



(19) 대한민국특허청(KR)
 (12) 공개특허공보(A)

(11) 공개번호 10-2008-0049735
 (43) 공개일자 2008년06월04일

(51) Int. Cl.

H03M 7/30 (2006.01) *G11B 20/10* (2006.01)
G10L 19/00 (2006.01)

(21) 출원번호 10-2008-7005994

(22) 출원일자 2008년03월11일

심사청구일자 2008년03월11일

번역문제출일자 2008년03월11일

(86) 국제출원번호 PCT/KR2006/003434

국제출원일자 2006년08월30일

(87) 국제공개번호 WO 2007/027055

국제공개일자 2007년03월08일

(30) 우선권주장

1020060004055 2006년01월13일 대한민국(KR)

(뒷면에 계속)

(71) 출원인

엘지전자 주식회사

서울특별시 영등포구 여의도동 20번지

(72) 발명자

방희석

서울 서초구 양재동 14-10 101호

오현오

경기 고양시 일산서구 주엽1동 강선마을3단지아파트 306동 403호

(뒷면에 계속)

(74) 대리인

김용인, 박영복

전체 청구항 수 : 총 6 항

(54) 오디오 신호의 디코딩 방법 및 장치

(57) 요 약

오디오 신호 디코딩 방법 및 장치가 개시된다. 이 디코딩 방법은 공간 정보 신호와 다운믹스 신호를 포함하는 오디오 신호를 수신하는 단계, 오디오 신호에 포함된 타임슬롯의 수 및 파라미터의 수를 이용하여 타임슬롯의 위치 정보를 획득하는 단계, 타임슬롯의 위치 정보에 근거하여, 공간 정보 신호를 다운믹스 신호에 적용하여 다채널 오디오 신호를 생성하는 단계 및 출력채널에 대응하여 다채널 오디오 신호에 대한 다채널 배열을 수행하는 단계를 포함하여 오디오 신호를 보다 효율적으로 압축 및 전송할 수 있다.

(72) 발명자

김동수

서울 관악구 남현동 602-265 우림빌라 1502호

임재현

서울 관악구 남현동 1062-20 파크빌오피스텔 609호

정양원

서울 강남구 도곡동 역삼한신아파트 2동 803호

(30) 우선권주장

1020060004056 2006년01월13일 대한민국(KR)

1020060004065 2006년01월13일 대한민국(KR)

1020060056480 2006년06월22일 대한민국(KR)

60/712,119 2005년08월30일 미국(US)

60/719,202 2005년09월22일 미국(US)

60/723,007 2005년10월04일 미국(US)

60/726,228 2005년10월14일 미국(US)

60/729,225 2005년10월24일 미국(US)

60/735,628 2005년11월12일 미국(US)

60/748,607 2005년12월09일 미국(US)

60/762,536 2006년01월27일 미국(US)

60/803,825 2006년06월02일 미국(US)

특허청구의 범위

청구항 1

오디오 설명어를 포함하는 오디오 신호를 수신하는 단계;

상기 오디오 설명어를 이용하여, 상기 오디오 신호가 다운믹스 신호 및 공간 정보 신호를 포함하는 것을 인식하는 단계;

상기 공간 정보 신호를 이용하여, 상기 다운믹스 신호를 다채널 오디오 신호로 업믹싱하는 단계; 및

상기 다채널 오디오 신호를 미리 정해진 배열에 따라 출력채널에 매핑하는 단계를 포함하고

상기 공간 정보 신호는 주기적 또는 비주기적인 시간적 간격마다 헤더를 포함하는 것을 특징으로 하는 오디오 신호 디코딩 방법.

청구항 2

제1 항에 있어서, 상기 오디오 신호 디코딩 방법은

상기 공간 정보 신호에 상기 헤더가 포함되어 있는 경우, 상기 헤더에 포함된 환경 설정 정보를 이용하여 상기 다운믹스 신호를 상기 다채널 신호로 업믹싱하는 단계를 더 포함하는 것을 특징으로 하는 오디오 신호 디코딩 방법.

청구항 3

제2 항에 있어서, 상기 오디오 신호 디코딩 방법은

상기 공간 정보 신호에 포함되어 있는 헤더가 이전에 추출한 헤더와 다르면 헤더에 에러가 발생했음을 감지하는 단계를 더 포함하는 것을 특징으로 하는 오디오 신호 디코딩 방법.

청구항 4

제1 항에 있어서, 상기 오디오 신호 디코딩 방법은

상기 공간정보 신호에 상기 헤더가 포함되어 있지 않은 경우, 이전에 추출한 환경 설정 정보를 이용하여 상기 다운믹스 신호를 상기 다채널 신호로 업믹싱하는 단계를 더 포함하는 것을 특징으로 하는 오디오 신호 디코딩 방법.

청구항 5

제1 항, 제2 항 또는 제4 항 중 어느 한 항에 있어서, i번째 상기 오디오 신호를 출력채널에 매핑하는데 필요한 정보량은 전체 오디오 신호의 갯수와 상기 의 차에 1을 더한 값을, 2를 밑으로 취하는 로그 함수에 대입할 때 상기 로그 함수 값보다 크거나 같은 최소한의 정수인 것을 특징으로 하는 오디오 신호 디코딩 방법.

청구항 6

제1 항에 있어서, 상기 오디오 신호를 출력채널에 매핑하는 단계는 상기 공간정보 신호에 상기 헤더가 포함되어 있는 경우, 상기 헤더에 포함된 환경 설정 정보로부터 추출된 오디오 신호 매핑 정보를 이용하여 수행되는 것을 특징으로 하는 오디오 신호 디코딩 방법.

명세서

기술분야

<1>

본 발명은 오디오 신호의 처리에 관한 것으로서, 특히 오디오 신호 디코딩 방법 및 그 장치에 관한 것이다.

배경기술

<2>

일반적으로 오디오 신호의 경우, 인코딩 장치는 다채널 오디오 신호를 각각 압축하는 대신 오디오 신호를 모노 혹은 스테레오 형태의 다운믹스 신호로 압축하고, 압축된 다운믹스 신호를 공간 정보 신호(spatial information

signal)와 함께 디코딩 장치로 전송하거나 저장 매체에 저장한다. 여기서, 공간 정보 신호는 다채널 오디오 신호를 다운믹싱할 때 추출되는 것으로 다운믹스 신호로부터 원래의 다채널 오디오 신호를 복원할 때 사용된다.

<3> 일반적으로 환경 설정 정보는 불변이므로 이 정보를 포함하는 헤더는 오디오 신호에 초기에 한 번 삽입되어 전송되므로 임의의 순간부터 오디오 신호를 재생할 경우 오디오 신호 디코딩 장치는 환경 설정 정보가 없어 공간 정보를 디코딩하지 못하는 문제가 있다.

<4> 또한, 오디오 신호 인코딩 장치는 다운믹스 신호와 공간 정보 신호를 함께 또는 각각 비트스트림 형태로 생성하여 오디오 신호 디코딩 장치로 전송하므로 공간 정보 신호에 불필요한 정보 등이 포함될 경우 신호 압축 및 전송 효율이 떨어진다는 문제가 있다.

발명의 상세한 설명

<5> 기술적 과제

<6> 본 발명이 이루고자 하는 기술적 과제는, 공간 정보 신호에 헤더를 선택적으로 포함시켜 임의의 순간부터 오디오 신호를 재생할 수 있는 오디오 신호 디코딩 방법 및 그 장치를 제공하는 데 있다.

<7> 또한, 본 발명이 이루고자 하는 다른 기술적 과제는, 파라미터 세트가 적용되는 타임슬롯의 위치를 가변 비트 수를 사용하여 효율적으로 표현하는 오디오 신호 디코딩 방법 및 그 장치를 제공하는 데 있다.

<8> 또한, 본 발명이 이루고자 하는 또 다른 기술적 과제는 다운믹스 신호 배열을 수행하거나 다채널을 스피커와 매핑할 때 요구되는 정보량을 최소한의 가변 비트 수로 표현하여 오디오 신호 압축 및 전송 효율이 높아지는 오디오 신호 디코딩 방법 및 그 장치를 제공하는 데 있다.

<9> 또한, 본 발명이 이루고자 하는 또 다른 기술적 과제는 다운믹스 신호 배열을 수행하지 않고 다채널을 스피커에 매핑함으로써 신호 배열에 소요되는 정보량을 감소시키는 오디오 신호 디코딩 방법 및 그 장치를 제공하는 데 있다.

<10> 기술적 해결방법

<11> 상기의 목적을 구체적으로 실현할 수 있는 본 발명의 실시 예를 첨부한 도면을 참조하여 설명한다.

<12> 도 1은 본 발명의 일 실시 예에 따라 오디오 신호 인코딩 장치에서 오디오 신호 디코딩 장치로 전송되는 오디오 신호의 구성도이다. 도 1에 따른 오디오 신호는 오디오 설명어(101), 다운믹스 신호(103) 그리고 공간 정보 신호(105)를 포함한다.

