(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0032746
(43) 공개일자 2024년03월12일

(51) 국제특허분류(Int. Cl.)
G10L 19/008 (2014.01) *G10L 19/035* (2013.01)
(52) CPC특허분류
G10L 19/008 (2020.08)
G10L 19/035 (2013.01)
(21) 출원번호 10-2023-7044255
(22) 출원일자(국제) 2022년07월08일
심사청구일자 없음
(85) 번역문제출일자 2023년12월21일
(86) 국제출원번호 PCT/JP2022/027053
(87) 국제공개번호 WO 2023/286698
국제공개일자 2023년01월19일
(30) 우선권주장
JP-P-2021-115100 2021년07월12일 일본(JP)
JP-P-2022-014722 2022년02월02일 일본(JP)

(71) 출원인
소니그룹주식회사
일본국 도쿄도 미나토쿠 코난 1-7-1
(72) 발명자
고노 아키후미
일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 그룹 주식회사 내
치넨 도루
일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 그룹 주식회사 내
(뒷면에 계속)
(74) 대리인
양영준, 이중희

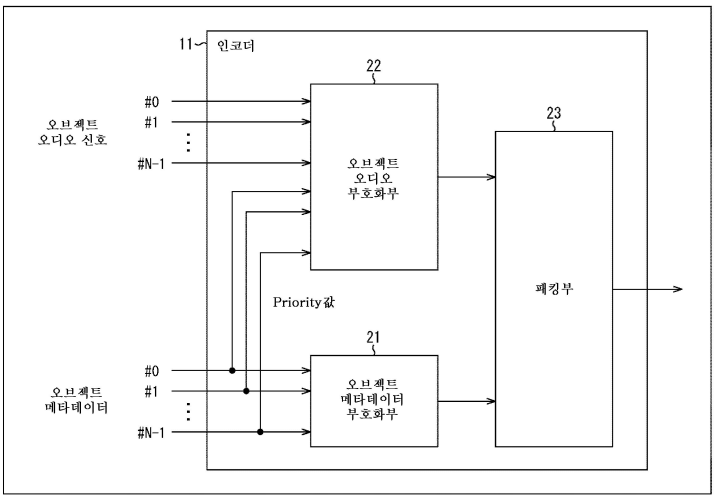
전체 청구항 수 : 총 43 항

(54) 발명의 명칭 **부호화 장치 및 방법, 복호 장치 및 방법, 그리고 프로그램**

(57) 요 약

본 기술은, 실시간 동작을 유지한 상태에서 부호화 효율을 향상시킬 수 있도록 하는 부호화 장치 및 방법, 복호 장치 및 방법, 그리고 프로그램에 관한 것이다. 부호화 장치는, 오디오 신호 및 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하는 우선도 정보 생성부와, 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하는 시간 주파수 변환부와, 복수의 오디오 신호에 대해서, 우선도 정보에 의해 나타내지는 우선도가 높은 오디오 신호로부터 차례로, 오디오 신호의 MDCT 계수의 양자화를 행하는 비트 얼로케이션부를 구비한다. 본 기술은 부호화 장치에 적용할 수 있다.

대 표 도



(72) 발명자

혼마 히로유키

일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 그룹 주식회사 내

하타나카 미츠유키

일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니 그룹 주식회사 내

명 세 서

청구범위

청구항 1

오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하는 우선도 정보 생성부와,

상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하는 시간 주파수 변환부와,

복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는 비트 얼로케이션부

를 구비하는 부호화 장치.

청구항 2

제1항에 있어서, 상기 비트 얼로케이션부는, 상기 복수의 상기 오디오 신호의 상기 MDCT 계수에 대해서 필요 최소한의 양자화 처리를 행함과 함께, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 필요 최소한의 양자화 처리의 결과에 기초하여 상기 MDCT 계수를 양자화하는 부가적인 양자화 처리를 행하는, 부호화 장치.

청구항 3

제2항에 있어서, 상기 비트 얼로케이션부는, 소정의 제한 시간 내에 모든 상기 오디오 신호에 대해서 상기 부가적인 양자화 처리를 행할 수 없었을 경우, 상기 부가적인 양자화 처리가 완료되지 않은 상기 오디오 신호의 양자화 결과로서, 상기 필요 최소한의 양자화 처리의 결과를 출력하는, 부호화 장치.

청구항 4

제3항에 있어서, 상기 비트 얼로케이션부는, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 필요 최소한의 양자화 처리를 행하는, 부호화 장치.

청구항 5

제4항에 있어서, 상기 비트 얼로케이션부는, 상기 제한 시간 내에 모든 상기 오디오 신호에 대해서 상기 필요 최소한의 양자화 처리를 행할 수 없었을 경우, 상기 필요 최소한의 양자화 처리가 완료되지 않은 상기 오디오 신호의 양자화 결과로서, 제로 데이터의 양자화 값을 출력하는, 부호화 장치.

청구항 6

제5항에 있어서, 상기 비트 얼로케이션부는, 상기 오디오 신호의 양자화 결과가 상기 제로 데이터의 양자화 값인지를 나타내는 뮤트 정보를 또한 출력하는, 부호화 장치.

청구항 7

제3항에 있어서, 상기 비트 얼로케이션부는, 상기 비트 얼로케이션부의 후단에 있어서 필요해지는 처리 시간에 기초하여 상기 제한 시간을 결정하는, 부호화 장치.

청구항 8

제7항에 있어서, 상기 비트 얼로케이션부는, 지금까지 행한 상기 필요 최소한의 양자화 처리의 결과 또는 상기 부가적인 양자화 처리의 결과에 기초하여, 상기 제한 시간을 동적으로 변경하는, 부호화 장치.

청구항 9

제2항에 있어서, 상기 우선도 정보 생성부는, 상기 오디오 신호의 음압, 상기 오디오 신호의 스펙트럼 형상 또는 복수의 상기 오디오 신호간의 상기 스펙트럼 형상의 상관에 기초하여, 상기 우선도 정보를 생성하는, 부호화

장치.

청구항 10

제2항에 있어서, 상기 메타데이터에는, 미리 생성된 상기 오디오 신호의 우선도를 나타내는 Priority값이 포함되어 있는, 부호화 장치.

청구항 11

제2항에 있어서, 상기 메타데이터에는, 상기 오디오 신호에 기초하는 소리의 음원 위치를 나타내는 위치 정보가 포함되어 있고,

상기 우선도 정보 생성부는, 적어도 상기 위치 정보와, 유저의 청취 위치를 나타내는 청취 위치 정보에 기초하여 상기 우선도 정보를 생성하는, 부호화 장치.

청구항 12

제2항에 있어서, 상기 복수의 상기 오디오 신호에는, 오브젝트의 상기 오디오 신호 및 채널의 상기 오디오 신호 중 적어도 어느 한 쪽이 포함되어 있는, 부호화 장치.

청구항 13

제2항에 있어서, 상기 오디오 신호에 기초하여 청각 심리 파라미터를 계산하는 청각 심리 파라미터 계산부를 더 구비하고,

상기 비트 얼로케이션부는, 상기 청각 심리 파라미터에 기초하여, 상기 필요 최소한의 양자화 처리 및 상기 부가적인 양자화 처리를 행하는, 부호화 장치.

청구항 14

제2항에 있어서, 상기 비트 얼로케이션부로부터 출력된, 상기 오디오 신호의 양자화 결과를 부호화하는 부호화부를 더 구비하는, 부호화 장치.

청구항 15

제13항에 있어서, 상기 청각 심리 파라미터 계산부는, 상기 오디오 신호와, 상기 오디오 신호에 관한 마스킹 역치에 관한 설정 정보에 기초하여 상기 청각 심리 파라미터를 계산하는, 부호화 장치.

청구항 16

부호화 장치가,

오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하고,

상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는, 부호화 방법.

청구항 17

오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하고,

상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는 처리를 컴퓨터에 실행시키는 프로그램.

청구항 18

복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초

하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고, 상기 부호화 오디오 신호를 복호하는 복호부를 구비하는, 복호 장치.

청구항 19

제18항에 있어서, 상기 복호부는, 상기 오디오 신호의 양자화 결과가 제로 데이터의 양자화 값인지를 나타내는 뮤트 정보를 또한 취득하고, 상기 뮤트 정보에 따라, 상기 복호에 의해 얻어진 상기 MDCT 계수에 기초하여 상기 오디오 신호를 생성하거나 또는 상기 MDCT 계수를 0으로 해서 상기 오디오 신호를 생성하는, 복호 장치.

청구항 20

복호 장치가,

복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고,

상기 부호화 오디오 신호를 복호하는, 복호 방법.

청구항 21

복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고,

상기 부호화 오디오 신호를 복호하는 처리를 컴퓨터에 실행시키는 프로그램.

청구항 22

오디오 신호를 부호화하여, 부호화 오디오 신호를 생성하는 부호화부와,

프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 보유하는 버퍼와,

처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는 삽입부

를 구비하는 부호화 장치.

청구항 23

제22항에 있어서, 상기 오디오 신호의 MDCT 계수의 양자화를 행하는 비트 얼로케이션부를 더 구비하고,

상기 부호화부는, 상기 MDCT 계수의 양자화 결과를 부호화하는, 부호화 장치.

청구항 24

제23항에 있어서, 상기 부호화 무음 데이터를 생성하는 생성부를 더 구비하는, 부호화 장치.

청구항 25

제24항에 있어서, 상기 생성부는, 무음 데이터의 MDCT 계수의 양자화 값을 부호화함으로써 상기 부호화 무음 데이터를 생성하는, 부호화 장치.

청구항 26

제24항에 있어서, 상기 생성부는, 1프레임분의 상기 무음 데이터에만 기초하여 상기 부호화 무음 데이터를 생성하는, 부호화 장치.

청구항 27

제24항에 있어서, 상기 오디오 신호는, 채널 또는 오브젝트의 오디오 신호이며,

상기 생성부는, 채널수 및 오브젝트수의 적어도 어느 것에 기초하여, 상기 부호화 무음 데이터를 생성하는, 부호화 장치.

청구항 28

제22항에 있어서, 상기 삽입부는, 상기 처리 대상의 프레임 종별에 따라서 상기 부호화 무음 데이터의 삽입을 행하는, 부호화 장치.

청구항 29

제28항에 있어서, 상기 삽입부는, 상기 처리 대상의 프레임이 랜덤 액세스 가능한 프레임의 프리롤 프레임일 경우, 상기 랜덤 액세스 가능한 프레임에 관한 상기 프리롤 프레임의 상기 부호화 오디오 신호로서 상기 부호화 무음 데이터를 상기 비트 스트림에 삽입하는, 부호화 장치.

