). 바이패스 모드(606) 또는 수행 모드(608)(단일 광대역 이득 값들 또는 주파수-종속적 다중-대역 이득 값들일 수 있음) 중 하나에 대한 적절한 이득 값들이 그 후 디코더에 적용된다(612).

[0102] 비트스트림 포맷

[0103] 이전에 기술된 바와 같이, 라우드니스 및 동적 범위를 최적화하기 위한 시스템은 인코더 및 디코더 사이에서, 또는 소스 및 렌더링/재생 디바이스들 사이에서 비트스트림에 송신된 메타데이터 및 오디오 콘텐츠가 서로로부터 분리되지 않았거나 또는 그 외 네트워크들 또는 서비스 제공자 인터페이스들과 같은 다른 독점 장비 등을 통한 송신 동안 변질되지 않았음을 보장하기 위해 안전한 확장 가능한 메타데이터 포맷을 이용한다. 이러한 비트스트림은 적절한 프로파일 정보를 통해 오디오 콘텐츠 및 출력 디바이스 특성들을 맞추도록 오디오 신호의 라우드니스 및 동적 범위를 적응시키기 위해 인코더 및/또는 디코더 구성요소들을 시그널링하기 위한 메커니즘을 제공한다. 실시예에서, 시스템은 인코더 및 디코더 사이에서 송신될 저 비트 레이트 인코딩된 비트스트림을 결정하도록 구성되며, 메타데이터를 통해 인코딩된 라우드니스 정보는 하나 이상의 출력 프로파일들에 대한 특성들을 포함한다. 실시예 하에서 라우드니스 및 동적 범위 최적화 시스템과 함께 사용하기 위한 비트스트림 포맷의 설명이 이어진다.

[0104] AC-3 인코딩된 비트스트림은 오디오 콘텐츠에 대한 1 대 6 채널들 및 메타데이터를 포함한다. 상기 오디오 콘텐츠는 지각적 오디오 코딩을 사용하여 압축된 오디오 데이터이다. 상기 메타데이터는 청취 환경에 전달된 프로그램의 사운드를 변경하는데 사용하기 위해 의도되는 여러 개의 오디오 메타데이터 파라미터들을 포함한다. AC-3 인코딩된 오디오 비트스트림의 각각의 프레임은 디지털 오디오의 1536개의 샘플들에 대한 오디오 콘텐츠 및 메타데이터를 포함한다. 48 kHz의 샘플링 레이트에 대해, 이것은 32 밀리초들의 디지털 오디오 또는 오디오의 초당 31.25 프레임들의 레이트를 나타낸다.

[0105] E-AC-3 인코딩된 오디오 비트스트림의 각각의 프레임은 프레임이 각각 오디오 데이터의 1, 2, 3, 또는 6개의 블록들을 포함하는지에 의존하여, 디지털 오디오의 256, 512, 768, 또는 1536 샘플들에 대한 오디오 콘텐츠 및 메타데이터를 포함한다. 48 kHz의 샘플링 레이트에 대해, 이것은 각각 5.333, 10.667, 16 또는 32 밀리초들의 디지털 오디오 또는 각각 오디오의 초당 189.9, 93.75, 62.5 또는 31.25 프레임들의 레이트를 나타낸다.

[0106] 도 4에 표시된 바와 같이, 각각의 AC-3 프레임은 다음을 포함한 섹션들(세그먼트들)로 분할된다: 동기화 워드(SW) 및 두 개의 에러 교정 워드들 중 첫 번째(CRC1)를 (도 5에 도시된 바와 같이) 포함하는 동기화 정보(SI) 섹션; 메타데이터의 대부분을 포함하는 비트스트림 정보(BSI) 섹션; 데이터 압축 오디오 콘텐츠를 포함하는(및 또한 메타데이터를 포함할 수 있는) 6개의 오디오 블록들(AB0 내지 AB5); 오디오 콘텐츠가 압축된 후 남겨진 임의의 사용되지 않은 비트들을 포함하는 낭비(waste) 비트들(W); 보다 많은 메타데이터를 포함할 수 있는 보조(AUX) 정보 섹션; 및 두 개의 에러 교정 워드들 중 두 번째(CRC2).