<13> 오디오 신호를 재생하는 코딩 방법을 방송 등에 이용하는 경우 오디오 신호는 오디오 설명어(101), 다운믹스 신호(103) 외에 부가 정보(ancillary data)를 포함할 수 있다. 본 발명은 부가 정보로 공간 정보 신호(105)를 포함한다. 오디오 신호는 오디오 신호 디코딩 장치가 오디오 신호를 분석하지 않고도 오디오 코덱의 기본적인 정보를 알 수 있도록 오디오 설명어(또는 오디오 기술자)(audio descriptor)(101)를 선택적으로 포함할 수 있다. 오디오 설명어(101)는 전송되는 오디오 신호의 전송률, 채널 수, 압축 데이터의 샘플링 주파수, 사용하고 있는 오디오 코덱을 나타내는 식별자 등 오디오 디코딩에 필요한 기초적인 소수의 정보로 구성된다. 오디오 신호 디코딩 장치는 오디오 설명어(101)를 이용하여 오디오 신호가 사용하는 코덱의 종류가 무엇인지를 알 수 있다. 즉, 오디오 신호 디코딩 장치는 오디오 설명어(101)를 이용하여 오디오 신호가 공간 정보 신호(105)와 다운믹스 신호(103)를 이용하여 멀티 채널을 형성하는지 여부 등을 알 수 있다. 오디오 설명어(101)는 오디오 신호에 포함되어 있는 다운믹스 신호(103) 또는 공간 정보 신호(105)와는 독립적으로 위치한다. 예를 들면 오디오 설명어(101)는 오디오 신호를 표시하는 별도의 필드 안에 위치한다. 다운믹스 신호(103)에 헤더가 없는 경우 오디오 신호 디코딩 장치는 오디오 설명어(101)를 이용하여 다운믹스 신호(103)를 디코딩할 수 있다.

<14> 다운믹스 신호(103)는 멀티 채널을 다운 믹싱하여 생성되는 신호로 오디오 신호 인코딩 장치에 포함된 다운믹싱 부를 통해 생성되거나 또는 인위적으로 생성될 수도 있다. 다운믹스 신호(103)는 헤더를 포함하는 경우와 그렇지 않은 경우로 나뉜다. 다운믹스 신호(103)가 헤더를 포함하는 경우에는 프레임 단위로 매 프레임에 헤더가 포함되어 있다. 다운믹스 신호(103)가 헤더를 포함하지 않는 경우에는 앞에서 언급한 바와 같이 오디오 설명어(101)를 이용하여 다운믹스 신호(103)를 디코딩할 수 있다. 다운믹스 신호(103)는 프레임마다 헤더를 포함하거나 또는 프레임에 헤더를 포함하지 않거나 둘 중 하나의 형태로 콘텐츠가 끝날 때까지 동일하게 오디오 신호에 포함된다.

<15> 공간 정보 신호(105) 또한 헤더(107)와 공간 정보(111)를 포함하는 경우와 헤더(107)를 포함하지 않고 공간 정

보(111)만 포함하는 경우로 나뉜다. 공간 정보 신호(105)의 헤더(107)는 매 프레임에 동일하게 들어가야 하는 것은 아니라는 점에서 다운믹스 신호(103)의 헤더와는 구별된다. 공간 정보 신호(105)는 헤더(107)를 포함하는 프레임과 그렇지 않은 프레임을 함께 사용할 수 있다. 공간 정보 신호(105)의 헤더(107)에 포함되는 대부분의 정보는 공간 정보(111)를 해독하여 공간 정보(111)를 디코딩하는 정보인 환경 설정 정보(109)이다. 공간 정보(111)는 프레임으로 구성되고 각 프레임은 타임슬롯으로 구성된다. 타임슬롯은 공간 정보(111) 프레임을 시간 간격으로 나눌 때 각각의 시간 간격을 의미한다. 하나의 프레임에 포함되는 타임슬롯 개수는 환경 설정 정보(109)에 포함되어 있다.

- <16> 환경 설정 정보(109)에는 타임슬롯 개수 외에도 신호 배열 정보, 신호 변환부 개수, 채널 구성 정보, 스피커 매핑 정보 등이 포함되어 있다. 신호 배열 정보는 복호화된 다운믹스 신호(103)를 다채널로 복원하기 전에 업믹싱을 위해 오디오 신호를 배열할 것인가를 표시하는 식별자이다.
- <17> 신호 변환부는 다운믹스 신호(103)를 업믹싱하여 다채널을 생성할 때, 한 개의 다운믹스 신호(103)를 두 개의 신호로 또는 두 개의 다운믹스 신호(103)를 세 개의 신호로 변환하기 위해 사용되는 OTT(One-To-Two) BOX 또는 TTT(Two-To-Three) BOX 등을 의미한다. OTT BOX 또는 TTT BOX는 오디오 신호 디코딩 장치의 업믹싱부(미도시)에 포함되어 다채널을 복원할 때 사용되는 개념적인 BOX이다. 공간 정보 신호(105)에는 신호 변환부의 종류 및 개수 등의 정보가 포함되어 있다.
- <18> 채널 구성 정보는 오디오 신호 디코딩 장치에 포함된 업믹싱부의 구성을 표현하는 정보이다. 채널 구성 정보는 오디오 신호가 신호 변환부를 거치는지 여부를 표현하는 식별자로 구성되어 있다. 오디오 신호 디코딩 장치는 채널 구성 정보를 이용하여 업믹싱부로 입력되는 오디오 신호가 신호 변환부를 거치는지 여부 등을 알 수 있다. 오디오 신호 디코딩 장치는 신호 변환부에 대한 정보, 채널 구성 정보 등을 이용하여 다운믹스 신호(103)를 다채널 오디오 신호로 업믹싱한다. 오디오 신호 디코딩 장치는 공간 정보(111)에 포함된 상기 신호 변환부 정보, 채널 구성 정보 등을 이용하여 다운믹스 신호(103)를 업믹싱하여 다채널을 생성한다.
- <19> 스피커 매핑 정보는 업믹싱하여 생성된 다채널 오디오 신호를 스피커로 출력할 때 다채널 오디오 신호를 각각 어느 스피커에 매핑할 것인가를 표시하는 정보이다. 오디오 신호 디코딩 장치는 환경 설정 정보(109)에 포함된 스피커 매핑 정보를 이용하여 다채널 오디오 신호를 스피커로 출력한다.
- <20> 공간 정보(111)는 다운믹스 신호와 결합하여 다채널 오디오 신호를 생성할 때 공간감을 주기 위해 사용되는 정보이다. 공간 정보(111)에는 오디오 신호들 사이의 에너지 차이를 나타내는 CLD(Channel Level Differences), 오디오 신호들 간의 긴밀성 내지는 유사성을 나타내는 ICC(Interchannel Correlations), 다른 신호들을 이용하여 오디오 신호 값을 예상하는 계수를 나타내는 CPC(Channel Prediction Coefficients)등의 파라미터들이 포함되어 있다. 파라미터 세트는 이러한 파라미터들의 묶음이다.
- <21> 공간 정보(111)에는 파라미터 외에도 파라미터 세트가 적용되는 타임슬롯의 위치가 고정되어 있는지 여부를 표현하는 프레임 식별자, 하나의 프레임에 적용되는 파라미터 세트의 개수, 파라미터 세트가 적용되는 타임슬롯의 위치 정보 등이 포함되어 있다.
- <22> 도 2는 본 발명의 다른 실시 예에 따른 오디오 신호 디코딩 방법을 나타내는 순서도이다. 오디오 신호 디코딩 장치는 오디오 신호 인코딩 장치가 비트 스트림 형태로 전송한 공간 정보 신호(105)를 수신한다(단계 201). 공간 정보 신호(105)는 다운믹스 신호(103)와는 별도의 스트림 형태로 전송되거나 다운믹스 신호(103)의 보조 데이터 또는 부가 데이터에 포함되어 전송된다. 공간 정보 신호(105)가 다운믹스 신호(103)와 결합되어 전송되는 경우, 오디오 신호의 역다중화부(미도시)는 수신한 오디오 신호를 인코딩된 다운믹스 신호(103)와 인코딩된 공간 정보 신호(105)로 분리한다. 인코딩된 공간 정보 신호는 헤더(107)와 공간 정보(111)를 포함한다. 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 헤더(107)가 포함되어 있는지를 판단하고(단계 203), 공간 정보 신호(105)에 헤더(107)가 포함되어 있는 경우 헤더(107)로부터 환경 설정 정보(109)를 추출한다(단계 205). 오디오 신호 디코딩 장치는 환경 설정 정보(109)가 공간 정보 신호(105)에 포함된 첫 번째 헤더(107)로부터 추출된 환경 설정 정보(109)인지를 판단한다(단계 207). 환경 설정 정보(109)가 공간 정보 신호(105)에서 첫 번째로 추출한 헤더(107)로부터 추출된 경우, 환경 설정 정보(109)를 디코딩하고(단계 215) 디코딩된 환경 설정 정보(109)에 따라 환경 설정 정보(109) 뒤로 전송되는 공간 정보(111)를 디코딩한다.
- <23> 오디오 신호에서 추출된 헤더(107)가 공간 정보 신호(105)에서 첫 번째로 추출된 헤더(107)가 아닌 경우, 헤더(107)에서 추출된 환경 설정 정보(109)가 첫 번째 헤더(107)로부터 추출된 환경 설정 정보(109)와 동일한지를 판단한다(단계 209). 환경 설정 정보(109)가 첫 번째 헤더(107)로부터 추출된 환경 설정 정보(109)와 동일한 경