청구항 30

제28항에 있어서, 상기 삽입부는, 상기 처리 대상의 프레임이 랜덤 액세스 가능한 프레임일 경우, 상기 처리 대상의 프레임에 관한 프리롤 프레임의 상기 부호화 오디오 신호로서 상기 부호화 무음 데이터를 상기 비트 스트림에 삽입하는, 부호화 장치.

청구항 31

제23항에 있어서, 상기 삽입부는, 상기 비트 얼로케이션부에서, 상기 MDCT 계수에 대해서 필요 최소한의 양자화 처리만을 행하거나, 또는 상기 MDCT 계수에 대해서 상기 필요 최소한의 양자화 처리 후에 행해지는 부가적인 양자화 처리를 도중에 중단하면, 상기 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료될 경우, 상기 부호화 무음 데이터의 삽입을 행하지 않는, 부호화 장치.

청구항 32

제22항에 있어서, 상기 부호화부는, 상기 오디오 신호에 대해서 가변 길이 부호화를 행하는, 부호화 장치.

청구항 33

제32항에 있어서, 상기 가변 길이 부호화는, 컨텍스트 베이스의 산술 부호화인, 부호화 장치.

청구항 34

부호화 장치가,

오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고,

프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 버퍼에 보유하고,

처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는, 부호화 방법.

청구항 35

오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고,

프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 버퍼에 보유하고,

처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는 처리를 컴퓨터에 실행시키는 프로그램.

청구항 36

오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상

기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는 복호부를 구비하는, 복호 장치.

청구항 37

복호 장치가,

오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는, 복호 방법.

청구항 38

오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는 처리를 컴퓨터에 실행시키는 프로그램.

청구항 39

오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하는 시간 주파수 변환부와,

상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산하는 청각 심리 파라미터 계산부와,

상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는 비트 얼로케이션부

를 구비하는 부호화 장치.

청구항 40

제39항에 있어서, 상기 설정 정보에는, 주파수마다 설정된 상기 마스킹 역치의 상한값을 나타내는 정보가 포함되어 있는, 부호화 장치.

청구항 41

제39항에 있어서, 상기 설정 정보에는, 1개 또는 복수의 상기 오브젝트마다 설정된 상기 마스킹 역치의 상한값을 나타내는 정보가 포함되어 있는, 부호화 장치.

청구항 42

부호화 장치가,

오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산하고,

상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는, 부호화 방법.

청구항 43

오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산

하고,

상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는 스텝을 포함하는 처리를 컴퓨터에 실행시키는 프로그램.

발명의 설명

기 술 분 야

[0001] 본 기술은, 부호화 장치 및 방법, 복호 장치 및 방법, 그리고 프로그램에 관한 것으로, 특히, 실시간 동작을 유지한 상태에서 부호화 효율을 향상시킬 수 있도록 한 부호화 장치 및 방법, 복호 장치 및 방법, 그리고 프로그램에 관한 것이다.

배 경 기 술

[0002] 종래, 국제 표준 규격인 MPEG(Moving Picture Experts Group)-D USAC(Unified Speech and Audio Coding) 규격이나, MPEG-D USAC 규격을 Core Coder로 한 MPEG-H 3D Audio 규격의 부호화 기술 등이 알려져 있다(예를 들어, 비특허문헌 1 내지 비특허문헌 3 참조).

선행기술문헌

비특허문헌

[0003] (비특허문헌 0001) ISO/IEC 23003-3, MPEG-D USAC

(비특허문헌 0002) ISO/IEC 23008-3, MPEG-H 3D Audio

(비특허문헌 0003) ISO/IEC 23008-3:2015/AMENDMENT3, MPEG-H 3D Audio Phase 2

발명의 내용

해결하려는 과제

[0004] MPEG-H 3D Audio 규격 등에서 취급되는 3D Audio에서는, 음 소재(오브젝트)의 위치를 나타내는 수평 각도나 수직 각도, 거리, 오브젝트에 관한 게인 등과 같은 오브젝트마다의 메타데이터를 갖고, 3차원적인 소리의 방향이나 거리, 확산 등을 재현할 수 있다. 그 때문에, 3D Audio에서는, 종래의 스테레오 재생에 비하여, 보다 현장감이 있는 오디오 재생이 가능해진다.

[0005] 그러나, 3D Audio로 실현되는 다수의 오브젝트의 데이터를 전송하기 위해서는, 보다 많은 오디오 채널을 압축 효율 좋게 고속으로 복호 가능한 부호화 기술이 필요해진다. 즉, 부호화 효율의 향상이 요망되고 있다.

[0006] 또한, 라이브나 콘서트를 3D Audio로 라이브 배신하기 위해서는, 부호화 효율의 향상과 실시간성의 양립이 필요해진다.

[0007] 본 기술은, 이러한 상황을 감안하여 이루어진 것으로, 실시간 동작을 유지한 상태에서 부호화 효율을 향상시킬 수 있도록 하는 것이다.

과제의 해결 수단

[0008] 본 기술의 제1 측면의 부호화 장치는, 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하는 우선도 정보 생성부와, 상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하는 시간 주파수 변환부와, 복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는 비트 얼로케이션부를 구비한다.

[0009] 본 기술의 제1 측면의 부호화 방법 또는 프로그램은, 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하고, 상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고, 복수의 상기 오디오 신호에 대해서, 상기 우선도 정보

에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는 스텝을 포함한다.

[0010] 본 기술의 제1 측면에서는, 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보가 생성되고, 상기 오디오 신호에 대한 시간 주파수 변환이 행해져서 MDCT 계수가 생성되고, 복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화가 행해진다.

[0011] 본 기술의 제2 측면의 복호 장치는, 복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고, 상기 부호화 오디오 신호를 복호하는 복호부를 구비한다.

[0012] 본 기술의 제2 측면의 복호 방법 또는 프로그램은, 복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고, 상기 부호화 오디오 신호를 복호하는 스텝을 포함한다.

[0013] 본 기술의 제2 측면에서는, 복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호가 취득되고, 상기 부호화 오디오 신호가 복호된다.

[0014] 본 기술의 제3 측면의 부호화 장치는, 오디오 신호를 부호화하여, 부호화 오디오 신호를 생성하는 부호화부와, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 보유하는 버퍼와, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는 삽입부를 구비한다.

[0015] 본 기술의 제3 측면의 부호화 방법 또는 프로그램은, 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 버퍼에 보유하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는 스텝을 포함한다.

[0016] 본 기술의 제3 측면에서는, 오디오 신호가 부호화되어 부호화 오디오 신호가 생성되고, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림이 버퍼에 보유되고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터가 상기 비트 스트림에 삽입된다.

[0017] 본 기술의 제4 측면의 복호 장치는, 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는 복호부를 구비한다.

[0018] 본 기술의 제4 측면의 복호 방법 또는 프로그램은, 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는 스텝을 포함한다.

[0019] 본 기술의 제4 측면에서는, 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림이 취득되고, 상기 부호화 오디오 신호가 복호

된다.

[0020] 본 기술의 제5 측면의 부호화 장치는, 오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하는 시간 주파수 변환부와, 상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산하는 청각 심리 파라미터 계산부와, 상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는 비트 얼로케이션부를 구비한다.

[0021] 본 기술의 제5 측면의 부호화 방법 또는 프로그램은, 오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고, 상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산하고, 상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는 스텝을 포함한다.

[0022] 본 기술의 제5 측면에서는, 오브젝트의 오디오 신호에 대한 시간 주파수 변환이 행해져서, MDCT 계수가 생성되고, 상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터가 계산되고, 상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리가 행해져서, 양자화 MDCT 계수가 생성된다.

도면의 간단한 설명

[0023] 도 1은 인코더의 구성예를 도시하는 도면이다.

도 2는 오브젝트 오디오 부호화부의 구성예를 도시하는 도면이다.

도 3은 부호화 처리를 설명하는 흐름도이다.

도 4는 비트 얼로케이션 처리를 설명하는 흐름도이다.

도 5는 메타데이터의 Config의 신택스 예를 도시하는 도면이다.

도 6은 디코더의 구성예를 도시하는 도면이다.

도 7은 언패킹/복호부의 구성예를 도시하는 도면이다.

도 8은 복호 처리를 설명하는 흐름도이다.

도 9는 선택 복호 처리를 설명하는 흐름도이다.

도 10은 오브젝트 오디오 부호화부의 구성예를 도시하는 도면이다.

도 11은 콘텐츠 배신 시스템의 구성예를 도시하는 도면이다.

도 12는 입력 데이터의 예에 대해서 설명하는 도면이다.

도 13은 컨텍스트의 계산에 대해서 설명하는 도면이다.

도 14는 인코더의 구성예를 도시하는 도면이다.

도 15는 오브젝트 오디오 부호화부의 구성예를 도시하는 도면이다.

도 16은 초기화부의 구성예를 도시하는 도면이다.

도 17은 진척 정보와 처리 완료 가부 판정의 예에 대해서 설명하는 도면이다.

도 18은 부호화 데이터를 포함하는 비트 스트림의 예에 대해서 설명하는 도면이다.

도 19는 부호화 데이터의 신택스 예를 도시하는 도면이다.

도 20은 확장 데이터의 예를 도시하는 도면이다.

도 21은 세그먼트 데이터에 대해서 설명하는 도면이다.

도 22는 AudioPreRoll()의 구성예를 도시하는 도면이다.

도 23은 초기화 처리를 설명하는 흐름도이다.

도 24는 부호화 처리를 설명하는 흐름도이다.

도 25는 부호화 Mute 데이터 삽입 처리를 설명하는 흐름도이다.

도 26은 언패킹/복호부의 구성예를 도시하는 도면이다.

도 27은 복호 처리를 설명하는 흐름도이다.

도 28은 인코더의 구성예를 도시하는 도면이다.

도 29는 오브젝트 오디오 부호화부의 구성예를 도시하는 도면이다.

도 30은 부호화 처리를 설명하는 흐름도이다.

도 31은 컴퓨터의 구성예를 도시하는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0024] 이하, 도면을 참조하여, 본 기술을 적용한 실시 형태에 대해서 설명한다.

[0025] <제1 실시 형태>

[0026] <본 기술에 대해서>

[0027] 본 기술은, 오브젝트(음성)의 중요도를 고려한 부호화 처리를 행함으로써, 실시간 동작을 유지한 상태에서 부호화 효율을 향상시켜, 전송 가능한 오브젝트수를 증가시킬 수 있도록 하는 것이다.

[0028] 예를 들어 라이브 배신을 실현하기 위해서는, 부호화 처리를 실시간 처리로 행할 것이 요구된다. 즉, 1초간에 f프레임의 음성을 배신할 경우, 1프레임의 부호화 및 비트 스트림 출력을 1/f초 이내에 완료해야 한다.