[0107] 도 7에 표시된 바와 같이, 각각의 E-AC-3 프레임은 다음을 포함한 섹션들(세그먼트들)로 분할된다: (도 5에 도시된 바와 같이) 동기화 워드(SW)를 포함하는 동기화 정보(SI) 섹션; 메타데이터의 대부분을 포함하는 비트스트림 정보(BSI) 섹션; 데이터 압축 오디오 콘텐츠를 포함하는(및 또한 메타데이터를 포함할 수 있는) 1 및 6 사이의 오디오 블록들(AB0 내지 AB5); 오디오 콘텐츠가 압축된 후 남겨진 임의의 사용되지 않은 비트들을 포함하는 낭비 비트들(W); 보다 많은 메타데이터를 포함할 수 있는 보조(AUX) 정보 섹션; 및 에러 교정 워드(CRC).

[0108] AC-3(또는 E-AC-3) 비트스트림에서, 청취 환경에 전달된 프로그램의 사운드를 변경하는데 특별히 사용하도록 의도되는 여러 개의 오디오 메타데이터 파라미터들이 있다. 메타데이터 파라미터들 중 하나는 Dialnorm 파라미터이며, 이것은 BSI 세그먼트에 포함된다.

[0109] 도 6에 도시된 바와 같이, AC-3 프레임의 BSI 세그먼트는 프로그램에 대한 Dialnorm 값을 표시한 5-비트 파라미터("Dialnorm")를 포함한다. 동일한 AC-3 프레임으로 운반된 제 2 오디오 프로그램에 대한 Dialnorm 값을 표시한 5-비트 파라미터("Dialnorm2")는 AC-3 프레임의 오디오 코딩 모드("acmod")가, 이중-모노 또는 "1+1" 채널 구성이 사용 중임을 나타내는, "0"이면 포함된다.

[0110] BSI 세그먼트는 또한 "addbsie" 비트에 이어 부가적인 비트 스트림 정보의 존재(또는 부재)를 표시한 플래그("addbsie"), "addbsil" 값에 이어 임의의 부가적인 비트 스트림 정보의 길이를 표시한 파라미터("addbsil"), 및 "addbsil" 값에 이어 최대 64 비트들까지의 부가적인 비트 스트림 정보("addbsi")를 포함한다. 상기 BSI 세

그먼트는 도 6에 특별히 도시되지 않은 다른 메타데이터 값들을 포함할 수 있다.

[0111] 여기에 설명된 하나 이상의 실시예들의 양상들은 소프트웨어 지시들을 실행하는 하나 이상의 컴퓨터들 또는 프로세싱 디바이스들을 포함하는 네트워크에 걸친 송신을 위해 오디오 신호들을 프로세싱하는 오디오 시스템에서 구현될 수 있다. 설명된 실시예들 중 어떠한 것도 단독으로 또는 임의의 조합으로 서로 함께 사용될 수 있다. 다양한 실시예들이 종래 기술이 가진 다양한 결점들에 의해 동기 부여가 되었을 수 있고 이러한 것이 본 명세서에서 하나 이상의 부분에서 논의되거나 또는 시사될 수 있지만, 실시예들은 이들 결점들 중 어떠한 것도 반드시 다루지는 않는다. 다시 말해서, 상이한 실시예들이 본 명세서에서 논의될 수 있는 상이한 결점들을 다룰 수 있다. 일부 실시예들은 본 명세서에서 논의될 수 있는 몇몇 결점들 또는 단지 하나의 결점만을 단지 부분적으로 다룰 수 있으며, 일부 실시예들은 이들 결점들의 어떠한 것도 다루지 않을 수 있다.

[0112] 여기에 설명된 시스템들의 양상들은 디지털 또는 디지털화된 오디오 파일들을 프로세싱하기 위한 적절한 컴퓨터-기반 사운드 프로세싱 네트워크 환경에서 구현될 수 있다. 적응적 오디오 시스템의 부분들은 컴퓨터들 중에서 송신된 데이터를 버퍼링하고 라우팅하도록 작용하는 하나 이상의 라우터들(도시되지 않음)을 포함하여, 임의의 원하는 수의 개개의 기계들을 포함하는 하나 이상의 네트워크들을 포함할 수 있다. 이러한 네트워크는 다양한 상이한 네트워크 프로토콜들 상에서 수립될 수 있으며, 인터넷, 광역 네트워크(WAN), 근거리 네트워크(LAN), 또는 그것의 임의의 조합일 수 있다.