우에는 첫 번째 헤더(107)로부터 추출하여 디코딩한 환경 설정 정보(109)를 이용하여 공간 정보(111)를 디코딩 한다. 추출한 환경 설정 정보(109)가 첫 번째 헤더(107)로부터 추출된 환경 설정 정보(109)와 동일하지 않은 경우에는 오디오 신호 인코딩 장치로부터 오디오 신호 디코딩 장치로 전송되는 경로 상에서 오디오 신호에 에러가 발생한 것인지를 판단한다(단계 211). 환경 설정 정보(109)가 가변인 경우에는 환경 설정 정보(109)가 첫 번째 헤더(107)로부터 추출된 환경 설정 정보(109)와 동일하지 않아도 에러가 발생한 것이 아니므로 헤더(107)를 가변인 헤더(107)로 간주하고(단계 213), 간주한 헤더(107)에서 추출된 환경 설정 정보(109)를 디코딩한다(단계 215). 오디오 신호 디코딩 장치는 디코딩한 환경 설정 정보(109)에 따라 환경 설정 정보(109) 뒤로 전송되는 공간 정보(111)를 디코딩한다. 환경 설정 정보(109)가 가변이 아님에도 첫 번째 헤더(107)로부터 추출된 환경 설정 정보(109)와 동일하지 않은 경우 오디오 신호 전송 경로 상에서 에러가 발생한 것이므로 에러가 발생한 환경 설정 정보(109)를 포함하는 공간 정보 신호(105)에 포함된 공간 정보(111)를 제거하거나 또는 공간 정보(111)의 에러를 정정한다(단계 217).

<24> 도 3은 본 발명의 또 다른 실시 예에 따른 오디오 신호 디코딩 방법을 나타내는 순서도이다. 오디오 신호 디코딩 장치는 오디오 신호 인코딩 장치로부터 다운믹스 신호(103)와 공간 정보 신호(105)를 포함하는 오디오 신호를 수신한다(단계 301). 오디오 신호 디코딩 장치는 수신한 오디오 신호를 공간 정보 신호(105)와 다운믹스 신호(103)로 분리하고(단계 303), 분리된 공간 정보 신호(105)와 다운믹스 신호(103)를 각각 코어 디코딩부(미도시)와 공간 정보 디코딩부(미도시)로 보낸다.

<25> 오디오 신호 디코딩 장치는 공간 정보 신호(105)로부터 타임슬롯 개수와 파라미터 세트 개수를 추출한다. 오디오 신호 디코딩 장치는 추출한 타임슬롯 개수와 파라미터 세트 개수를 이용하여 파라미터 세트가 적용되는 타임슬롯의 위치를 구한다. 몇 번째 파라미터 세트인가에 따라 파라미터 세트가 적용되는 타임슬롯의 위치는 가변적 비트 수로 표현된다. 파라미터 세트가 적용될 타임슬롯의 위치를 표시하는 비트 수를 감소시킴으로써 공간 정보 신호(105)를 효율적으로 표현할 수 있다. 파라미터 세트가 적용되는 타임슬롯의 위치에 대해서는 이후 살펴볼 도 4 내지 5에서 상세히 설명하기로 한다. 오디오 신호 디코딩 장치는 타임슬롯 위치가 구해지면, 그 위치에 파라미터 세트를 적용하여 공간 정보 신호(105)를 디코딩한다(단계 305). 또한, 오디오 신호 디코딩 장치는 다운믹스 신호(103)를 코어 디코딩부에서 디코딩한다(단계 305).

<26> 오디오 신호 디코딩 장치는 디코딩된 다운믹스 신호(103)를 그대로 업믹싱하여 다채널을 생성할 수도 있으나 디코딩된 다운믹스 신호(103) 순서를 배열한 후에 업믹싱 할 수도 있다(단계 307).

<27> 오디오 신호 디코딩 장치는 디코딩된 다운믹스 신호(103)와 디코딩된 공간 정보 신호(105)를 이용하여 다채널을 생성한다(단계 309). 오디오 신호 디코딩 장치는 다운믹스 신호(103)를 다채널로 생성하기 위해 공간 정보 신호(105)를 이용하는데, 공간 정보 신호(105)는 앞에서도 설명한 바와 같이 신호 변환부 개수, 다운믹스 신호(103)가 업믹싱될 때 신호 변환부를 거치는지 또는 신호 변환부를 거치지 않고 출력되는지 등을 표현하는 채널 구성 정보를 포함한다. 오디오 신호 디코딩 장치는 신호 변환부 개수, 채널 구성 정보 등을 이용하여 다운믹스 신호(103)를 업믹싱한다(단계 309). 채널 구성 정보를 표현하는 방법 및 보다 적은 비트 수를 이용하여 채널 구성 정보를 표현하는 방법에 대해서는 이후 살펴볼 도 6 내지 7을 이용하여 설명하기로 한다.

<28> 오디오 신호 디코딩 장치는 생성된 다채널 오디오 신호를 출력하기 위해 미리 정해진 순서대로 다채널 오디오 신호를 스피커에 매핑(mapping)한다(단계 311). 이때, 매핑하는 오디오 신호 순서가 증가함에 따라 다채널 오디오 신호를 스피커에 매핑하기 위한 비트 수는 감소한다. 즉, 다채널 오디오 신호에 순서대로 번호를 부여할 경우, 첫 번째 오디오 신호는 전체 스피커 중 어느 하나의 스피커에 매핑될 수 있으므로 오디오 신호를 스피커에 매핑하기 위해 요구되는 정보량이 두 번째 이후의 오디오 신호를 매핑하기 위해 요구되는 정보량보다 크다. 두 번째 이후의 오디오 신호는 이전 오디오 신호와 매핑된 스피커를 제외한 나머지 스피커 중 하나의 스피커에 매핑되므로 매핑하기 위해 요구되는 정보량이 감소한다. 즉, 매핑하는 오디오 신호 순서가 증가함에 따라 오디오 신호를 매핑하기 위해 요구되는 정보량을 표현하는 비트 수를 감소시킴으로써 공간 정보 신호(105)를 효율적으로 표현할 수 있다. 이 같은 방법은 단계 307에서 다운믹스 신호(103)를 배열할 경우에도 사용될 수 있다.

<29> 도 4는 본 발명의 일실시 예에 따른 파라미터 세트가 적용되는 타임슬롯의 위치 정보를 나타내는 선택스이다. 도 4를 참조하면, 도 4의 선택스는 'FramingInfo'(401)에 대한 것으로 이는 파라미터 세트 수 및 파라미터 세트가 적용되는 타임슬롯에 대한 정보를 표현한다. 'bsFramingType' 필드(403)는 공간 정보 신호(105)에 포함된 프레임이 고정 프레임(fixed frame)인지, 또는 가변 프레임(variable frame)인지를 나타낸다. 고정 프레임이란 파라미터 세트가 적용되는 타임슬롯 위치가 미리 정해져 있는 프레임을 의미한다. 즉, 미리 정해진 규칙에 따라 파라미터 세트가 적용될 타임슬롯의 위치가 결정되어 있다. 가변 프레임은 파라미터 세트를 적용할 타임슬롯의

위치가 미리 정해져 있지 않은 프레임을 의미한다. 따라서 가변 프레임은 파라미터 세트가 적용될 타임슬롯의 위치를 표현하는 타임슬롯 위치 정보가 더 필요하다. 이하 'bsFramingType'(403)을 프레임이 고정 프레임인지, 가변 프레임인지를 표시하는 '프레임 식별자'라 부르기로 한다.

<30> 가변 프레임인 경우, 'bsParamSlot' 필드(407, 411)는 파라미터 세트가 적용될 타임슬롯의 위치 정보를 나타낸다. 'bsParamSlot[0]'(407)는 첫 번째 파라미터 세트가 적용될 타임슬롯의 위치를 표현하고 'bsParamSlot[ps]'(411)는 두 번째 이후의 파라미터 세트가 적용될 타임슬롯의 위치를 표현한다. 첫 번째 파라미터 세트가 적용될 타임슬롯의 위치는 초기값으로 표현되고, 두 번째 이후의 파라미터 세트가 적용될 타임슬롯의 위치는 차이값 ('bsDiffParamSlot[ps]')(409), 즉 'bsParamSlot[ps]'과 'bsParamSlot[ps-1]'의 차로 표현된다. 여기서 ps는 파라미터 세트를 의미한다. 첫 번째 파라미터 세트는 ps=0으로 표현된다. ps는 0부터 전체 파라미터 세트 수보다 작은 값까지 표현된다.