[0029] 이와 같이, 부호화 처리를 실시간 처리로 행한다는 목표를 달성하기 위해서는, 이하의 어프로치가 유효하다.

[0030] ·부호화 처리를 단계적으로 행한다.

[0031] 우선은 필요 최소한의 부호화를 완료하고, 그 후에 부호화 효율을 높인 부가적인 부호화 처리를 행한다. 사전에 정한 소정의 제한 시간이 경과한 시점에서 부가적인 부호화 처리가 완료되지 않았을 경우, 그 시점에서 처리를 중단하고, 직전의 단계에서의 부호화 처리의 결과를 출력한다.

[0032] ·또한, 소정의 제한 시간이 지난 시점에서 필요 최소한의 부호화도 완료되지 않았을 경우, 처리를 중단하고, 사전에 준비해 둔 Mute 데이터의 비트 스트림을 출력한다.

[0033] 그런데, 멀티채널이나 복수의 오브젝트의 오디오 신호가 동시에 재생될 경우, 그러한 오디오 신호에 의해 재생되는 음성 중에는, 다른 음성과 비교했을 때 중요한 음성과, 그다지 중요하지 않은 음성이 존재한다. 예를 들어 중요하지 않은 음성이란, 전체 음성 중에서, 어떤 특정 음성이 재생되지 않았다고 해도, 그에 의해 청취자에게 위화감을 부여하지 않을 정도의 음성 등이다.

[0034] 음성의 중요도, 즉 채널이나 오브젝트의 중요도를 고려하지 않는 처리순으로 부호화 효율을 높인 부가적인 부호화 처리를 행하면, 중요한 음성임에도 불구하고 처리가 중단되어 음질이 열화되어 버리는 경우가 있다.

[0035] 그래서, 본 기술에서는, 음성의 중요도 순으로 부호화 효율을 높인 부가적인 부호화 처리를 행함으로써, 실시간 동작을 유지한 상태에서 콘텐츠 전체의 부호화 효율을 향상시킬 수 있도록 하였다.

[0036] 이와 같이 하면, 중요도가 높은 음성일수록 부가적인 부호화 처리가 완료되고, 부가적인 부호화 처리가 완료되지 않고, 필요 최소한의 부호화만이 행해지는 것은, 보다 중요도가 낮은 음성으로 되므로, 콘텐츠 전체의 부호화 효율을 향상시킬 수 있다. 이에 의해, 전송 가능한 오브젝트수를 증가시킬 수 있다.

[0037] 이상과 같이 본 기술에서는, 멀티채널을 구성하는 각 채널의 오디오 신호, 및 오브젝트의 오디오 신호의 부호화에 있어서, 각 채널의 오디오 신호와 각 오브젝트의 오디오 신호의 우선도가 높은 순으로, 부호화 효율을 높인 부가적인 부호화 처리가 행해진다. 이에 의해, 실시간 처리에서의 콘텐츠 전체의 부호화 효율을 향상시킬 수 있다.

[0038] 또한, 이하에서는, 오브젝트의 오디오 신호가 MPEG-H 규격에 따라서 부호화되는 경우에 대해서 설명하지만, 채널의 오디오 신호를 포함한 MPEG-H 규격에 따라서 부호화가 행해지는 경우나, 다른 방식으로 부호화가 행해지는 경우에도 마찬가지의 처리가 행해진다.

[0039] <인코더의 구성예>

[0040] 도 1은, 본 기술을 적용한 인코더의 일 실시 형태의 구성예를 도시하는 도면이다.

[0041] 도 1에 도시하는 인코더(11)는, 예를 들어 인코더(부호화 장치)로서 기능하는 컴퓨터 등의 신호 처리 장치 등을 포함한다.

[0042] 도 1에 도시하는 예는, N개의 오브젝트의 오디오 신호와, 그러한 N개의 오브젝트의 메타데이터가 인코더(11)에 입력되어, MPEG-H 규격에 따라서 부호화가 행해지는 예로 되어 있다. 또한, 도 1에서 #0 내지 #N-1은, N개의 각 오브젝트를 나타내는 오브젝트 번호를 나타내고 있다.

[0043] 인코더(11)는, 오브젝트 메타데이터 부호화부(21), 오브젝트 오디오 부호화부(22) 및 패킹부(23)를 갖고 있다.

[0044] 오브젝트 메타데이터 부호화부(21)는, 공급된 N개의 각 오브젝트의 메타데이터를 MPEG-H 규격에 따라서 부호화하고, 그 결과 얻어진 부호화 메타데이터를 패킹부(23)에 공급한다.

[0045] 예를 들어 오브젝트의 메타데이터에는, 3차원 공간에서의 오브젝트의 위치를 나타내는 오브젝트 위치 정보, 오브젝트의 우선 정도(중요 정도)를 나타내는 Priority값, 및 오브젝트의 오디오 신호의 게인 보정을 위한 게인을 나타내는 게인값이 포함되어 있다. 특히, 이 예에서는, 메타데이터에는 적어도 Priority값이 포함되어 있다.

[0046] 여기서, 오브젝트 위치 정보는, 예를 들어 수평 각도(Azimuth), 수직 각도(Elevation) 및 거리(Radius)를 포함한다.

[0047] 수평 각도 및 수직 각도는, 3차원 공간에서의 기준이 되는 청취 위치에서 본 오브젝트의 위치를 나타내는 수평 방향 및 수직 방향의 각도이다. 또한, 거리(Radius)는 3차원 공간에서의 오브젝트의 위치를 나타내는, 기준이 되는 청취 위치부터 오브젝트까지의 거리를 나타내고 있다. 이러한 오브젝트 위치 정보는, 오브젝트의 오디오 신호에 기초하는 소리의 음원 위치를 나타내는 정보라고 할 수 있다.

[0048] 그 밖에, 오브젝트의 메타데이터에는, 오브젝트의 음상을 확장하는 스프레드 처리를 위한 파라미터 등이 포함되어 있어도 된다.

[0049] 오브젝트 오디오 부호화부(22)는, 공급된 각 오브젝트의 메타데이터에 포함되어 있는 Priority값에 기초하여, 공급된 N개의 각 오브젝트의 오디오 신호를 MPEG-H 규격에 따라서 부호화하고, 그 결과 얻어진 부호화 오디오 신호를 패킹부(23)에 공급한다.

[0050] 패킹부(23)는, 오브젝트 메타데이터 부호화부(21)로부터 공급된 부호화 메타 데이터와, 오브젝트 오디오 부호화부(22)로부터 공급된 부호화 오디오 신호를 패킹하고, 그 결과 얻어진 부호화 비트 스트림을 출력한다.

[0051] <오브젝트 오디오 부호화부의 구성예>

[0052] 또한, 오브젝트 오디오 부호화부(22)는, 예를 들어 도 2에 도시하는 바와 같이 구성된다.

[0053] 도 2의 예에서는, 오브젝트 오디오 부호화부(22)는, 우선도 정보 생성부(51), 시간 주파수 변환부(52), 청각 심리 파라미터 계산부(53), 비트 얼로케이션부(54) 및 부호화부(55)를 갖고 있다.

[0054] 우선도 정보 생성부(51)는, 공급된 각 오브젝트의 오디오 신호와, 공급된 각 오브젝트의 메타데이터에 포함되는 Priority값의 적어도 어느 한쪽에 기초하여, 각 오브젝트의 우선도, 즉 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하여, 비트 얼로케이션부(54)에 공급한다.

[0055] 예를 들어 우선도 정보 생성부(51)는, 오디오 신호의 음압이나 스펙트럼 형상, 복수의 각 오브젝트나 채널의 오디오 신호와의 사이의 스펙트럼 형상의 상관 등에 기초하여, 오브젝트의 오디오 신호가 어느 정도의 우선 정도인지의 해석을 행한다. 그리고 우선도 정보 생성부(51)는, 그 해석 결과에 기초하여 우선도 정보를 생성한다.

[0056] 또한, 예를 들어 MPEG-H의 오브젝트의 메타데이터에는, 오브젝트의 우선도를 나타내는 파라미터인 Priority값이 0부터 7까지의 3bit의 정수로서 포함되어 있고, Priority값이 클수록, 우선도가 높은 오브젝트인 것을 나타내고 있다.

[0057] 이 Priority값에 대해서는, 콘텐츠 제작자가 의도해서 Priority값을 설정하는 경우도 있고, 메타데이터를 생성하는 애플리케이션이 각 오브젝트의 오디오 신호를 해석해서 자동적으로 Priority값이 설정되는 경우도 있을 수 있다. 또한, 콘텐츠 제작자의 의도도 오디오 신호의 해석도 없이, 애플리케이션의 디폴트로서, 예를 들어 Priority값에 최고 우선도 「7」 등의 고정값이 설정되어 있을 수도 있다.

[0058] 따라서, 우선도 정보 생성부(51)에서 오브젝트(오디오 신호)의 우선도 정보를 생성할 때는, Priority값을 사용하지 않고 오디오 신호의 해석 결과만을 사용하도록 해도 되고, Priority값과 해석 결과 양쪽을 사용하도록 해도 된다.

[0059] 예를 들어 Priority값과 해석 결과 양쪽을 사용하는 경우, 오디오 신호의 해석 결과가 동일하여도, 보다 큰(높은) Priority값을 갖는 오브젝트의 우선도가, 보다 높아지도록 할 수 있다.

[0060] 시간 주파수 변환부(52)는, 공급된 각 오브젝트의 오디오 신호에 대해서 MDCT(Modified Discrete Cosine Transform)(수정 이산 코사인 변환)를 사용한 시간 주파수 변환을 행한다.

[0061] 시간 주파수 변환부(52)는, 시간 주파수 변환에 의해 얻어진, 각 오브젝트의 주파수 스펙트럼 정보인 MDCT 계수를 비트 얼로케이션부(54)에 공급한다.

[0062] 청각 심리 파라미터 계산부(53)는, 공급된 각 오브젝트의 오디오 신호에 기초하여, 인간의 청각 특성(청각 마스킹)을 고려하기 위한 청각 심리 파라미터를 계산하여, 비트 얼로케이션부(54)에 공급한다.

[0063] 비트 얼로케이션부(54)는, 우선도 정보 생성부(51)로부터 공급된 우선도 정보, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수, 및 청각 심리 파라미터 계산부(53)로부터 공급된 청각 심리 파라미터에 기초하여, 비트 얼로케이션 처리를 행한다.