[0113] 구성요소들, 블록들, 프로세스들, 또는 다른 기능 구성요소들 중 하나 이상이 시스템의 프로세서-기반 계산 디바이스의 실행을 제어하는 컴퓨터 프로그램을 통해 구현될 수 있다. 여기에 개시된 다양한 기능들은 그것들의 행동, 레지스터 전송, 논리적 구성요소, 및/또는 다른 특성들에 대하여, 하드웨어, 펌웨어의 임의의 수의 조합들을 사용하여 및/또는 다양한 기계-관독 가능한 또는 컴퓨터-관독 가능한 미디어에서 구체화된 데이터 및/또는 지시들로서 설명될 수 있다는 것이 또한 주시되어야 한다. 이러한 포맷팅된 데이터 및/또는 지시들이 구체화될 수 있는 컴퓨터-관독 가능한 미디어는 이에 제한되지 않지만, 광학적, 자기, 또는 반도체 저장 미디어와 같은, 다양한 형태들에서 물리적(비-일시적), 비-휘발성, 저장 미디어를 포함한다.

[0114] 맥락이 달리 명확하게 요구하지 않는다면, 설명 및 청구항들 전체에 걸쳐, 단어들 "포함하다", "포함하는" 등은 배타적 또는 철저한 의미와는 대조적으로 포괄적인 의미로, 즉, "이에 제한되지 않지만, 이를 포함하는"의 의미로 해석될 것이다. 단수형 또는 복수형 숫자를 사용한 단어들은 또한 각각 복수 또는 단수 숫자를 포함한다. 부가적으로, 단어들 "여기에서", "아래에", "상기", "이하에서" 및 유사한 의미의 단어들은 본 출원 명세서의 임의의 특정한 부분들이 아닌, 전체로서 본 출원 명세서를 나타낸다. 단어 "또는" 이 둘 이상의 아이тем들의 리스트를 참조하여 사용될 때, 상기 단어는 상기 단어의 다음의 해석들 모두를 커버한다: 리스트에서의 아이тем들 중 임의의 것. 상기 리스트에서의 아이тем들의 모두 및 상기 리스트에서의 아이тем들의 임의의 조합.

[0115] 하나 이상의 구현들이 예로서 및 특정한 실시예들에 대하여 설명되었지만, 하나 이상의 구현들은 개시된 실시예들에 제한되지 않는다는 것이 이해될 것이다. 반대로, 그것은 이 기술분야의 숙련자들에게 명백할 바와 같이, 다양한 변경들 및 유사한 배열들을 커버하도록 의도된다. 그러므로, 첨부된 청구항들의 범위는 모든 이러한 변경들 및 유사한 배열들을 포괄하도록 가장 광범위한 해석에 부합되어야 한다.

부호의 설명

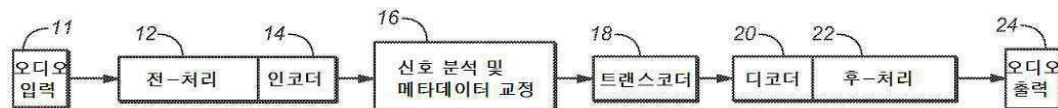
[0116]

10: 시스템	11: 입력
12: 전-처리 유닛	14: 인코더
16: 신호 분석 및 메타데이터 교정 유닛	18: 트랜스코더
20: 디코더	24: 후-처리 유닛
100: 인코더	101: 디코더
102: 오디오 상태 검증기	103: 라우드니스 프로세싱 스테이지
104: 오디오 스트림 선택 스테이지	105: 인코더
106: 메타데이터 발생기	107: 스테리오/포맷터 스테이지
108: 다이얼로그 라우드니스 측정 서브시스템	109: 프레임 버퍼