<31> (□) 파라미터 세트가 적용되는 타임슬롯의 위치(407, 409)는 ps 값이 커짐에 따라 증가하고($bsParamSlot[ps] > bsParamSlot[ps-1]$), (□) 첫 번째 파라미터 세트가 적용되는 타임슬롯 위치의 최대값은 타임슬롯 개수와 파라미터 세트 수의 차에 1을 더한 값이며 타임슬롯 위치는 ' $nBitsParamSlot(0)$ '(413)의 정보량으로 표현된다. (□) 두 번째 이후의 파라미터 세트에 대하여, N번째 파라미터 세트가 적용되는 타임슬롯 위치는 N-1번째 파라미터 세트가 적용되는 타임슬롯 위치보다 1 이상 크고, 타임슬롯 개수에서 파라미터 세트 수를 뺀 값에 N 값을 더한 값까지 가질 수 있다. 두 번째 이후의 파라미터 세트가 적용될 타임슬롯의 위치($bsParamSlot[ps]$)는 차이값 (' $bsDiffParamSlot[ps]$ ')(409)으로 표현되며 이 값은 ' $nBitsParamSlot(ps)$ '(409)의 정보량으로 표현된다. 위 (□) 내지 (□)을 이용하여 파라미터 세트가 적용되는 타임슬롯 위치를 구할 수 있다.

<32> 예를 들어 하나의 공간 프레임에 포함되는 타임슬롯이 10개이고 파라미터 세트가 3개인 경우, 첫 번째 파라미터 세트(ps=0)가 적용되는 타임슬롯의 위치는 전체 타임슬롯 개수에서 전체 파라미터 세트 수를 뺀 값에 1을 더한 타임슬롯 위치까지 적용될 수 있다. 즉, 1부터 최대 8 중 어느 한 타임슬롯에 적용될 수 있다. 이는 파라미터 세트가 적용되는 타임슬롯의 위치가 파라미터 세트 번호에 따라 증가함을 고려하면, 나머지 두 개의 파라미터 세트가 적용될 수 있는 타임슬롯의 위치는 최대한 각각 9, 10 이 됨을 생각하면 이해할 수 있다. 따라서 첫 번째 파라미터 세트가 적용될 타임슬롯의 위치(407)는 1 내지 8을 표시하기 위해 3비트가 소요된다. 이를 수식으로 표현하면 $\text{ceil}(\log_2(k-i+1))$ 이 될 수 있다. 여기서, k는 타임슬롯의 수, i는 파라미터의 수를 의미한다.

<33> 만일 첫 번째 파라미터 세트가 적용되는 타임슬롯의 위치(407)가 5라면, 두 번째 파라미터 세트가 적용될 타임슬롯 위치 ($bsParamSlot[1]$)는 상기 (□)에 의해 '5+1=6' 내지 '10-3+2=9' 사이의 값 중에서 선택되어야 한다. 즉, 두 번째 파라미터 세트가 적용될 타임슬롯 위치는 첫 번째 파라미터 세트가 적용될 타임슬롯 위치에 1을 더한 값에 차이값($bsDiffParamSlot[ps]$)(409)을 더한 값으로 표현할 수 있다. 차이값(409)은 따라서 0부터 3이 될 수 있고 이는 2비트로 표현할 수 있다. 두 번째 이후의 파라미터 세트에 대해서는 파라미터 세트가 적용될 타임슬롯의 위치를 직접 표시하지 않고 차이값(409)으로 표현함으로써 비트 수를 감소시킬 수 있다. 앞 예에서 보면, 타임슬롯 위치를 직접 표시하면 6 내지 9 중 어느 하나를 표시하기 위해 4비트가 소요되나 차이값으로 표시하면 2비트만 소요될 수 있다.

<34> 따라서 파라미터 세트가 적용될 타임슬롯의 위치 정보 표시량 ($nBitsParamSlot(0)$ (413), $nBitsParamSlot(ps)$ (415)은 고정된 비트가 아닌 가변적인 비트 수로 표현될 수 있다.

<35> 도 5는 본 발명의 다른 실시 예에 따른 파라미터 세트를 타임슬롯에 적용하여 공간 정보 신호를 디코딩하는 방법을 나타내는 순서도이다. 도 5를 참조하면, 오디오 신호 디코딩 장치는 다운믹스 신호(103)와 공간 정보 신호(105)를 포함하는 오디오 신호를 수신한다(단계 501). 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 헤더(107)가 있는 경우 헤더(107)에 포함된 환경 설정 정보(109)로부터 프레임에 포함되는 타임슬롯의 개수를 추출한다(단계 503). 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 헤더(107)가 포함되어 있지 않은 경우에는 이전에 추출한 헤더(107)에 포함된 환경 설정 정보(109)로부터 타임슬롯의 개수를 추출한다. 오디오 신호 디코딩 장치는 공간 정보 신호(105)로부터 프레임에 적용될 파라미터 세트의 개수를 추출한다(단계 505). 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 포함되어 있는 프레임 식별자를 이용하여 프레임에 파라미터 세트들이 적용되는 타임슬롯의 위치가 고정되어 있는지 가변인지를 판단한다(단계 507). 프레임이 고정 프레임인 경우 오디오 신호 디코딩 장치는 미리 정해진 규칙에 따라 파라미터 세트를 타임슬롯에 적용하여 공간 정보 신호(105)를 디코딩한다(단계 513). 프레임이 가변 프레임인 경우 오디오 신호 디코딩 장치는 첫 번째 파라미터 세트가 적용되는 타임슬롯 위치 정보를 추출한다(단계 509). 앞에서 설명한 바와 같이, 첫 번째 파라미터 세트가 적용

되는 타임슬롯 위치는 타임슬롯 개수와 파라미터 세트 수의 차에 1을 더한 값까지 최대 적용될 수 있다. 오디오 신호 디코딩 장치는 첫 번째 파라미터 세트가 적용되는 타임슬롯 위치 정보를 이용하여 두 번째 이후의 파라미터 세트가 적용되는 타임슬롯의 위치 정보를 구한다(단계 511). N이 2보다 크거나 같은 자연수라고 할 때, N번째 파라미터 세트가 적용되는 타임슬롯 위치는 N-1번째 파라미터 세트가 적용되는 타임슬롯 위치보다 1 이상 크고, 타임슬롯 개수에서 파라미터 세트 수를 뺀 값에 N 값을 더한 값까지 가질 수 있다는 것을 이용하여 파라미터 세트가 적용되는 타임슬롯의 위치를 최소 비트 수로 표현할 수 있다. 오디오 신호 디코딩 장치는 구해진 타임슬롯의 위치에 파라미터 세트를 적용하여 공간 정보 신호를 디코딩한다(단계 513).

<36> 도 6 및 7은 본 발명의 일 실시 예에 따른 오디오 신호 디코딩 장치의 업믹싱부를 나타낸 것이다. 오디오 신호 디코딩 장치는 오디오 신호 인코딩 장치로부터 수신한 오디오 신호를 다운믹스 신호(103)와 공간 정보 신호(105)로 분리하고 다운믹스 신호(103)와 공간 정보 신호(105)를 각각 디코딩한다. 앞에서 설명한 바와 같이 오디오 신호 디코딩 장치는 타임슬롯에 파라미터를 적용하여 공간 정보 신호(105)를 디코딩한다. 오디오 신호 디코딩 장치는 디코딩된 다운믹스 신호(103)와 공간 정보 신호(105)를 이용하여 다채널 오디오 신호를 생성한다.

<37> 오디오 신호 인코딩 장치가 N개의 입력 채널을 M개의 오디오 신호로 압축하여 비트스트림 형태로 오디오 신호 디코딩 장치로 전송하면, 오디오 신호 디코딩 장치는 원래의 N개의 채널을 복원하여 출력하는데 이러한 구성을 N-M-N 구조라고 한다. 경우에 따라, 오디오 신호 디코딩 장치가 N개의 채널을 복원할 수 없는 경우, 공간 정보 신호(105)를 고려하지 않고 다운믹스 신호(103)만을 두 개의 스테레오 신호로 출력하는 경우도 있으나 여기서는 논외로 한다. N, M의 값이 고정된 값으로 정해진 구조를 고정 채널 구조라고 하고 고정되지 않은 임의의 값으로 표현되는 경우를 임의 채널 구조라고 부르기로 한다. 5-1-5, 5-2-5, 7-2-7 등의 고정 채널 구조인 경우, 오디오 신호 인코딩 장치는 오디오 신호에 채널 구조를 포함시켜 전송하고 오디오 신호 디코딩 장치는 이를 읽어 오디오 신호를 디코딩한다.

<38> 오디오 신호 디코딩 장치는 M개의 오디오 신호를 N개의 다채널로 복원하기 위해 신호 변환부를 포함하는 업믹싱부를 이용한다. 신호 변환부는 다운믹스 신호(103)를 업믹싱하여 다채널을 생성할 때, 한 개의 다운믹스 신호(103)를 두 개의 신호로 또는 두 개의 다운믹스 신호를 세 개의 신호로 변환하기 위해 사용되는 개념적인 BOX이다.