[0064] 비트 얼로케이션 처리에서는, 각 스케일 팩터 밴드의 양자화 비트와 양자화 노이즈의 계산 및 평가를 행하는, 청각 심리 모델에 기초한 비트 얼로케이션이 행해진다. 그리고 그 비트 얼로케이션의 결과에 기초하여 스케일 팩터 밴드마다 MDCT 계수가 양자화되어, 양자화 MDCT 계수가 얻어진다.

[0065] 비트 얼로케이션부(54)는, 이와 같이 하여 얻어진 각 오브젝트의 스케일 팩터 밴드마다의 양자화 MDCT 계수를, 각 오브젝트의 양자화 결과, 보다 상세하게는 각 오브젝트의 MDCT 계수의 양자화 결과로서 부호화부(55)에 공급한다.

[0066] 여기서, 스케일 팩터 밴드란, 인간의 청각 특성에 기초하여 소정 대역폭의 서브 밴드(여기서는 MDCT의 분해능)를 복수 묶어서 얻어지는 대역(주파수 대역)이다.

[0067] 이상과 같은 비트 얼로케이션 처리에 의해, MDCT 계수의 양자화로 발생해 버리는 양자화 노이즈가 마스크되어 지각되지 않는 스케일 팩터 밴드의 양자화 비트의 일부가, 양자화 노이즈가 지각되기 쉬운 스케일 팩터 밴드에 할당된다(돌려진다). 이에 의해, 전체적으로 음질의 열화를 억제하고, 효율적인 양자화를 행할 수 있다. 즉, 부호화 효율을 향상시킬 수 있다.

[0068] 또한, 비트 얼로케이션부(54)는, 실시간 처리를 위한 제한 시간 내에 양자화 MDCT 계수를 얻을 수 없었던 오브젝트에 대해서는, 그 오브젝트의 양자화 결과로서, 미리 준비된 Mute 데이터를 부호화부(55)에 공급한다.

[0069] Mute 데이터는, 각 스케일 팩터 밴드의 MDCT 계수의 값 「0」을 나타내는 제로 데이터이며, 보다 상세하게는 Mute 데이터의 양자화 값, 즉 MDCT 계수 「0」의 양자화 MDCT 계수가 부호화부(55)에 출력된다. 또한, 여기서는 Mute 데이터를 부호화부(55)에 출력하고 있지만, Mute 데이터를 공급하는 대신에, 양자화 결과(양자화 MDCT 계수)가 Mute 데이터인지 여부를 나타내는 Mute 정보를 부호화부(55)에 공급해도 된다. 그 경우, 부호화부(55)에서 Mute 정보에 따라서, 통상의 부호화 처리를 행할지 직접 MDCT 계수 「0」의 양자화 MDCT 계수의 부호화를 행할지를 전환한다. 또한 MDCT 계수 「0」의 양자화 MDCT 계수의 부호화를 행하는 대신에, 사전에 준비해 둔 MDCT 계수 「0」의 부호화 완료 데이터를 사용해도 된다.

[0070] 또한, 비트 얼로케이션부(54)는, 예를 들어 오브젝트마다, 양자화 결과(양자화 MDCT 계수)가 Mute 데이터인지 여부를 나타내는 Mute 정보를 패킹부(23)에 공급한다. 패킹부(23)는, 비트 얼로케이션부(54)로부터 공급된 Mute 정보를 부호화 비트 스트림의 부가 영역 등에 저장한다.

[0071] 부호화부(55)는, 비트 얼로케이션부(54)로부터 공급된 각 오브젝트의 스케일 팩터 밴드마다의 양자화 MDCT 계수를 부호화하고, 그 결과 얻어진 부호화 오디오 신호를 패킹부(23)에 공급한다.

[0072] <부호화 처리의 설명>

[0073] 계속해서, 인코더(11)의 동작에 대해서 설명한다. 즉, 이하, 도 3의 흐름도를 참조하여, 인코더(11)에 의한 부호화 처리에 대해서 설명한다.

[0074] 스텝 S11에서 오브젝트 메타데이터 부호화부(21)는, 공급된 각 오브젝트의 메타데이터를 부호화하고, 그 결과

얻어진 부호화 메타데이터를 패킹부(23)에 공급한다.

[0075] 스텝 S12에서 우선도 정보 생성부(51)는, 공급된 각 오브젝트의 오디오 신호와, 공급된 각 오브젝트의 메타데이터의 Priority값의 적어도 어느 한쪽에 기초하여 각 오브젝트의 우선도 정보를 생성하여, 비트 얼로케이션부(54)에 공급한다.

[0076] 스텝 S13에서 시간 주파수 변환부(52)는, 공급된 각 오브젝트의 오디오 신호에 대해서 MDCT를 사용한 시간 주파수 변환을 행하고, 그 결과 얻어진 스케일 팩터 밴드마다의 MDCT 계수를 비트 얼로케이션부(54)에 공급한다.

[0077] 스텝 S14에서 청각 심리 파라미터 계산부(53)는, 공급된 각 오브젝트의 오디오 신호에 기초하여 청각 심리 파라미터를 계산하여, 비트 얼로케이션부(54)에 공급한다.

[0078] 스텝 S15에서 비트 얼로케이션부(54)는, 우선도 정보 생성부(51)로부터 공급된 우선도 정보, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수, 및 청각 심리 파라미터 계산부(53)로부터 공급된 청각 심리 파라미터에 기초하여, 비트 얼로케이션 처리를 행한다.

[0079] 비트 얼로케이션부(54)는, 비트 얼로케이션 처리에 의해 얻어진 양자화 MDCT 계수를 부호화부(55)에 공급함과 함께, 패킹부(23)에 대해서 Mute 정보를 공급한다. 또한, 비트 얼로케이션 처리의 상세는 후술한다.

[0080] 스텝 S16에서 부호화부(55)는, 비트 얼로케이션부(54)로부터 공급된 양자화 MDCT 계수를 부호화하고, 그 결과 얻어진 부호화 오디오 신호를 패킹부(23)에 공급한다.

[0081] 예를 들어 부호화부(55)에서는, 양자화 MDCT 계수에 대해서 컨텍스트 베이스의 산술 부호화가 행해지고, 부호화된 양자화 MDCT 계수가 부호화 오디오 신호로서 패킹부(23)에 출력된다. 또한, 부호화 방식은 산술 부호화에 한정되지 않는다. 예를 들어, 허프만 코딩화나 기타 부호화 방식에 의해 부호화되어도 된다.

[0082] 스텝 S17에서 패킹부(23)는, 오브젝트 메타데이터 부호화부(21)로부터 공급된 부호화 메타데이터, 및 부호화부(55)로부터 공급된 부호화 오디오 신호의 패킹을 행한다.

[0083] 이때, 패킹부(23)는, 비트 얼로케이션부(54)로부터 공급된 Mute 정보를 부호화 비트 스트림의 부가 영역 등에 저장한다.

[0084] 그리고 패킹부(23)는, 패킹에 의해 얻어진 부호화 비트 스트림을 출력하고, 부호화 처리는 종료된다.

[0085] 이상과 같이 해서 인코더(11)는, 오브젝트의 오디오 신호나 Priority값에 기초하여 우선도 정보를 생성하고, 우선도 정보를 사용하여 비트 얼로케이션 처리를 행한다. 이와 같이 함으로써, 실시간 처리에서의 콘텐츠 전체의 부호화 효율을 향상시켜, 보다 많은 오브젝트의 데이터를 전송할 수 있다.

[0086] <비트 얼로케이션 처리의 설명>

[0087] 이어서, 도 4의 흐름도를 참조하여, 도 3의 스텝 S15의 처리에 대응하는 비트 얼로케이션 처리에 대해서 설명한다.

[0088] 스텝 S41에서 비트 얼로케이션부(54)는, 우선도 정보 생성부(51)로부터 공급된 우선도 정보에 기초하여, 우선도 정보에 의해 나타내지는 우선도가 높은 오브젝트순으로, 각 오브젝트의 처리 순번(처리순)을 설정한다.

[0089] 이 예에서는, 전부 N개의 오브젝트 중, 가장 우선도가 높은 오브젝트의 처리순이 「0」으로 되고, 가장 우선도가 낮은 오브젝트의 처리순이 「N-1」로 된다. 또한, 처리순의 설정은 이것에 한정되지 않고, 예를 들어 가장 우선도가 높은 오브젝트의 처리순이 「1」로 되고, 가장 우선도가 낮은 오브젝트의 처리순이 「N」으로 되어도 되고, 숫자 이외의 기호로 우선도가 나타내지도록 해도 된다.

[0090] 이후에 있어서는, 우선도가 높은 오브젝트부터 차례로, 필요 최소한의 양자화 처리, 즉 필요 최소한의 부호화 처리가 행해진다.

[0091] 즉, 스텝 S42에서 비트 얼로케이션부(54)는, 처리 대상의 오브젝트를 나타내는 처리 대상 ID를 「0」으로 한다.

[0092] 이 처리 대상 ID의 값은, 「0」부터 1씩 인크리먼트되어 갱신되어 간다. 또한, 처리 대상 ID의 값을 n으로 하면, 그 처리 대상 ID에 의해 나타내지는 오브젝트는, 스텝 S41에서 설정된 처리순이 n번째인 오브젝트로 된다.

[0093] 따라서, 비트 얼로케이션부(54)에서는, 각 오브젝트가 스텝 S41에서 설정된 처리순으로 처리되어 가게 된다.

[0094] 스텝 S43에서 비트 얼로케이션부(54)는, 처리 대상 ID의 값이 N 미만인지 여부를 판정한다.

[0095] 스텝 S43에서 처리 대상 ID의 값이 N 미만이라고 판정된 경우, 즉, 아직 모든 오브젝트에 대해서 양자화의 처리를 행하지 않은 경우, 스텝 S44의 처리가 행해진다.

[0096] 즉, 스텝 S44에서 비트 얼로케이션부(54)는, 처리 대상 ID에 의해 나타내지는 처리 대상의 오브젝트의 스케일 팩터 밴드마다의 MDCT 계수에 대해서, 필요 최소한의 양자화 처리를 행한다.

[0097] 여기서, 필요 최소한의 양자화 처리란, 비트 얼로케이션 루프 처리 전에 행해지는 1회째의 양자화 처리이다.

[0098] 구체적으로는, 비트 얼로케이션부(54)는, 청각 심리 파라미터나 MDCT 계수에 기초하여, 각 스케일 팩터 밴드의 양자화 비트와 양자화 노이즈의 계산 및 평가를 행한다. 이에 의해, 스케일 팩터 밴드마다, 양자화 MDCT 계수가 목표로 하는 비트수(양자화 비트수)가 결정된다.