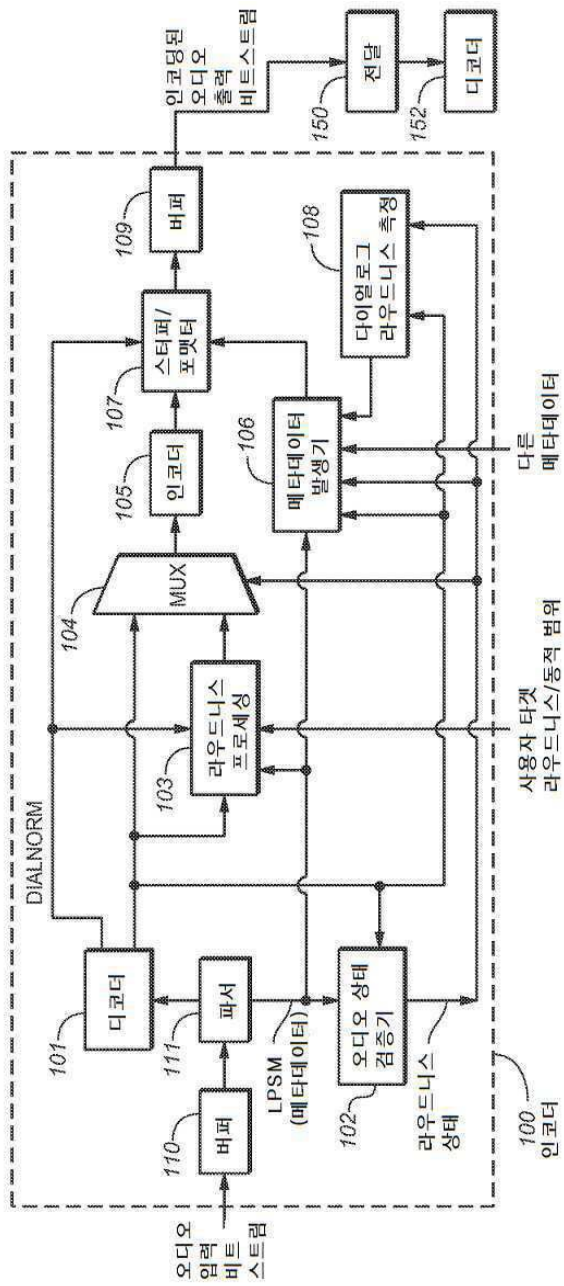
- | | |
|--------------------|---------------------|
| 110: 프레임 버퍼 | 111: 파서 |
| 150: 서브시스템 | 152: 디코더 |
| 200: 디코더 | 201: 프레임 버퍼 |
| 202: 오디오 디코더 | 203: 오디오 상태 검증기 |
| 204: 제어 비트 발생 스테이지 | 205: 파서 |
| 300: 후-처리기 | 301: 프레임 버퍼 |
| 303: 오디오 입력 | 304: 코어 인코더 구성요소 |
| 306: 다중화기 | 308: 트랜스코더 |
| 310: 역다중화기 | 312: 디코더 |
| 314: 이득 구성요소 | 316: 디지털-대-아날로그 변환기 |
| 320: 전체 범위 스피커 | 322: 소형 스피커 |
| 324: 헤드폰 | 401: 입력 오디오 |
| 402: 인코더 | 406: 디코더 |
| 408: 이득 산출 유닛 | 409: 다중화기 |
| 410: 증폭기 | 411: 역다중화기 |
| 416: DACS 유닛 | 501: 오디오 입력 |
| 502: 인코더 | 506: 프로파일 |
| 510: 컴퓨터 | 512: 이동 전화 |
| 514: AVR | 516: 텔레비전 |