<39> 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 포함된 환경 설정 정보(109)로부터 채널 구성 정보를 추출하여 업믹싱부의 구조를 파악할 수 있다. 앞에서 설명한 바와 같이 채널 구성 정보는 오디오 신호 디코딩 장치에 포함된 업믹싱부의 구성을 표현하는 정보이다. 채널 구성 정보는 오디오 신호가 신호 변환부를 거치는지 여부를 표현하는 식별자로 구성되어 있다. 즉, 채널 구성 정보는 디코딩된 다운믹스 신호가 업믹싱부에서 신호 변환부를 거치는 경우에는 신호 변환부의 입, 출력 신호의 개수가 변하므로 분할 식별자로, 디코딩된 다운믹스 신호가 업믹싱부에서 신호 변환부를 거치지 않는 경우에는 신호 변환부의 입력 신호가 그대로 출력되므로 미분할 식별자로 표현될 수 있다. 본 발명에서는 분할 식별자를 '1'로 미분할 식별자를 '0'으로 표현하기로 한다.

<40> 채널 구성 정보를 표현하는 방법은 크게 수평 방법과 수직 방법이 있다. 수평 방법은 오디오 신호가 신호 변환부를 거치는 경우 즉, 채널 구성 정보가 1인 경우에는 신호 변환부를 거쳐 나온 하위 계층 신호가 다시 신호 변환부를 거치는지 여부를 분할 식별자 또는 미분할 식별자로 순차적으로 표시하고, 채널 구성 정보가 0인 경우에는 동일한 계층 또는 상위 계층의 다음 번째의 오디오 신호가 신호 변환부를 거치는지 여부를 분할 식별자 또는 미분할 식별자로 표시하는 방법이다. 수직 방법은 상위 계층의 오디오 신호가 신호 변환부를 거치는지 여부와 무관하게 상위 계층 오디오 신호 전체에 대해 각각의 오디오 신호가 신호 변환부를 거치는지 여부를 분할 식별자 또는 미분할 식별자로 순차적으로 표현한 후, 하위 계층의 오디오 신호에 대해 신호 변환부를 거치는지를 표현하는 방법이다.

<41> 이하 동일한 업믹싱부의 구조에 대해 도 6에서는 채널 구성 정보를 수평 방법으로, 도 7에서는 채널 구성 정보를 수직 방법으로 표현한 경우를 예로 들어 설명하기로 한다. 또한, 도 6과 도 7에서는 신호 변환부로 OTT 박스를 예로 들어 설명하기로 한다. 도 6을 참조하면, X₁부터 X₄까지 4개의 오디오 신호가 업믹싱부로 들어간다. X₁은 제1 신호 변환부로 들어가서 두 신호(601, 601)로 변환된다. 업믹싱부에 포함된 신호 변환부는 CLD, ICC 등의 공간 파라미터를 이용하여 오디오 신호를 변환한다. 제1 신호 변환부에서 변환된 신호(601, 603)는 각각 제2 신호 변환부와 제3 신호 변환부로 들어가서 Y₁부터 Y₄까지의 다채널 오디오 신호로 출력된다. X₂는 제4 신호 변환부로 들어가서 각각 Y₅, Y₆으로 출력된다. X₃, X₄는 신호 변환부를 거치지 않고 바로 출력된다.

<42> X₁ 이 제1 신호 변환부를 거치므로 채널 구성 정보는 분할 식별자 1로 표현된다. 도 6은 채널 구성 정보를 수평

방법으로 표현하고 있으므로, 채널 구성 정보가 분할 식별자로 표현되면 제1 신호 변환부를 거쳐 나온 두 신호(601, 603)가 신호 변환부를 거치는지 여부를 분할 식별자 또는 미분할 식별자로 순차적으로 표시한다. 제1 변환부의 두 출력 신호 중 위에 있는 신호(601)는 다시 제2 신호 변환부를 거치므로 분할 식별자 1로 표현된다. 제2 신호 변환부를 거친 신호는 신호 변환부를 거치지 않고 그대로 출력되므로 미분할 식별자 0으로 표현된다. 채널 구성 정보가 0인 경우 동일 계층 또는 상위 계층의 다음 번째의 오디오 신호에 대해 신호 변환부를 거치는지 여부를 분할 식별자 또는 미분할 식별자로 표시하므로 상위 계층의 X_2 신호에 대해 채널 구성 정보를 표현한다. X_2 는 제4 신호 변환부를 거치므로 분할 식별자 1로 표현되고 제4 신호 변환부를 거친 신호가 각각 Y_5 , Y_6 으로 그대로 출력되므로 미분할 식별자 0으로 표현된다. X_3 , X_4 는 신호 변환부를 거치지 않고 바로 출력되므로 미분할 식별자 0으로 표현된다. 따라서 수평 방법으로 채널 구성 정보를 표현하면 110010010000가 된다. 이해를 돋기 위해 여기서는 업믹싱부의 구성을 통해 채널 구성 정보를 추출했지만 오디오 신호 디코딩 장치는 역으로 채널 구성 정보를 읽고 업믹싱부의 구조를 파악한다.

<43> 도 7은 도 6과 마찬가지로 X_1 부터 X_4 까지 4개의 오디오 신호가 업믹싱부로 들어간다. 수직 방법은 채널 구성 정보를 상위 계층부터 하위 계층 순으로 분할 식별자 또는 미분할 식별자로 표현하므로 우선 최상위계층인 제1 계층(701)의 오디오 신호의 식별자를 순서대로 표시한다. 즉, X_1 , X_2 는 각각 제1, 제4 신호 변환부를 거치므로 채널 구성 정보는 각각 1이 되고 X_3 , X_4 는 신호 변환부를 거치지 않으므로 채널 구성 정보는 0이 된다. 따라서 제1 계층(701)의 채널 구성 정보는 1100이 된다. 이 같은 방법에 따라 제2 계층(703), 제3 계층(705)의 채널 구성 정보를 순서대로 표현하면 각각 1100, 0000이 된다. 따라서 수직 방법으로 표현된 전체 채널 구성 정보는 110011000000이 된다.

<44> 오디오 신호 디코딩 장치는 상기 채널 구성 정보를 읽고 업믹싱부를 구성한다. 오디오 신호 디코딩 장치가 업믹싱부를 구성하기 위해서는 채널 구성 정보가 수평 방법 또는 수직 방법 중 어느 방법으로 표현되었는지를 나타내는 식별자가 오디오 신호에 포함되어 있어야 한다. 또는 채널 구성 정보를 수평 방법으로 표현하는 것을 원칙으로 하되 수직 방법으로 표현하는 것이 효율적인 경우에는 오디오 신호 인코딩 장치는 채널 구성을 수직 방법으로 표현했음을 알려주는 식별자를 오디오 신호에 포함할 수도 있다.

<45> 오디오 신호 디코딩 장치는 수평 방법으로 표현된 패널 구성 정보를 읽고 업믹싱부를 구성할 수 있다. 그러나 수직 방법으로 표현된 채널 구성 정보인 경우, 오디오 신호 디코딩 장치는 업믹싱부에 포함되는 신호 변환부의 개수 또는 입출력 채널의 개수를 알아야 업믹싱부를 구성할 수 있다. 따라서 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 포함된 환경 설정 정보(109)로부터 신호 변환부의 개수 또는 입출력 채널 개수를 추출하여 업믹싱부를 구성할 수 있다.

<46> 오디오 신호 디코딩 장치는 채널 구성을 정보를 앞에서부터 순서대로 해독하는데 환경 설정 정보(109)로부터 추출한 신호 변환부의 개수만큼 채널 구성을 정보에 포함되어 있는 분할 식별자 1의 개수를 감지하면 더 이상의 채널 구성을 정보를 읽지 않아도 된다. 분할 식별자 1은 오디오 신호가 신호 변환부로 입력됨을 표시하므로 채널 구성을 정보에 포함된 분할 식별자 1의 개수는 업믹싱부에 포함된 신호 변환부의 개수와 같기 때문이다. 즉, 앞에서 예를 든 바와 같이 수직 방법으로 표현된 채널 구성을 정보가 110011000000인 경우, 채널 구성을 정보를 디코딩하기 위해 총 12비트를 읽어야 하나 오디오 신호 디코딩 장치가 신호 변환부의 개수가 4개임을 감지한 경우에는 채널 구성을 정보에 포함된 1의 개수가 네 번 나올 때까지 즉, 채널 구성을 정보 중 110011까지만 디코딩한다. 더 이상의 채널 구성을 정보를 이용하지 않아도 나머지 값들이 모두 미분할 식별자 0으로 표현되기 때문이다. 따라서 오디오 신호 디코딩 장치는 6비트만큼을 디코딩하지 않아도 되므로 디코딩 효율이 높아진다.

<47> 채널 구조가 이미 정해진 고정 채널 구조인 경우에는 신호 변환부의 개수 또는 입출력 채널 개수가 공간 정보 신호(105)에 포함된 환경 설정 정보에 포함되어 있으므로 별도의 정보가 필요 없으나 채널 구조가 정해지지 않은 임의 채널 구조인 경우에는 신호 변환부의 개수나 입출력 채널 개수가 공간 정보 신호(105)에 포함되어 있지 않으므로 신호 변환부의 개수나 입출력 채널 개수 등을 표현하기 위한 별도의 정보가 필요하다.