[0099] 비트 얼로케이션부(54)는, 각 스케일 팩터 밴드의 양자화 MDCT 계수가 목표로 하는 양자화 비트수 내의 데이터로 되도록, 스케일 팩터 밴드마다의 MDCT 계수를 양자화하여, 양자화 MDCT 계수를 구한다.

[0100] 또한, 비트 얼로케이션부(54)는, 처리 대상의 오브젝트에 대해서, 양자화 결과가 Mute 데이터가 아닌 취지의 Mute 정보를 생성하여, 보유한다.

[0101] 스텝 S45에서 비트 얼로케이션부(54)는, 실시간 처리를 위한 소정의 제한 시간 이내인지 여부를 판정한다.

[0102] 예를 들어 비트 얼로케이션 처리가 개시되고 나서 소정의 시간이 경과한 경우, 제한 시간 이내가 아니라고 판정된다.

[0103] 이 제한 시간은, 예를 들어 실시간으로 부호화 비트 스트림의 출력(배신)이 가능하게 되도록, 즉 부호화 처리를 실시간 처리로 행할 수 있도록, 비트 얼로케이션부(54) 후단의 부호화부(55) 및 패킹부(23)에서 필요해지는 처리 시간이 고려되어 비트 얼로케이션부(54)에 의해 설정(결정)되는 역치이다.

[0104] 또한, 이 제한 시간은, 비트 얼로케이션부(54)에서의 지금까지의 처리로 얻어진 오브젝트의 양자화 MDCT 계수의 값 등, 지금까지의 비트 얼로케이션 처리의 처리 결과에 기초하여, 동적으로 변경되도록 해도 된다.

[0105] 스텝 S45에서 제한 시간 이내라고 판정된 경우, 그 후, 처리는 스텝 S46으로 진행된다.

[0106] 스텝 S46에서 비트 얼로케이션부(54)는, 스텝 S44의 처리에 의해 얻어진 양자화 MDCT 계수를, 처리 대상의 오브젝트의 양자화 결과로서 보존(보유)함과 함께, 처리 대상 ID의 값에 「1」을 더한다. 이에 의해, 아직 필요 최소한의 양자화 처리가 행해지지 않은 새로운 오브젝트가 다음 처리 대상의 오브젝트로 된다.

[0107] 스텝 S46의 처리가 행해지면, 그 후, 처리는 스텝 S43으로 돌아가서, 상술한 처리가 반복해서 행해진다. 즉, 새로운 처리 대상의 오브젝트에 대해서, 필요 최소한의 양자화 처리가 행해진다.

[0108] 이와 같이 스텝 S43 내지 스텝 S46에서는, 우선도가 높은 순으로, 각 오브젝트에 대해서 필요 최소한의 양자화 처리가 행해진다. 이에 의해, 부호화 효율을 향상시킬 수 있다.

[0109] 또한, 스텝 S45에서 제한 시간 이내가 아니라고 판정된 경우, 즉 제한 시간이 되어 버렸을 경우, 오브젝트마다의 필요 최소한의 양자화 처리는 중단되고, 그 후, 처리는 스텝 S47로 진행된다. 즉, 이 경우, 처리 대상으로 되지 않은 오브젝트에 대해서는, 필요 최소한의 양자화 처리가 미완료인 채로 처리가 중단된다.

[0110] 스텝 S47에서 비트 얼로케이션부(54)는, 상술한 스텝 S43 내지 스텝 S46에서 처리 대상으로 되지 않은 오브젝트, 즉 필요 최소한의 양자화 처리가 미완료인 오브젝트에 대해서, 사전에 준비한 Mute 데이터의 양자화 값을, 그러한 각 오브젝트의 양자화 결과로서 보존(보유)한다.

[0111] 즉, 스텝 S47에서는, 필요 최소한의 양자화 처리가 미완료인 오브젝트에 대해서는, Mute 데이터의 양자화 값이 그 오브젝트의 양자화 결과로서 사용된다.

[0112] 또한, 비트 얼로케이션부(54)는, 필요 최소한의 양자화 처리가 미완료인 오브젝트에 대해서, 양자화 결과가 Mute 데이터인 취지의 Mute 정보를 생성하여, 보유한다.

[0113] 스텝 S47의 처리가 행해지면, 그 후, 처리는 스텝 S54로 진행된다.

[0114] 또한, 스텝 S43에서 처리 대상 ID의 값이 N 미만이 아니라고 판정된 경우, 즉, 모든 오브젝트에 대해서, 제한 시간 내에 필요 최소한의 양자화 처리가 완료된 경우, 스텝 S48의 처리가 행해진다.

[0115] 스텝 S48에서 비트 얼로케이션부(54)는, 처리 대상의 오브젝트를 나타내는 처리 대상 ID를 「0」으로 한다. 이

에 의해, 다시 우선도가 높은 것부터 차례로 처리 대상의 오브젝트로 되어, 이후의 처리가 행해지게 된다.

[0116] 스텝 S49에서 비트 얼로케이션부(54)는, 처리 대상 ID의 값이 N 미만인지 여부를 판정한다.

[0117] 스텝 S49에서 처리 대상 ID의 값이 N 미만이라고 판정된 경우, 즉, 아직 모든 오브젝트에 대해서 부가적인 양자화 처리(부가적인 부호화 처리)가 행해지지 않았을 경우, 스텝 S50의 처리가 행해진다.

[0118] 스텝 S50에서 비트 얼로케이션부(54)는, 처리 대상 ID에 의해 나타내지는 처리 대상의 오브젝트의 스케일 팩터 밴드마다의 MDCT 계수에 대해서, 부가적인 양자화 처리, 즉 부가적인 비트 얼로케이션 루프 처리를 1회 행하고, 필요에 따라 양자화 결과의 갱신 보존을 행한다.

[0119] 구체적으로는, 비트 얼로케이션부(54)는, 청각 심리 파라미터와, 필요 최소한의 양자화 처리 등의 지금까지의 처리로 얻어진 오브젝트의 스케일 팩터 밴드마다의 양자화 결과인 양자화 MDCT 계수에 기초하여, 각 스케일 팩터 밴드의 양자화 비트와 양자화 노이즈의 재계산 및 재평가를 행한다. 이에 의해, 스케일 팩터 밴드마다, 양자화 MDCT 계수가 목표로 하는 양자화 비트수가 새롭게 결정된다.

[0120] 비트 얼로케이션부(54)는, 각 스케일 팩터 밴드의 양자화 MDCT 계수가 목표로 하는 양자화 비트수 내의 데이터로 되도록, 스케일 팩터 밴드마다의 MDCT 계수를 다시 양자화하여, 양자화 MDCT 계수를 구한다.

[0121] 그리고 비트 얼로케이션부(54)는, 스텝 S50의 처리에 의해, 오브젝트의 양자화 결과로서 보유하고 있는 양자화 MDCT 계수보다, 보다 양자화 노이즈 등이 적은 고품질의 양자화 MDCT 계수가 얻어진 경우, 지금까지 보유하고 있었던 양자화 MDCT 계수를, 새롭게 얻어진 양자화 MDCT 계수로 바꾸어서 보존한다. 즉, 보유하고 있는 양자화 MDCT 계수가 갱신된다.

[0122] 스텝 S51에서 비트 얼로케이션부(54)는, 실시간 처리를 위한 소정의 제한 시간 이내인지 여부를 판정한다.

[0123] 예를 들어 스텝 S51에서는, 스텝 S45에서의 경우와 마찬가지로, 비트 얼로케이션 처리가 개시되고 나서 소정의 시간이 경과한 경우, 제한 시간 이내가 아니라고 판정된다.

[0124] 또한, 스텝 S51에서의 제한 시간은, 스텝 S45에서의 경우와 동일해도 되고, 상술한 바와 같이, 지금까지의 비트 얼로케이션 처리, 즉 필요 최소한의 양자화 처리나 부가적인 비트 얼로케이션 루프 처리의 처리 결과에 따라서 동적으로 변경되도록 해도 된다.

[0125] 스텝 S51에서 제한 시간 이내라고 판정된 경우, 제한 시간까지 아직 시간이 남아 있으므로, 처리는 스텝 S52로 진행된다.

[0126] 스텝 S52에서 비트 얼로케이션부(54)는, 부가적인 양자화 처리의 루프 처리, 즉 부가적인 비트 얼로케이션 루프 처리가 종료되었는지 여부를 판정한다.

[0127] 예를 들어 스텝 S52에서는, 부가적인 비트 얼로케이션 루프 처리가 미리 정해진 횟수만큼 반복해서 행해진 경우나, 최근 2회의 부가적인 비트 얼로케이션 루프 처리에서의 양자화 노이즈의 차분이 역치 이하인 경우 등에 루프 처리가 종료되었다고 판정된다.

[0128] 스텝 S52에서, 아직 루프 처리가 종료되지 않았다고 판정된 경우, 처리는 스텝 S50으로 돌아가서, 상술한 처리가 반복해서 행해진다.

[0129] 이에 반해, 스텝 S52에서 루프 처리가 종료되었다고 판정된 경우, 스텝 S53의 처리가 행해진다.

[0130] 스텝 S53에서 비트 얼로케이션부(54)는, 스텝 S50에서 갱신된 양자화 MDCT 계수를, 처리 대상의 오브젝트의 최종적인 양자화 결과로서 보존(보유)함과 함께, 처리 대상 ID의 값에 「1」을 더한다. 이에 의해, 아직 부가적인 양자화 처리가 행해지지 않은 새로운 오브젝트가 다음 처리 대상의 오브젝트로 된다.

[0131] 스텝 S53의 처리가 행해지면, 그 후, 처리는 스텝 S49로 돌아가서, 상술한 처리가 반복해서 행해진다. 즉, 새로운 처리 대상의 오브젝트에 대해서, 부가적인 양자화 처리가 행해진다.

[0132] 이와 같이 스텝 S49 내지 스텝 S53에서는, 우선도가 높은 순으로, 각 오브젝트에 대해서 부가적인 양자화 처리가 행해진다. 이에 의해, 부호화 효율을 더욱 향상시킬 수 있다.

[0133] 또한, 스텝 S51에서 제한 시간 이내가 아니라고 판정된 경우, 즉 제한 시간이 되어 버렸을 경우, 오브젝트마다의 부가적인 양자화 처리는 중단되고, 그 후, 처리는 스텝 S54로 진행된다.

[0134] 즉, 이 경우, 일부 오브젝트에 대해서는, 필요 최소한의 양자화 처리는 완료되었지만, 부가적인 양자화 처리에

대해서는 미완료인 채의 상태에서 처리가 중단된다. 그 때문에, 일부 오브젝트에 대해서는, 필요 최소한의 양자화 처리의 결과가, 최종적인 양자화 MDCT 계수로서 출력되게 된다.