도면

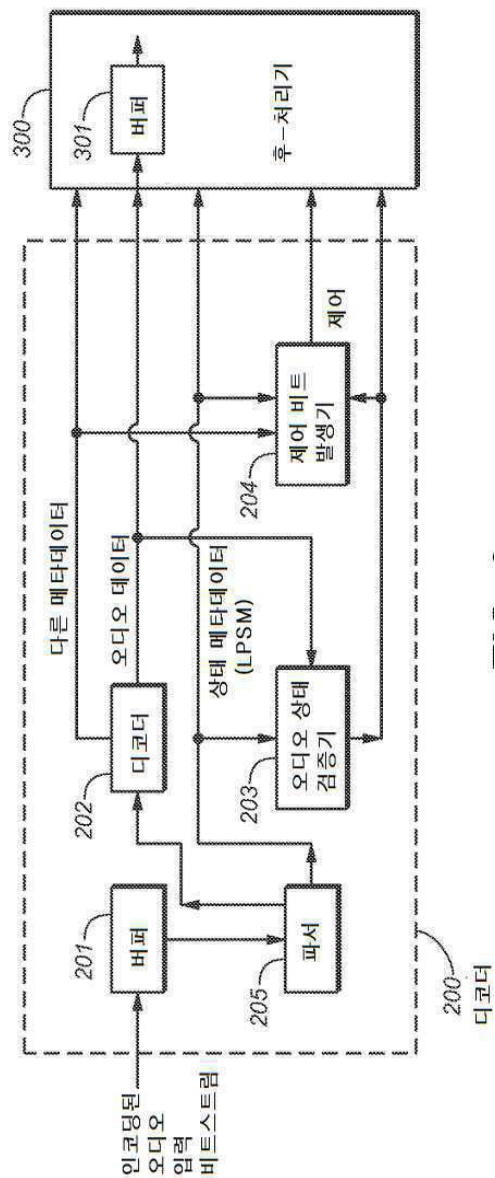
도면1



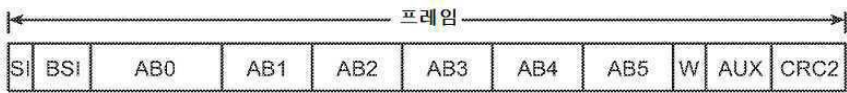
도면2



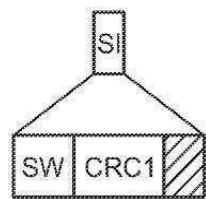
도면3



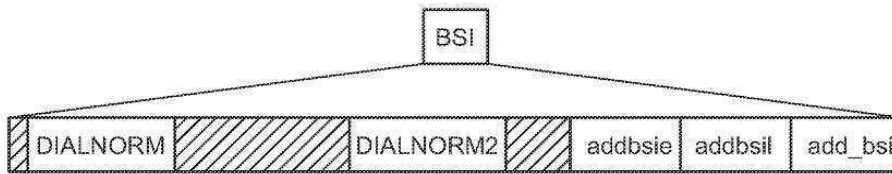
도면4



도면5



도면6



도면7

SI	BSI	AB0, AB1, AB2, AB3, AB4, AB5 중 1 내지 6	W	AUX	CRC2
----	-----	---------------------------------------	---	-----	------

도면8

파라미터	디스크립션	의무적/선택적
SYNC [ID]		M
코어 요소 버전		M
코어 요소 길이		M
코어 요소 기간 (xxx)		M
연장된 요소 카운트	코어 요소와 연관된 연장된 메타 데이터의 수를 표시함. 이 값은 비트스트림이 분배 및 최종 방출을 통한 생성으로부터 전달될 때 증가/감소할 수 있다	M
서브스트림 연관	코어 요소가 어떤 서브스트림(들)과 연관되는지를 설명함	M
서명 (HMAC 다이제스트)	전체 프레임의 오디오 데이터, 코어 요소, 및 모든 확장된 요소들에 걸쳐 계산된 256-비트 HMAC 다이제스트(SHA-2 알고리즘을 사용하여)	M
PGM 경계 카운트다운	스트림의 헤드 또는 테일에서 프레임들의 몇몇 수에 대해 나타난다. 따라서, 코어 요소 버전 변화는 이러한 파라미터의 포함을 시그널링하기 위해 사용될 수 있다	O
오디오 핑거프린트	코어 요소 기간 필드에 의해 표현된 PCM 오디오 샘플들의 몇몇 수에 걸쳐 취해진 오디오 핑거프린트	O
비디오 핑거프린트	코어 요소 기간 필드에 의해 표현된 압축된 비디오 샘플들(만약에 있다면)의 몇몇 수에 걸쳐 취해진 비디오 핑거프린트	O
URL/UUID	이 필드는 부가적인 프로그램 콘텐츠의 외부 위치(본질) 및/또는 비트스트림과 연관된 메타데이터를 참조하는 URL 및/또는 UUID(그것은 핑거프린트에 중복될 수 있다)를 운반하기 위해 정의된다	O