<48> 신호 변환부에 대한 정보를 예로 들어 설명하면, 신호 변환부로 OTT 박스만을 사용하는 경우, 신호 변환부를 표시하는 정보는 최대한 5비트로 표현될 수 있다. 업믹싱부로 들어가는 입력 신호는 OTT 박스 또는 TTT 박스를 거치는 경우 한 개의 입력 신호가 두 개로, 두 개의 입력 신호가 세 개로 변환되므로 출력 채널 수는 입력 신호에 OTT 박스 또는 TTT 박스 개수를 더한 값이 된다. 따라서 신호 변환부의 개수는 출력 채널 수에서 입력 신호 수와 TTT 박스 개수를 뺀 값이 된다. 일반적으로 출력 채널은 최대 32개까지 사용될 수 있으므로 신호 변환부를

표시하는 정보는 5비트 이내의 값으로 표현된다.

- <49> 따라서 채널 구성 정보가 수직 방법으로 표현되어 있고, 채널 구조 또한 임의 채널 구조인 경우에는 오디오 신호 인코딩 장치는 공간 정보 신호(105)에 신호 변환부의 개수를 최대 5비트로 별도로 표현해 주어야 한다. 상기 예에서는 6비트의 채널 구성 정보와 신호 변환부를 표시하는 정보 5비트가 소요되므로 총 11비트가 소요된다. 이는 수평 방법으로 표현된 채널 구성 정보보다 업믹싱부를 구성하기 위한 비트량이 감소함을 보여준다. 이와 같이 수직 방법으로 채널 구성 정보를 표현할 경우 비트 수가 감소하는 효과가 있다.
- <50> 도 8은 본 발명의 일 실시 예에 따른 오디오 신호 디코딩 장치를 나타내는 구성도이다. 도 8을 참조하면 오디오 신호 디코딩 장치는 수신부, 역다중화부, 코어 디코딩부, 공간 정보 디코딩부, 신호 배열부, 다채널 생성부, 스피커 매핑부를 포함한다. 수신부(801)는 오디오 신호 인코딩 장치(미도시)로부터 다운믹스 신호(103)와 공간 정보 신호(105)를 포함하는 오디오 신호를 수신한다. 역다중화부(803)는 수신부(801)가 수신한 오디오 신호를 인코딩된 다운믹스 신호(103)와 인코딩된 공간 정보 신호(105)로 파싱하여 각각 코어 디코딩부(805), 공간 정보 디코딩부(807)로 보낸다. 코어 디코딩부(805)와 공간 정보 디코딩부(807)는 인코딩된 다운믹스 신호와 인코딩된 공간 정보 신호를 각각 복호화한다. 공간 정보 디코딩부(807)는 앞에서 설명한 바와 같이 공간 정보 신호(105)로부터 프레임 식별자, 타임슬롯 개수, 파라미터 세트 개수, 타임슬롯 위치 정보 등을 추출하여 파라미터 세트를 타임슬롯에 적용하여 공간 정보 신호(105)를 디코딩한다.
- <51> 오디오 신호 디코딩 장치는 신호 배열부(809)를 포함할 수 있다. 신호 배열부(809)는 복호화된 다운믹스 신호(103)를 업믹싱하기 위해 복수의 다운믹스 신호(103)를 미리 정해진 배열에 따라 배열하는 곳이다. 즉, N-M-N 채널 구성에서 M 개의 다운믹스 신호를 M'의 오디오 신호로 배열한다. 오디오 신호 디코딩 장치는 코어 디코딩부(805)를 거친 순서 그대로 다운믹스 신호를 업믹싱할 수도 있지만 경우에 따라서는 다운믹스 신호의 순서를 배열하여 업믹싱을 수행하기도 한다. 상황에 따라서는 두 개의 다운믹스 신호를 세 개의 신호로 업믹싱하는 신호 변환부로 들어가는 신호에 대해서만 신호 배열을 수행할 수도 있다. 오디오 신호 인코딩 장치는 오디오 신호가 신호 배열을 수행하는 경우, 또는 TTT 박스의 입력 신호에 대해서만 신호 배열을 수행하는 경우에는 이를 표시하는 신호 배열 정보를 오디오 신호에 포함시켜야 한다. 신호 배열 정보는 오디오 신호를 다채널로 복원하기 전에 업믹싱을 위해 신호 순서를 배열할 것인가, 특정 신호에 대해서만 배열을 수행하는가 등을 표시하는 식별자이다. 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 헤더(107)가 포함된 경우, 헤더(107)로부터 추출한 환경 설정 정보(109)에 포함된 오디오 신호 배열 정보를 이용하여 다운믹스 신호를 배열한다. 오디오 신호 디코딩 장치는 공간 정보 신호(105)에 헤더(107)가 포함되어 있지 않은 경우에는 이전 헤더(107)에 포함되어 있는 환경 설정 정보(109)로부터 추출한 오디오 신호 배열 정보를 이용하여 오디오 신호를 배열할 수도 있다.
- <52> 오디오 신호 디코딩 장치는 다운믹스 신호 배열을 수행하지 않을 수도 있다. 즉, 오디오 신호 디코딩 장치는 다운믹스 신호 배열을 수행하지 않고 코어 디코딩부(805)가 복호화하여 다채널 생성부(811)로 전송한 신호를 그대로 업믹싱하여 다채널을 생성할 수도 있다. 생성된 다채널을 스피커에 매핑함으로써 신호 배열의 소기 목적은 달성되기 때문이다. 이 경우, 오디오 신호에 다운믹스 신호 배열에 대한 정보를 삽입하지 않음으로써 오디오 신호를 보다 효율적으로 압축 및 전송할 수 있다. 또한, 오디오 신호 디코딩 장치는 신호 배열을 별도로 수행하지 않음으로써 디코딩 장치의 복잡성(complexity)이 감소한다.
- <53> 신호 배열부(809)는 배열한 다운믹스 신호(103)를 다채널 생성부(811)로 보낸다. 공간 정보 디코딩부(809) 또한 복호화된 공간 정보 신호(105)를 다채널 생성부(811)로 보낸다. 다채널 생성부(811)는 다운믹스 신호(103)와 공간 정보 신호(105)를 이용하여 다채널 오디오 신호를 생성한다.
- <54> 오디오 신호 디코딩 장치는 다채널 생성부(811)를 거친 오디오 신호를 스피커로 출력하기 위해 스피커 매핑부(813)를 포함한다. 스피커 매핑부(813)는 다채널 오디오 신호를 각각 어느 스피커에 매핑하여 출력할 것인가를 결정한다. 일반적으로 오디오 신호를 출력하기 위해 사용되는 스피커 종류는 아래 표 1에 나타난 바와 같다.

<55>

표 1

bsOutputChannelPos	Loudspeaker
0	FL: Front Left
1	FR: Front Right
2	FC: Front Center
3	LFE: Low Frequency Enhancement
4	BL: Back Left
5	BR: Back Right
6	FLC: Front Left Center
7	FRC: Front Right Center
8	BC: Back Center
9	SL: Side Left
10	SR: Side Right
11	TC: Top Center
12	TFL: Top Front Left
13	TFC: Top Front Center
14	TFR: Top Front Right
15	TBL: Top Back Left
16	TBC: Top Back Center
17	TBR: Top Back Right
18...31	reserved

<56>

<57>

일반적으로 출력된 오디오 신호와 매핑할 수 있는 스피커는 최대 32개까지 가능하다. 그러므로 표 1에서와 같이 스피커 매핑부(813)는 다채널 오디오 신호에 0부터 31까지 중 특정한 번호(bsOutputChannelPos)를 부여하여 각각의 번호에 해당하는 스피커(Loudspeaker)에 오디오 신호가 매핑되도록 한다. 이때 다채널 생성부(811)에서 출력된 다채널 오디오 신호 중 첫 번째 오디오 신호를 스피커에 매핑하기 위해서는 전체 32개의 스피커 중 어느 하나의 스피커를 선택해야 하므로 5비트가 소요된다. 두 번째 오디오 신호를 스피커에 매핑하기 위해서는 남은 31개의 스피커 중 어느 하나의 스피커를 선택해야 하므로 역시 5비트가 소요된다. 이와 같은 방법에 따르면 열 일곱 번째 오디오 신호를 스피커에 매핑하기 위해서는 남은 16개의 스피커 중 하나의 스피커를 선택해야 하므로 4비트가 소요된다. 즉, 오디오 신호를 매핑하는 개수가 늘어남에 따라 오디오 신호와 매핑되는 스피커를 표시하기 위해 소요되는 정보량도 감소한다. 이를 수식으로 표현하면 오디오 신호를 스피커에 매핑하기 위해 소요되는 비트 수는 $\lceil \log_2(32-\text{bsOutputChannelPos}) \rceil$ 와 같이 표현된다. 이같이 배열할 오디오 신호의 개수가 늘어남에 따라 소요되는 비트 수가 감소하는 것은 신호 배열부(809)에서 배열하는 다운믹스 신호의 개수가 늘어나는 경우도 동일하다. 오디오 신호 디코딩 장치는 이 같은 방법으로 다채널 오디오 신호를 스피커에 매핑하여 출력한다.