[0135] 그러나, 스텝 S49 내지 스텝 S53에서는, 우선도가 높은 것부터 차례로 처리가 행해지기 때문에, 처리가 중단된 오브젝트는, 비교적 우선도가 낮은 오브젝트로 되어 있다. 즉, 우선도가 높은 오브젝트에 대해서는, 고품질의 양자화 MDCT 계수가 얻어졌으므로, 음질의 열화를 최소한으로 억제할 수 있다.

[0136] 또한, 스텝 S49에서 처리 대상 ID의 값이 N 미만이 아니라고 판정된 경우, 즉, 모든 오브젝트에 대해서, 제한 시간 내에 부가적인 양자화 처리가 완료된 경우, 처리는 스텝 S54로 진행된다.

[0137] 스텝 S47의 처리가 행해졌는데, 스텝 S49에서 처리 대상 ID의 값이 N 미만이 아니라고 판정되었거나, 또는 스텝 S51에서 제한 시간 이내가 아니라고 판정된 경우, 스텝 S54의 처리가 행해진다.

[0138] 스텝 S54에서 비트 얼로케이션부(54)는, 오브젝트마다 양자화 결과로서 보유하고 있는 양자화 MDCT 계수, 즉 보존 완료된 양자화 MDCT 계수를 부호화부(55)에 출력한다.

[0139] 이때, 필요 최소한의 양자화 처리가 미완료인 오브젝트에 대해서는, 양자화 결과로서 보유하고 있는 Mute 데이터의 양자화 값이 부호화부(55)에 출력되게 된다.

[0140] 또한, 비트 얼로케이션부(54)는, 각 오브젝트의 Mute 정보를 패킹부(23)에 공급하고, 비트 얼로케이션 처리는 종료된다.

[0141] 패킹부(23)에 대해서 Mute 정보가 공급되면, 상술한 도 3의 스텝 S17에서는, 패킹부(23)에 의해 Mute 정보가 부호화 비트 스트림에 저장된다.

[0142] Mute 정보는, 「0」 또는 「1」을 값으로서 갖는 플래그 정보 등으로 된다.

[0143] 구체적으로는, 예를 들어 오브젝트의 부호화 대상으로 되어 있는 프레임에서의 모든 양자화 MDCT 계수가 0일 경우, 즉 양자화 결과가 Mute 데이터일 경우에는, Mute 정보의 값은 「1」로 된다. 이에 반해, 양자화 결과가 Mute 데이터가 아닐 경우에는, Mute 정보의 값은 「0」으로 된다.

[0144] 이러한 Mute 정보는, 예를 들어 오브젝트의 메타데이터나, 부호화 비트 스트림의 부가 영역 등에 기술된다. 또한, Mute 정보는, 플래그 정보에 한정되지 않고, 알파벳이나 기타 기호, 「MUTE」 등의 문자열을 갖는 것이어도 된다.

[0145] 일례로서, MPEG-H의 ObjectMetadataConfig()에 Mute 정보가 추가된 신택스 예를 도 5에 나타낸다.

[0146] 도 5의 예에서는, 메타데이터의 Config에 있어서 오브젝트의 수(num_objects)만큼, Mute 정보 「mutedObjectFlag[o]」이 저장되어 있다.

[0147] 상술한 바와 같이, 오브젝트의 양자화 MDCT 계수가 모두 「0」일 경우에는 Mute 정보(mutedObjectFlag[o])로서 「1」이 세트되고, 그 이외의 경우에는 「0」이 세트된다.

[0148] 이러한 Mute 정보를 기술해 둠으로써, 복호측에서는 Mute 정보가 「1」인 오브젝트에 대해서는, IMDCT(Inverse Modified Discrete Cosine Transform)를 행하는 대신에 0 데이터(제로 데이터)를 IMDCT 출력으로서 사용할 수 있다. 이에 의해, 복호 처리의 고속화를 실현할 수 있다.

[0149] 이상과 같이 해서 비트 얼로케이션부(54)는, 우선도가 높은 오브젝트부터 차례로, 필요 최소한의 양자화 처리나 부가적인 양자화 처리를 행해 나간다.

[0150] 이와 같이 함으로써, 우선도가 높은 오브젝트일수록, 부가적인 양자화 처리(부가적인 비트 얼로케이션 루프 처리)를 완료시킬 수 있게 되어, 실시간 처리에서도 콘텐츠 전체의 부호화 효율을 향상시킬 수 있다. 이에 의해, 보다 많은 오브젝트의 데이터를 전송할 수 있다.

[0151] 또한, 이상에서는, 우선도 정보를 비트 얼로케이션부(54)에 입력하고, 시간 주파수 변환부(52)에서는, 모든 오브젝트에 대해서 시간 주파수 변환을 행하는 경우에 대해서 설명하였지만, 예를 들어 시간 주파수 변환부(52)에도 우선도 정보가 공급되도록 해도 된다.

[0152] 그러한 경우, 시간 주파수 변환부(52)는, 우선도 정보에 의해 나타내지는 우선도가 낮은 오브젝트에 대해서는 시간 주파수 변환을 행하지 않고, 각 스케일 팩터 밴드의 MDCT 계수를 모두 0 데이터(제로 데이터)로 바꾸어서 비트 얼로케이션부(54)에 공급한다.

[0153] 이와 같이 함으로써, 도 2에 도시한 구성에서의 경우와 비교하여, 우선도가 낮은 오브젝트의 처리 시간이나 처리량을 더욱 삭감하고, 우선도가 높은 오브젝트에 보다 많은 처리 시간을 확보할 수 있다.

[0154] <디코더의 구성예>

[0155] 계속해서, 도 1에 도시한 인코더(11)로부터 출력된 부호화 비트 스트림을 수신(취득)하여, 부호화 메타데이터나 부호화 오디오 신호를 복호하는 디코더에 대해서 설명한다.

[0156] 그러한 디코더는, 예를 들어 도 6에 도시하는 바와 같이 구성된다.

[0157] 도 6에 도시하는 디코더(81)는, 언패킹/복호부(91), 렌더링부(92) 및 믹싱부(93)를 갖고 있다.

[0158] 언패킹/복호부(91)는, 인코더(11)로부터 출력된 부호화 비트 스트림을 취득함과 함께, 부호화 비트 스트림의 언패킹 및 복호를 행한다.

[0159] 언패킹/복호부(91)는, 언패킹 및 복호에 의해 얻어진 각 오브젝트의 오디오 신호와, 각 오브젝트의 메타데이터를 렌더링부(92)에 공급한다. 이때, 언패킹/복호부(91)는, 부호화 비트 스트림에 포함되어 있는 Mute 정보에 따라서 각 오브젝트의 부호화 오디오 신호의 복호를 행한다.

[0160] 렌더링부(92)는, 언패킹/복호부(91)로부터 공급된 각 오브젝트의 오디오 신호 및 각 오브젝트의 메타데이터에 포함되어 있는 오브젝트 위치 정보에 기초하여 M 채널의 오디오 신호를 생성하여, 믹싱부(93)에 공급한다. 이때 렌더링부(92)는, 각 오브젝트의 음상이, 그러한 오브젝트의 오브젝트 위치 정보에 의해 나타내지는 위치에 정위하도록 M개의 각 채널의 오디오 신호를 생성한다.

[0161] 믹싱부(93)는, 렌더링부(92)로부터 공급된 각 채널의 오디오 신호를, 외부의 각 채널에 대응하는 스피커에 공급하여, 음성을 재생시킨다.

[0162] 또한, 부호화 비트 스트림에 채널마다의 부호화된 오디오 신호가 포함되어 있는 경우에는, 믹싱부(93)는, 언패킹/복호부(91)로부터 공급된 각 채널의 오디오 신호와, 렌더링부(92)로부터 공급된 각 채널의 오디오 신호를 채널마다 가중치 가산을 행하여, 최종적인 각 채널의 오디오 신호를 생성한다.

[0163] <언패킹/복호부의 구성예>

[0164] 또한, 도 6에 도시한 디코더(81)의 언패킹/복호부(91)는, 보다 상세하게는 예를 들어 도 7에 도시하는 바와 같이 구성된다.

[0165] 도 7에 도시하는 언패킹/복호부(91)는, Mute 정보 취득부(121), 오브젝트 오디오 신호 취득부(122), 오브젝트 오디오 신호 복호부(123), 출력 선택부(124), 0값 출력부(125) 및 IMDCT부(126)를 갖고 있다.

[0166] Mute 정보 취득부(121)는, 공급된 부호화 비트 스트림으로부터, 각 오브젝트의 오디오 신호의 Mute 정보를 취득해서 출력 선택부(124)에 공급한다.

[0167] 또한, Mute 정보 취득부(121)는, 공급된 부호화 비트 스트림으로부터 각 오브젝트의 부호화 메타데이터를 취득해서 복호하고, 그 결과 얻어진 메타데이터를 렌더링부(92)에 공급한다. 또한 Mute 정보 취득부(121)는, 공급된 부호화 비트 스트림을 오브젝트 오디오 신호 취득부(122)에 공급한다.

[0168] 오브젝트 오디오 신호 취득부(122)는, Mute 정보 취득부(121)로부터 공급된 부호화 비트 스트림으로부터 각 오브젝트의 부호화 오디오 신호를 취득하여, 오브젝트 오디오 신호 복호부(123)에 공급한다.

[0169] 오브젝트 오디오 신호 복호부(123)는, 오브젝트 오디오 신호 취득부(122)로부터 공급된 각 오브젝트의 부호화 오디오 신호를 복호하고, 그 결과 얻어진 MDCT 계수를 출력 선택부(124)에 공급한다.

[0170] 출력 선택부(124)는, Mute 정보 취득부(121)로부터 공급된 각 오브젝트의 Mute 정보에 기초하여, 오브젝트 오디오 신호 복호부(123)로부터 공급된 각 오브젝트의 MDCT 계수의 출력처를 선택적으로 전환한다.

[0171] 구체적으로는, 출력 선택부(124)는, 소정의 오브젝트에 관한 Mute 정보의 값이 「1」일 경우, 즉 양자화 결과가 Mute 데이터일 경우, 그 오브젝트의 MDCT 계수를 0으로 해서 0값 출력부(125)에 공급한다. 즉, 제로 데이터가 0값 출력부(125)에 공급된다.

[0172] 이에 반해, 출력 선택부(124)는, 소정의 오브젝트에 관한 Mute 정보의 값이 「0」일 경우, 즉 양자화 결과가 Mute 데이터가 아닐 경우, 오브젝트 오디오 신호 복호부(123)로부터 공급된, 그 오브젝트의 MDCT 계수를 IMDCT

부(126)에 공급한다.