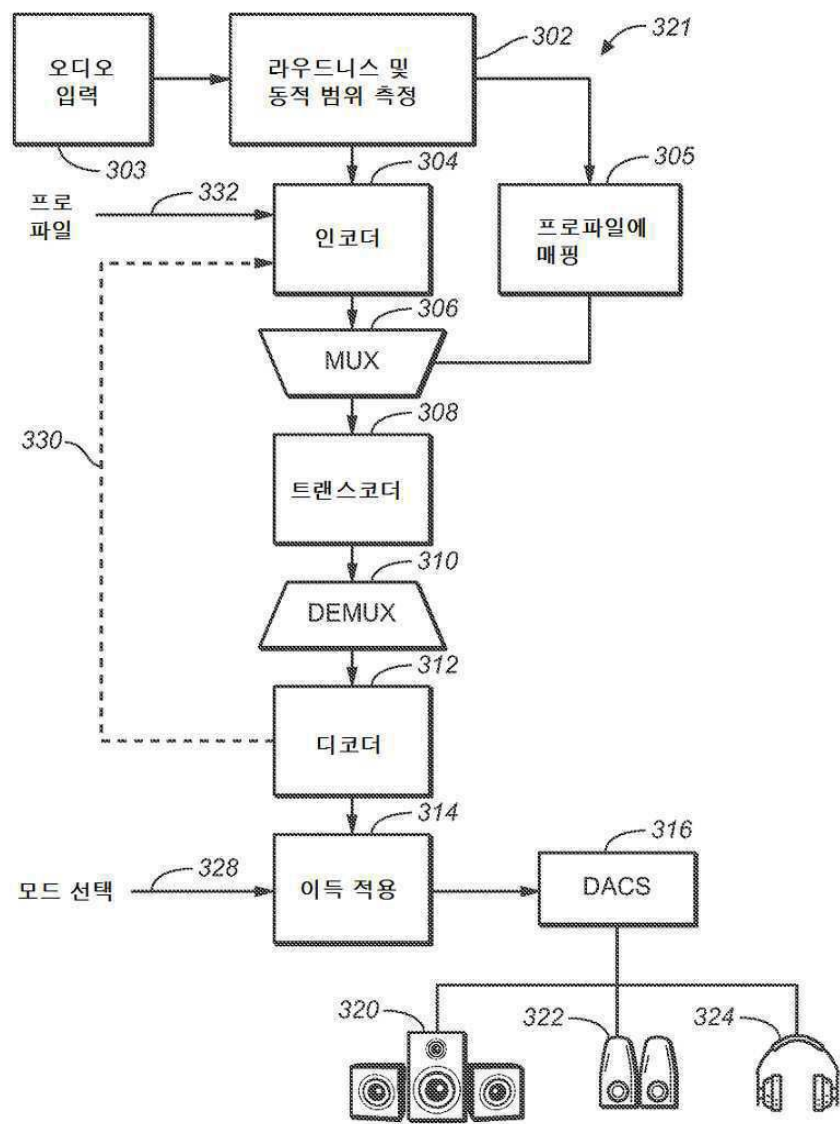
표 1

도면9

LPSM 파라미터 [기능적 라우드니스]	디스크립션	숫자 고유 상태들	의무적/ 선택적	삽입 레이트 (파라미터를 업데이트하는 기간)
LPSM 버전			M	
LPSM 기간 (xxx)	xxx 필드들 전용에 적용 가능함		M	
LPSM 카운트			M	
LPSM 서브스트림 연관			M	
다이얼로그 채널(들)	L, C 및 R 오디오 채널들의 어떤 조합이 이전 0.5초들에 걸쳐 스피치를 포함하는지를 표시함. 스피치가 임의의 L, C 또는 R 조합에 존재하지 않을 때, 이 파라미터는 "다이얼로그 없음"을 표시할 것임	8	M	~0.5 초들 (통상적)
라우드니스 규정 유형	연관된 오디오 데이터 스트림이 특정 세트의 규정들(예로서, ATSC A/85 또는 EBU R128)을 준수하는지를 표시한다	8	M	프레임
다이얼로그 게이팅 라우드니스 교정 플래그	연관된 오디오 스트림이 다이얼로그 게이팅에 기초하여 교정되었 는지를 표시한다	2	O (단지 라우드니스_ 규정_유형이 대응하 는 오디오가 교정되 지 않음을 표시하는 경우 존재함)	프레임
라우드니스 교정 유형	연관된 오디오 스트림이 무한 예전(파일-기반)을 갖고 또는 실 시간(RT) 라우드니스 및 동적 범위 제어를 갖고 교정되었는지 를 표시한다	2	O (단지 라우드니스_ 규정_유형이 대응하 는 오디오가 교정되 지 않음을 표시하는 경우에만 존재함)	프레임