<58>

지금까지 본 발명에 대하여 몇몇 실시 예들을 들어 구체적으로 설명하였으나, 상기 실시 예들은 본 발명을 이해하기 위한 설명을 위해 제시된 것이며, 본 발명의 범위가 상기 실시 예에 제한되는 것은 아니다. 당업자라면 본 발명의 기술적 사상의 범위를 벗어나지 않고도 다양한 변형이 가능함을 이해할 수 있을 것이며, 본 발명의 범위는 첨부된 특허청구범위에 의해서 해석되어야 할 것이다.

<59>

유리한 효과

<60>

이상에서 설명한 것과 같이, 본 발명에 의한 오디오 신호 디코딩 방법 및 장치는 공간 정보 신호에 헤더를 선택적으로 포함할 수 있다.

- <61> 또한, 본 발명에 의한 오디오 신호 디코딩 방법 및 장치는 파라미터 세트가 적용되는 타임슬롯의 위치를 가변적인 비트 수로 표현함으로써 전송되는 데이터 양을 줄일 수 있다.
- <62> 또한, 본 발명에 의한 오디오 신호 디코딩 방법 및 장치는 다운믹스 신호 배열을 수행하거나 다채널을 스피커와 매핑할 때 요구되는 정보량을 최소한의 가변 비트 수로 표현하여 오디오 신호 압축 및 전송 효율을 높이는 효과를 갖는다.
- <63> 또한, 본 발명에 의한 오디오 신호 디코딩 방법 및 장치는 다운믹스 신호 배열을 수행하지 않고 코어 디코딩부가 복호화하여 다채널 생성부로 전송한 신호를 순서대로 업믹싱함으로써 오디오 신호를 보다 효율적으로 압축 및 전송할 수 있고 오디오 신호 디코딩 장치의 복잡성이 감소하는 효과를 갖는다.

도면의 간단한 설명

- <64> 도 1은 본 발명의 일 실시 예에 따른 오디오 신호의 구성도이다.
- <65> 도 2는 본 발명의 다른 실시 예에 따른 오디오 신호 디코딩 방법을 나타내는 순서도이다.
- <66> 도 3은 본 발명의 또 다른 실시 예에 따른 오디오 신호 디코딩 방법을 나타내는 순서도이다.
- <67> 도 4는 본 발명의 일 실시 예에 따른 파라미터 세트가 적용되는 타임슬롯의 위치 정보를 나타내는 신택스이다.
- <68> 도 5는 본 발명의 다른 실시 예에 따른 파라미터 세트를 타임슬롯에 적용하여 공간 정보 신호를 디코딩하는 방법을 나타내는 순서도이다.
- <69> 도 6 및 7은 본 발명의 일 실시 예에 따른 오디오 신호 디코딩 장치의 업믹싱부를 나타낸 것이다.
- <70> 도 8은 본 발명의 일 실시 예에 따른 오디오 신호 디코딩 장치를 나타내는 구성도이다.

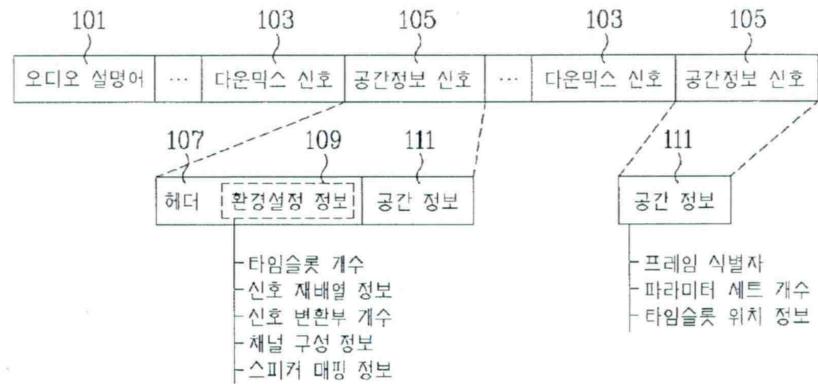
발명의 실시를 위한 조건의 형태

- <72> 상기 과정을 이루기 위한 본 발명에 의한 오디오 신호 디코딩 방법은 공간 정보 신호와 다운믹스 신호를 포함하는 오디오 신호를 수신하는 단계, 상기 오디오 신호에 포함된 타임슬롯의 수 및 파라미터의 수를 이용하여 타임슬롯의 위치 정보를 획득하는 단계, 상기 타임슬롯의 위치 정보에 근거하여, 상기 공간 정보 신호를 상기 다운믹스 신호에 적용하여 다채널 오디오 신호를 생성하는 단계 및 출력채널에 대응하여 상기 다채널 오디오 신호에 대한 다채널 배열을 수행하는 단계를 포함하는 것을 특징으로 하는 오디오 신호 디코딩 방법을 제공할 수 있다.
- <73> 바람직한 실시예에서, 상기 타임슬롯의 위치정보는 가변적 비트수로 표현되는 것을 특징으로 한다. 또한, 상기 위치 정보는 초기값 및 차이값을 포함하며, 상기 초기값은 첫 번째 파라미터가 적용되는 타임슬롯의 상기 위치 정보를 나타내고, 상기 차이값은 두 번째 이후의 파라미터가 적용되는 타임슬롯의 상기 위치 정보를 나타내는 것을 특징으로 한다. 또한, 상기 초기값은 상기 타임슬롯 수 및 상기 파라미터의 수 중 하나 이상을 이용하여 결정되는 가변 비트로 표현되는 것을 특징으로 한다. 또한, 상기 차이값은 상기 타임슬롯 수, 상기 파라미터의 수 및 이전 파라미터가 적용되는 타임슬롯의 위치 정보 중 하나 이상을 이용하여 결정되는 가변적 비트 수로 표현되는 것을 특징으로 한다. 또한, 상기 오디오 신호 디코딩 방법은 미리 정해진 방식으로, 상기 다운믹스 신호에 대한 다운믹스 신호 배열을 수행하는 단계를 더 포함하는 것을 특징으로 한다. 또한, 상기 다운믹스 신호 배열을 수행하는 단계는 두 개의 다운믹스 신호를 세 개의 신호로 업믹싱하는 신호 변환부로 들어가는 다운믹스 신호에 대해서만 수행되는 것을 특징으로 한다. 또한, 상기 다운믹스 신호 배열은 상기 공간 정보 신호에 헤더가 포함된 경우, 상기 헤더로부터 추출한 환경 설정 정보에 포함된 오디오 신호 배열 정보를 이용하여 상기 다운믹스 신호를 배열하는 것을 특징으로 한다. 또한, 제1 항, 제2 항 또는 제6 항 중 어느 한 항에 있어서, i번째 상기 오디오 신호를 매핑하는데 필요한 정보량 또는 i번째 상기 다운믹스 신호를 배열하는데 필요한 정보량은 전체 오디오 신호 또는 전체 다운믹스 신호의 갯수와 상기 i값의 차에 1을 더한 값을, 2를 밑으로 취하는 로그 함수에 대입할 때 상기 로그 함수 값보다 크거나 같은 최소한의 정수인 것을 특징으로 한다. 또한, 상기 다채널 배열 단계는 오디오 신호를 스피커에 대응하여 배열하는 단계를 더 포함하는 것을 특징으로 한다.
- <74> 본 발명의 다른 실시 예에 따르면, 오디오 신호를 다채널 오디오 신호로 업믹싱하는 업믹싱부 및 상기 다채널 오디오 신호를 미리 정해진 배열에 따라 출력채널에 매핑하는 다채널 배열부를 포함하는 것을 특징으로 하는 오디오 신호 디코딩 장치를 제공할 수 있다.
- <75> 본 발명의 또 다른 실시 예에 따르면, 인코딩된 다운믹스 신호를 복호화하는 코어 디코딩부, 상기 복호화된 오디오 신호를 미리 정해진 배열에 따라 배열하는 배열부 및 상기 배열된 오디오 신호를 다채널 오디오 신호로 업