[0173] 0값 출력부(125)는, 출력 선택부(124)로부터 공급된 MDCT 계수(제로 데이터)에 기초하여 오디오 신호를 생성하여, 렌더링부(92)에 공급한다. 이 경우, MDCT 계수는 0이므로, 무음의 오디오 신호가 생성된다.

[0174] IMDCT부(126)는, 출력 선택부(124)로부터 공급된 MDCT 계수에 기초하여 IMDCT를 행해서 오디오 신호를 생성하여, 렌더링부(92)에 공급한다.

[0175] <복호 처리의 설명>

[0176] 이어서, 디코더(81)의 동작에 대해서 설명한다.

[0177] 디코더(81)는, 인코더(11)로부터 1프레임분의 부호화 비트 스트림이 공급되면, 복호 처리를 행해서 오디오 신호를 생성하여, 스피커에 출력한다. 이하, 도 8의 흐름도를 참조하여, 디코더(81)에 의해 행해지는 복호 처리에 대해서 설명한다.

[0178] 스텝 S81에서, 언패킹/복호부(91)는, 인코더(11)로부터 송신되어 온 부호화 비트 스트림을 취득(수신)한다.

[0179] 스텝 S82에서, 언패킹/복호부(91)는 선택 복호 처리를 행한다.

[0180] 또한, 선택 복호 처리의 상세는 후술하지만, 선택 복호 처리에서는, 각 오브젝트의 부호화 오디오 신호가 Mute 정보에 기초하여 선택적으로 복호된다. 그리고 그 결과 얻어진 각 오브젝트의 오디오 신호가 렌더링부(92)에 공급된다. 또한, 부호화 비트 스트림으로부터 취득된 각 오브젝트의 메타데이터가 렌더링부(92)에 공급된다.

[0181] 스텝 S83에서, 렌더링부(92)는, 언패킹/복호부(91)로부터 공급된 각 오브젝트의 오디오 신호 및 각 오브젝트의 메타데이터에 포함되어 있는 오브젝트 위치 정보에 기초하여, 각 오브젝트의 오디오 신호의 렌더링을 행한다.

[0182] 예를 들어 렌더링부(92)는, 오브젝트 위치 정보에 기초하여 VBAP(Vector Base Amplitude Panning)에 의해, 각 오브젝트의 음상이 오브젝트 위치 정보에 의해 나타내지는 위치에 정위하도록 각 채널의 오디오 신호를 생성하여, 믹싱부(93)에 공급한다. 또한, 렌더링 방식은 VBAP에 한정되지 않고, 기타 형식을 사용해도 된다. 또한, 오브젝트의 위치 정보는 상술한 바와 같이, 예를 들어 수평 각도(Azimuth), 수직 각도(Elevation) 및 거리(Radius)를 포함하는데, 예를 들어 직교 좌표(X, Y, Z)에 의해 나타내져 있어도 된다.

[0183] 스텝 S84에서, 믹싱부(93)는, 렌더링부(92)로부터 공급된 각 채널의 오디오 신호를, 그러한 채널에 대응하는 스피커에 공급하여, 음성을 재생시킨다. 각 채널의 오디오 신호가 스피커에 공급되면, 복호 처리는 종료된다.

[0184] 이상과 같이 하여, 디코더(81)는, 부호화 비트 스트림으로부터 Mute 정보를 취득하고, 그 Mute 정보에 따라서 각 오브젝트의 부호화 오디오 신호를 복호한다.

[0185] <선택 복호 처리의 설명>

[0186] 계속해서, 도 9의 흐름도를 참조하여, 도 8의 스텝 S82의 처리에 대응하는 선택 복호 처리에 대해서 설명한다.

[0187] 스텝 S111에서, Mute 정보 취득부(121)는, 공급된 부호화 비트 스트림으로부터, 각 오브젝트의 오디오 신호의 Mute 정보를 취득해서 출력 선택부(124)에 공급한다.

[0188] 또한, Mute 정보 취득부(121)는, 부호화 비트 스트림으로부터 각 오브젝트의 부호화 메타데이터를 취득해서 복호하고, 그 결과 얻어진 메타데이터를 렌더링부(92)에 공급함과 함께, 부호화 비트 스트림을 오브젝트 오디오 신호 취득부(122)에 공급한다.

[0189] 스텝 S112에서, 오브젝트 오디오 신호 취득부(122)는, 처리 대상으로 하는 오브젝트의 오브젝트 번호에 0을 설정하고, 보유한다.

[0190] 스텝 S113에서, 오브젝트 오디오 신호 취득부(122)는, 보유하고 있는 오브젝트 번호가 오브젝트수 N 미만인지 여부를 판정한다.

[0191] 스텝 S113에서, 오브젝트 번호가 N 미만이라고 판정된 경우, 스텝 S114에서, 오브젝트 오디오 신호 복호부(123)는, 처리 대상의 오브젝트의 부호화 오디오 신호를 복호한다.

[0192] 즉, 오브젝트 오디오 신호 취득부(122)는, Mute 정보 취득부(121)로부터 공급된 부호화 비트 스트림으로부터, 처리 대상의 오브젝트의 부호화 오디오 신호를 취득해서 오브젝트 오디오 신호 복호부(123)에 공급한다.

[0193] 그러면, 오브젝트 오디오 신호 복호부(123)는, 오브젝트 오디오 신호 취득부(122)로부터 공급된 부호화 오디오

신호를 복호하고, 그 결과 얻어진 MDCT 계수를 출력 선택부(124)에 공급한다.

[0194] 스텝 S115에서, 출력 선택부(124)는, Mute 정보 취득부(121)로부터 공급된 처리 대상의 오브젝트의 Mute 정보의 값이 「0」인지 여부를 판정한다.

[0195] 스텝 S115에서, Mute 정보의 값이 「0」이라고 판정된 경우, 출력 선택부(124)는, 오브젝트 오디오 신호 복호부(123)로부터 공급된, 처리 대상의 오브젝트의 MDCT 계수를 IMDCT부(126)에 공급하고, 처리는 스텝 S116으로 진행된다.

[0196] 스텝 S116에서, IMDCT부(126)는, 출력 선택부(124)로부터 공급된 MDCT 계수에 기초해서 IMDCT를 행하여, 처리 대상의 오브젝트의 오디오 신호를 생성하고, 렌더링부(92)에 공급한다. 오디오 신호가 생성되면, 그 후, 처리는 스텝 S117로 진행된다.

[0197] 이에 반해, 스텝 S115에서, Mute 정보의 값이 「0」이 아닌, 즉 Mute 정보의 값이 「1」이라고 판정된 경우, 출력 선택부(124)는, MDCT 계수를 0으로 해서 0값 출력부(125)에 공급한다.

[0198] 0값 출력부(125)는, 출력 선택부(124)로부터 공급된 0인 MDCT 계수로부터, 처리 대상의 오브젝트의 오디오 신호를 생성하여, 렌더링부(92)에 공급한다. 따라서, 0값 출력부(125)에서는, 실질적으로는 IMDCT 등의 오디오 신호를 생성하기 위한 처리는 아무것도 행해지지 않는다.

[0199] 또한, 0값 출력부(125)에 의해 생성되는 오디오 신호는 무음 신호이다. 오디오 신호가 생성되면, 그 후, 처리는 스텝 S117로 진행된다.

[0200] 스텝 S115에서 Mute 정보의 값이 「0」이 아니라고 판정되었거나, 또는 스텝 S116에서 오디오 신호가 생성되면, 스텝 S117에서, 오브젝트 오디오 신호 취득부(122)는, 보유하고 있는 오브젝트 번호에 1을 더하여, 처리 대상의 오브젝트의 오브젝트 번호를 갱신한다.

[0201] 오브젝트 번호가 갱신되면, 그 후, 처리는 스텝 S113으로 돌아가서, 상술한 처리가 반복해서 행해진다. 즉, 새로운 처리 대상의 오브젝트의 오디오 신호가 생성된다.

[0202] 또한, 스텝 S113에서, 처리 대상의 오브젝트의 오브젝트 번호가 N 미만이 아니라고 판정된 경우, 모든 오브젝트에 대해서 오디오 신호가 얻어졌으므로 선택 복호 처리는 종료되고, 그 후, 처리는 도 8의 스텝 S83으로 진행된다.

[0203] 이상과 같이 하여, 디코더(81)는, 각 오브젝트에 대해서, Mute 정보에 기초하여, 처리 대상의 프레임의 오브젝트마다 부호화 오디오 신호의 복호를 행할지 여부를 판정하면서, 부호화 오디오 신호를 복호한다.

[0204] 즉, 디코더(81)에서는, 각 오디오 신호의 Mute 정보에 따라, 필요한 부호화 오디오 신호만이 복호된다. 이에 의해, 오디오 신호에 의해 재생되는 음성의 음질 열화를 최소한으로 억제하면서, 복호의 계산량을 저감시킬 수 있을 뿐만 아니라, 렌더링부(92) 등에서의 처리 등, 그 후의 처리의 계산량도 저감시킬 수 있다.

[0205] <제2 실시 형태>

[0206] <오브젝트 오디오 부호화부의 구성예>

[0207] 또한, 상술한 제1 실시 형태는, 고정 시점 3DAudio의 콘텐츠(오디오 신호)를 배신하는 예로 되어 있다. 이 경우, 유저의 청취 위치는 고정 위치가 된다.

[0208] 그런데, MPEG-I의 자유 시점 3DAudio에서는, 유저의 청취 위치는 고정 위치가 아니라, 유저는 임의의 위치로 이동할 수 있다. 그 때문에, 각 오브젝트의 우선도도 유저의 청취 위치와 오브젝트의 위치의 관계(위치 관계)에 따라서 변화하게 된다.

[0209] 그래서, 배신되는 콘텐츠(오디오 신호)가 자유 시점 3DAudio의 것인 경우, 오브젝트의 오디오 신호, 메타데이터의 Priority값, 오브젝트 위치 정보, 및 유저의 청취 위치를 나타내는 청취 위치 정보를 고려해서 우선도 정보가 생성되도록 해도 된다.

[0210] 그러한 경우, 인코더(11)의 오브젝트 오디오 부호화부(22)는, 예를 들어 도 10에 도시하는 바와 같이 구성된다. 또한, 도 10에서 도 2에서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있으며, 그 설명은 적절하게 생략한다.

[0211] 도 10에 도시하는 오브젝트 오디오 부호화부(22)는, 우선도 정보 생성부(51), 시간 주파수 변환부(52), 청각 심

리 파라미터 계산부(53), 비트 얼로케이션부(54) 및 부호화부(55)를 갖고 있다.