ITU 상대적 게이팅 라우드니스 (INF)	적용된 메타데이터 없이 연관된 오디오 스트림의 ITU-R BS.1770-2 통합 라우드니스를 표시한다(예로서, 7 비트들: -58 → +5.5 LKFS 0.5 LKFS 스텝들)	128	O	1 초
ITU 스피치 게이팅 라우드니스 (INF)	적용된 메타데이터 없이 연관된 오디오 스트림의 ITU-R BS.1770-2 통합 라우드니스를 표시한다(예로서, 7 비트들: -58 → +5.5 LKFS 0.5 LKFS 스텝들)	128	O	1 초
ITU (EBU 3341) 단기 3s 라우드니스	적용된 메타데이터 없이 연관된 오디오 스트림의 3-초 게이팅되 지 않은 ITU 라우드니스를 표시한다(슬라이딩 윈도우) @ ~10Hz 삽입 레이트(예로서, 8 비트들: 116 → +11.5 LKFS 0.5 LKFS 스텝들)	256	O	0.1 초
실제 피크 값	적용된 메타데이터 없이 연관된 오디오 스트림의 ITU-R BS.1770 부록 2 실제피크 값(dB TP)을 표시한다(즉, 요소 기 간 필드에서 시그널링된 프레임 기간에 걸쳐 최대 값) 116 → +11.5 LKFS 0.5 LKFS 스텝들	256	O	0.5 초