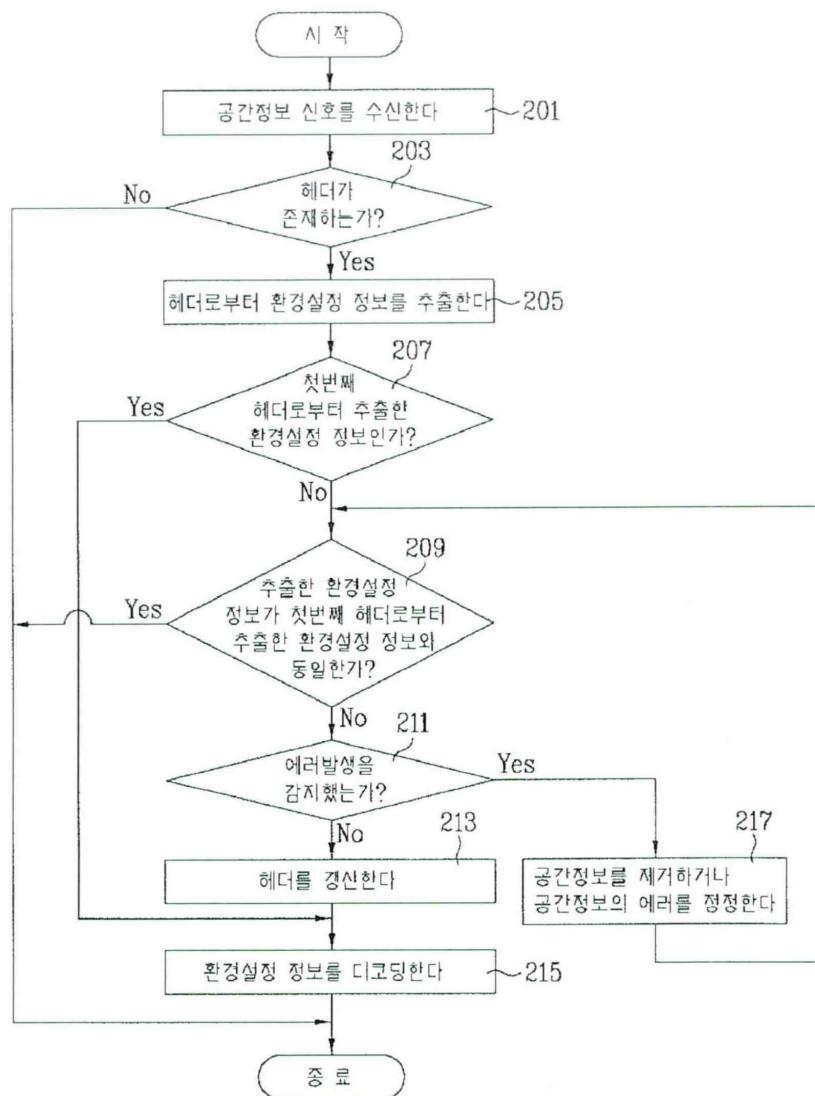
믹싱하는 업믹싱부를 포함하는 것을 특징으로 하는 오디오 신호 디코딩 장치를 제공할 수 있다.

도면

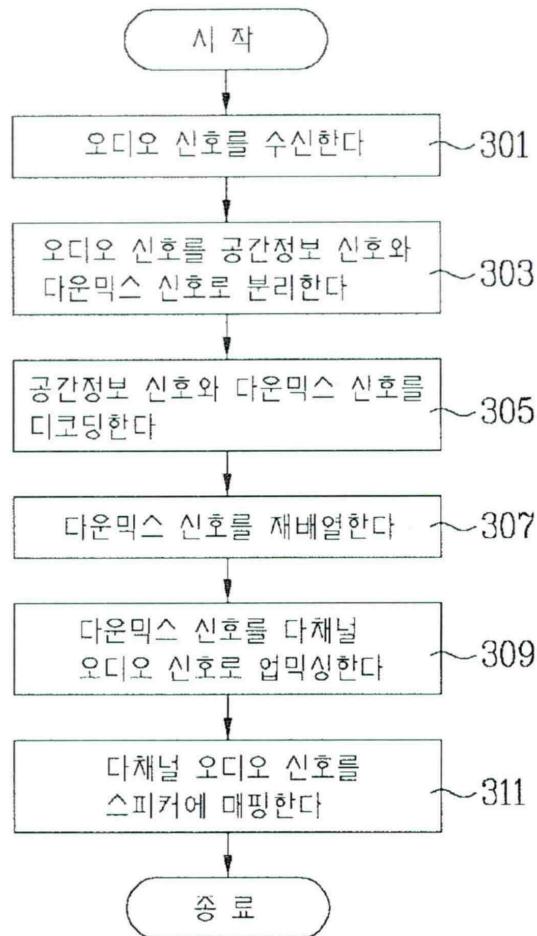
도면1



도면2



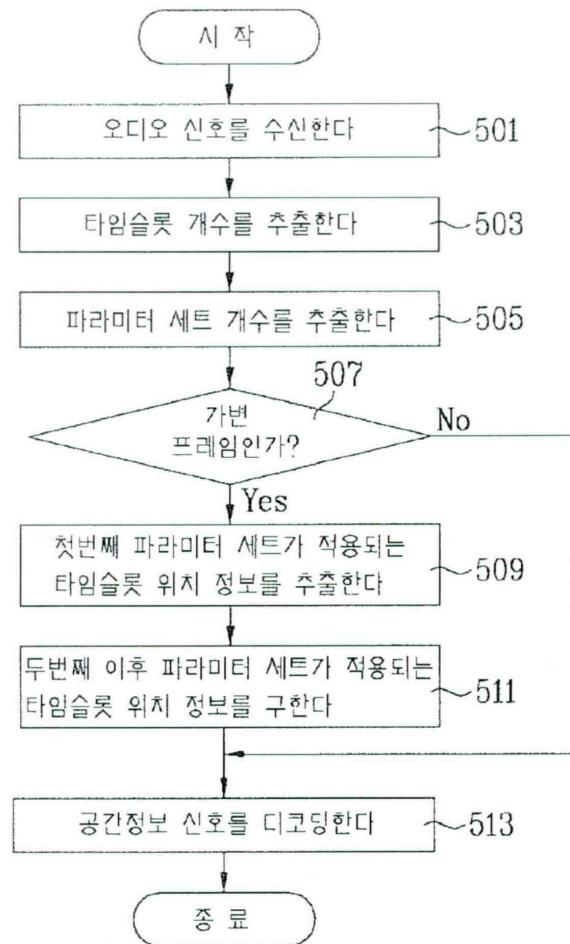
도면3



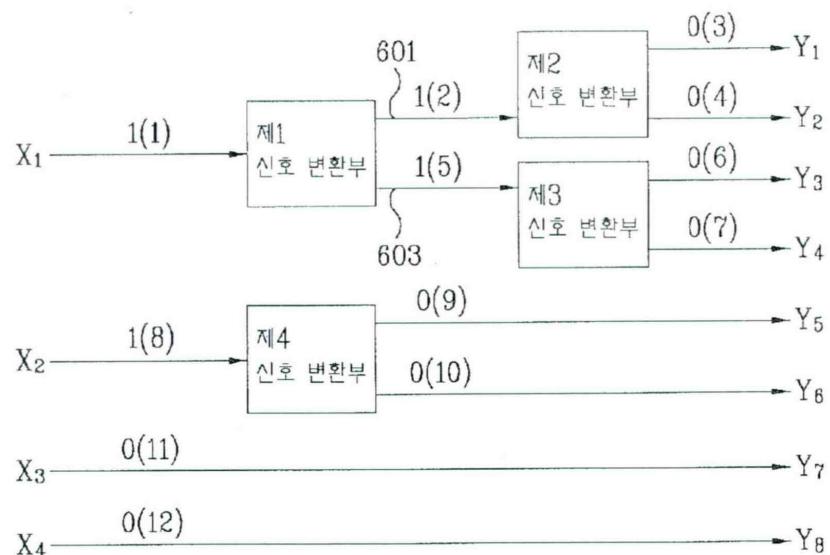
도면4

	Syntax	No. of bits
401	FramingInfo()	
	{	
403	bsFramingType;	
405	bsNumParamSets:	1
	if (bsFramingType) {	3
	for (ps=0; ps<numParamSets; ps++) {	
	if(ps==0){	
407	bsParamSlot[0];	nBitsParamSlot(0) ~ 413
	else {	
409	bsDiffParamSlot[ps];	nBitsParamSlot(ps) ~ 415
411	bsParamSlot[ps] = bsParamSlot[ps-1] + bsDiffParamSlot[ps] - 1;	
	}	
	}	
	}	
	}	

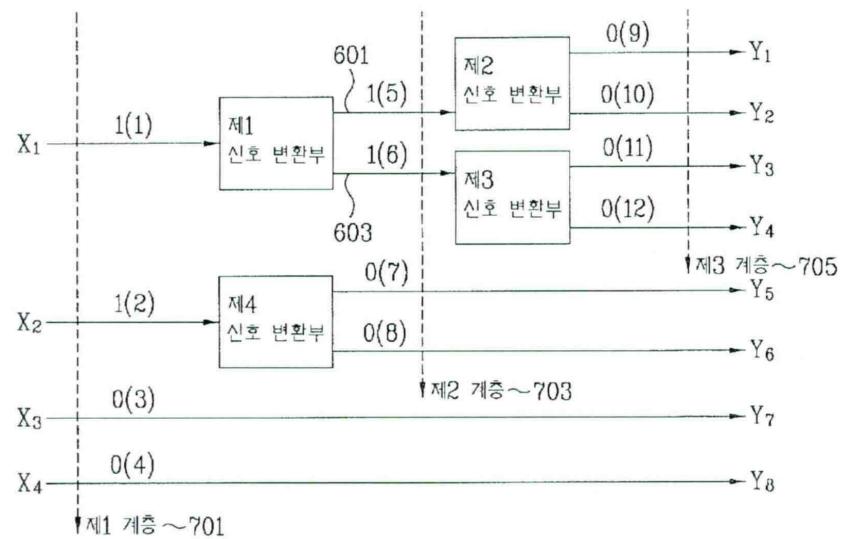
도면5



도면6



도면7



도면8