표 2

도면10

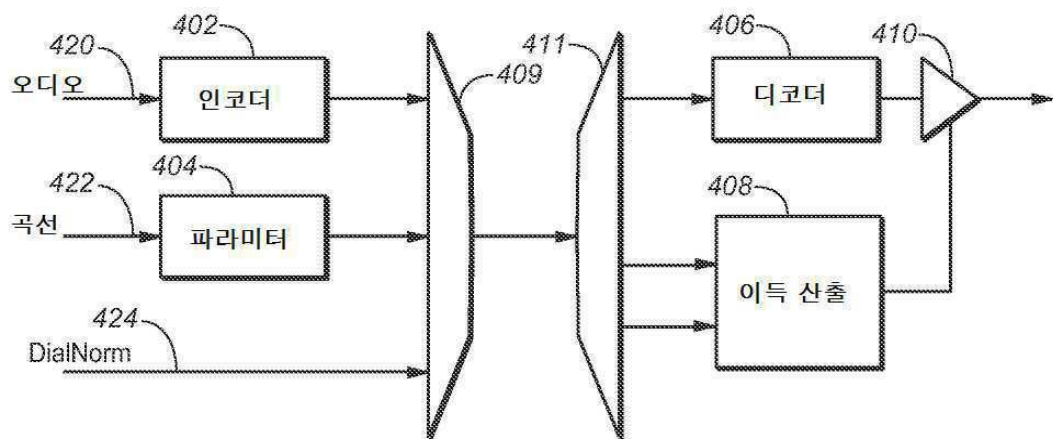


도면11

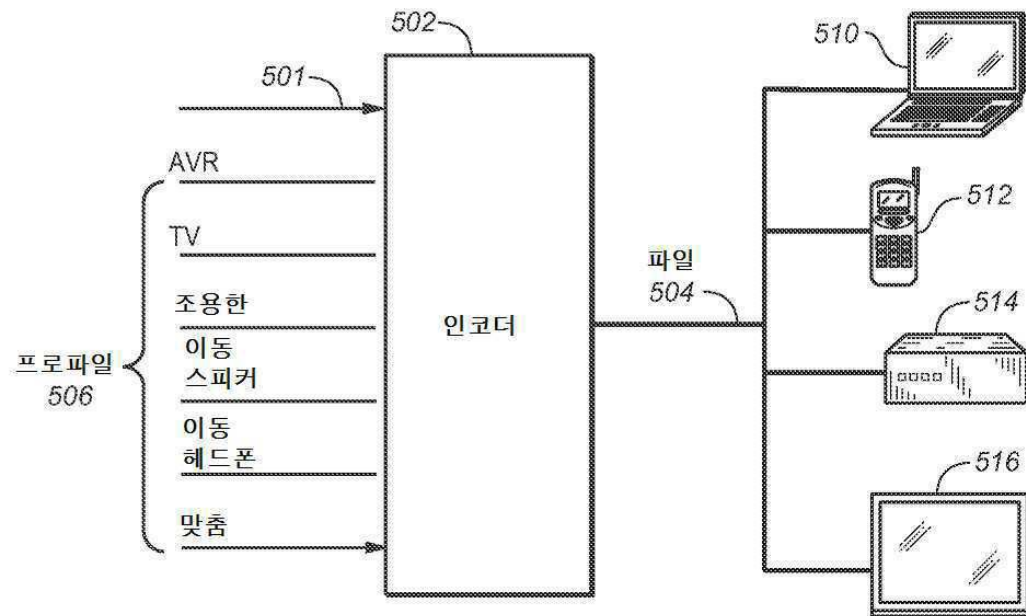
스피커들	재생 조건	
	조용한	시끄러운
소형/미니	낮은 동적 범위	낮은 동적 범위
헤드폰들	중간 동적 범위	낮은 동적 범위
대형	높은 동적 범위	낮은 동적 범위

표 3

도면12



도면13

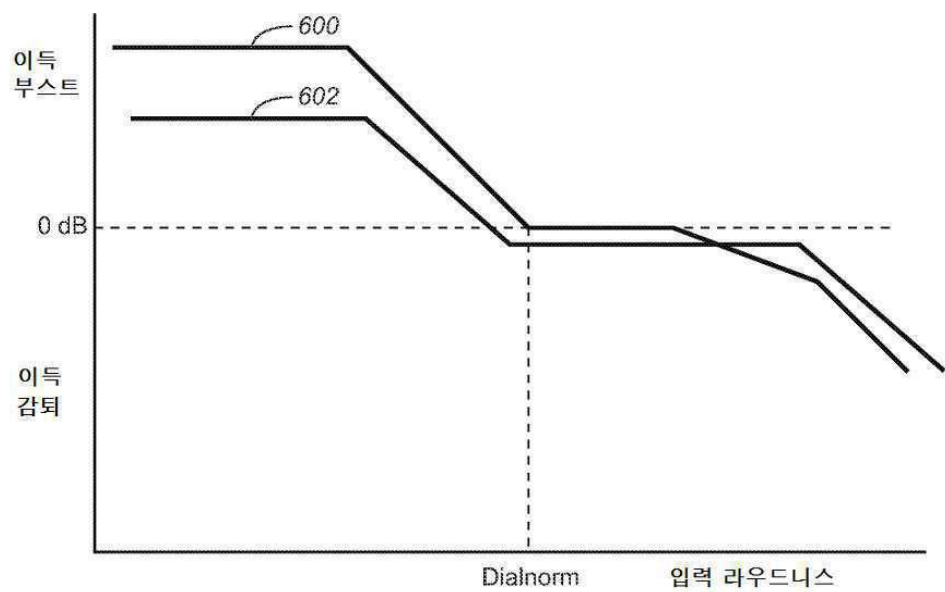


도면14

장기 라우드니스	DRC 이득들	오버로드 압축
1	프로파일 1	1
2	프로파일 2	2
3	프로파일 3	3
⋮	⋮	⋮
N	프로파일 N	N

표 4

도면15



도면16