[0212] 도 10의 오브젝트 오디오 부호화부(22)의 구성은, 기본적으로는 도 2에 도시한 구성과 동일하지만, 우선도 정보 생성부(51)에 대해서, Priority값에 더해서 오브젝트 위치 정보 및 청취 위치 정보도 공급되는 점에서 도 2에 도시한 예와 다르게 되어 있다.

[0213] 즉, 도 10의 예에서는, 우선도 정보 생성부(51)에는, 각 오브젝트의 오디오 신호, 각 오브젝트의 메타데이터에 포함되어 있는 Priority값과 오브젝트 위치 정보, 및 3차원 공간에서의 유저의 청취 위치를 나타내는 청취 위치 정보가 공급된다.

[0214] 예를 들어 청취 위치 정보는, 인코더(11)에 의해, 콘텐츠의 배신처인 디코더(81)로부터 수신(취득)된다.

[0215] 또한, 여기서는 콘텐츠가 자유 시점 3DAudio의 것이기 때문에, 메타데이터에 포함되어 있는 오브젝트 위치 정보는, 예를 들어 3차원 공간에서의 음원 위치, 즉 오브젝트의 절대적인 위치를 나타내는 좌표 정보 등으로 된다. 또한, 이것에 한정되지 않고, 오브젝트 위치 정보는, 오브젝트의 상대적인 위치를 나타내는 좌표 정보이어도 된다.

[0216] 우선도 정보 생성부(51)는, 각 오브젝트의 오디오 신호, 각 오브젝트의 Priority값 및 각 오브젝트의 오브젝트 위치 정보와 청취 위치 정보(메타데이터와 청취 위치 정보) 중 적어도 어느 하나에 기초하여 우선도 정보를 생성하여, 비트 얼로케이션부(54)에 공급한다.

[0217] 예를 들어, 오브젝트와 유저(청취자)의 거리가 가까울 때에 비하여, 오브젝트와 유저의 거리가 멀어질수록 오브젝트의 음량이 낮아지고, 그 오브젝트의 우선도도 낮아지는 경향이 있다.

[0218] 그래서, 예를 들어 우선도 정보 생성부(51)가 오브젝트의 오디오 신호와 Priority값에 기초해서 구한 우선도에 대해서, 오브젝트와 유저의 청취 위치의 거리가 멀어질수록 우선도가 낮아지는 저차의 비선형 함수를 사용하여 우선도를 조정하고, 조정 후의 우선도를 나타내는 우선도 정보를 최종적인 우선도 정보로 하도록 해도 된다. 이와 같이 함으로써, 보다 주관에 적합한 우선도 정보를 얻을 수 있다.

[0219] 오브젝트 오디오 부호화부(22)가 도 10에 도시한 구성으로 되는 경우에도, 인코더(11)에서는, 도 3을 참조하여 설명한 부호화 처리가 행해진다.

[0220] 단, 스텝 S12에서는, 필요에 따라 오브젝트 위치 정보와 청취 위치 정보도 사용되어 우선도 정보가 생성된다. 즉, 오디오 신호, Priority값 및 오브젝트 위치 정보와 청취 위치 정보의 적어도 어느 것에 기초하여 우선도 정보가 생성된다.

[0221] <제3 실시 형태>

[0222] <콘텐츠 배신 시스템의 구성예>

[0223] 그런데, 라이브나 콘서트의 라이브 배신에 있어서 제1 실시 형태와 같은 부호화 효율을 향상시키는 실시간 처리를 위한 제한 처리를 행하고 있어도, 인코더를 실현하는 하드웨어에 있어서 OS(Operating System)의 인터럽트 등으로 급격하게 처리 부하가 높아져 버리는 경우가 있다. 그러한 경우, 실시간 처리의 제한 시간 내에 처리가 완료되지 않는 오브젝트가 증가하여, 청감상의 위화감을 부여해 버리는 경우가 생각된다. 즉, 음질이 열화되어 버리는 경우가 있다.

[0224] 그래서, 그러한 청감상의 위화감의 발생, 즉 음질의 열화를 억제하기 위해서, 프리렌더링에 의해 오브젝트수가 다른 복수의 입력 데이터를 준비하여, 각각 별도의 하드웨어에 의해 입력 데이터의 인코드(부호화)를 행하도록 해도 된다.

[0225] 이 경우, 예를 들어 실시간 처리를 위한 제한 처리가 발생하지 않은 부호화 비트 스트림 중에서, 가장 오브젝트수가 많은 부호화 비트 스트림이 디코더(81)에 출력된다. 따라서, 복수의 하드웨어 중에 OS의 인터럽트 등에 의한 급격한 처리 부하 증가가 발생한 하드웨어가 있었던 경우에도, 청감상의 위화감의 발생을 억제할 수 있다.

[0226] 이와 같이, 미리 복수의 입력 데이터를 준비할 경우, 콘텐츠를 배신하는 콘텐츠 배신 시스템은, 예를 들어 도 11에 도시하는 바와 같이 구성된다.

[0227] 도 11에 도시하는 콘텐츠 배신 시스템은, 인코더(201-1) 내지 인코더(201-3) 및 출력부(202)를 갖고 있다.

[0228] 예를 들어 콘텐츠 배신 시스템에서는, 동일한 콘텐츠를 재생하기 위한 데이터로서, 서로 오브젝트수가 다른 3개

의 입력 데이터 D1 내지 입력 데이터 D3이 미리 준비되어 있다.

[0229] 여기서는, 입력 데이터 D1은, N개의 각 오브젝트의 오디오 신호 및 메타데이터를 포함하는 데이터이며, 예를 들어 입력 데이터 D1은, 프리렌더링이 행해지지 않은 오리지널 데이터 등으로 된다.

[0230] 또한, 입력 데이터 D2는, 입력 데이터 D1보다 적은 16개의 각 오브젝트의 오디오 신호 및 메타데이터를 포함하는 데이터이며, 예를 들어 입력 데이터 D2는, 입력 데이터 D1에 대해서 프리렌더링을 행함으로써 얻어진 데이터 등으로 된다.

[0231] 마찬가지로, 입력 데이터 D3은, 입력 데이터 D2보다 적은 10개의 각 오브젝트의 오디오 신호 및 메타데이터를 포함하는 데이터이며, 예를 들어 입력 데이터 D3은, 입력 데이터 D1에 대해서 프리렌더링을 행함으로써 얻어진 데이터 등으로 된다.

[0232] 이러한 입력 데이터 D1 내지 입력 데이터 D3의 어느 것을 사용하여 콘텐츠(오디오)의 재생을 행해도, 기본적으로는 동일한 소리가 재생된다.

[0233] 콘텐츠 배신 시스템에서는, 인코더(201-1)에 대해서 입력 데이터 D1이 공급(입력)되고, 인코더(201-2)에 대해서 입력 데이터 D2가 공급되고, 인코더(201-3)에 대해서 입력 데이터 D3이 공급된다.

[0234] 인코더(201-1) 내지 인코더(201-3)는, 서로 다른 컴퓨터 등의 하드웨어에 의해 실현된다. 바꾸어 말하면, 인코더(201-1) 내지 인코더(201-3)는, 서로 다른 OS에 의해 실현된다.

[0235] 인코더(201-1)는, 공급된 입력 데이터 D1에 대해서 부호화 처리를 행함으로써 부호화 비트 스트림을 생성하여, 출력부(202)에 공급한다.

[0236] 마찬가지로, 인코더(201-2)는, 공급된 입력 데이터 D2에 대해서 부호화 처리를 행함으로써 부호화 비트 스트림을 생성해서 출력부(202)에 공급하고, 인코더(201-3)는, 공급된 입력 데이터 D3에 대해서 부호화 처리를 행함으로써 부호화 비트 스트림을 생성해서 출력부(202)에 공급한다.

[0237] 또한, 이하, 인코더(201-1) 내지 인코더(201-3)를 특별히 구별하는 필요가 없을 경우, 단순히 인코더(201)라고도 칭하기로 한다.

[0238] 각 인코더(201)는, 예를 들어 도 1에 도시한 인코더(11)와 동일한 구성을 갖고 있으며, 도 3을 참조하여 설명한 부호화 처리를 행함으로써, 부호화 비트 스트림을 생성한다.

[0239] 또한, 여기서는 콘텐츠 배신 시스템에는, 3개의 인코더(201)가 마련되는 예에 대해서 설명하지만, 이에 한정하지 않고, 2개 또는 4개 이상의 인코더(201)가 마련되도록 해도 된다.

[0240] 출력부(202)는, 복수의 각 인코더(201)로부터 공급된 부호화 비트 스트림 중 1개를 선택하고, 그 선택한 부호화 비트 스트림을 디코더(81)에 송신한다.

[0241] 예를 들어 출력부(202)는, 복수의 부호화 비트 스트림 중에, 값이 「1」인 Mute 정보가 포함되어 있지 않은 부호화 비트 스트림이 있는지, 즉 모든 오브젝트의 Mute 정보의 값이 「0」인 부호화 비트 스트림이 있는지를 특정한다.

[0242] 그리고 출력부(202)는, 값이 「1」인 Mute 정보가 포함되어 있지 않은 부호화 비트 스트림이 있을 경우에는, 값이 「1」인 Mute 정보가 포함되어 있지 않은 부호화 비트 스트림 중, 가장 오브젝트수가 많은 것을 선택하여, 디코더(81)에 송신한다.

[0243] 또한, 값이 「1」인 Mute 정보가 포함되어 있지 않은 부호화 비트 스트림이 없을 경우, 예를 들어 출력부(202)는, 가장 오브젝트수가 많은 것이나, Mute 정보가 「0」인 오브젝트의 수가 가장 많은 것 등을 선택해서 디코더(81)에 송신한다.

[0244] 이와 같이, 복수의 부호화 비트 스트림 중의 1개를 선택해서 출력함으로써, 청감상의 위화감의 발생을 억제하고, 고품질의 오디오 재생을 실현할 수 있다.

[0245] 여기서, 도 12를 참조하여, 콘텐츠의 오리지널 데이터로서, N(단, N>16)개의 오브젝트의 메타데이터 및 오디오 신호를 포함하는 데이터가 준비되어 있을 경우에 있어서의, 입력 데이터 D1 내지 입력 데이터 D3의 구체적인 예에 대해서 설명한다.

[0246] 이 예에서는, 입력 데이터 D1 내지 입력 데이터 D3의 어느 것에서든, 원래의(오리지널의) 데이터는 동일한 것이

며, 그 데이터에서의 오브젝트수는 N개로 되어 있다.

[0247] 특히, 입력 데이터 D1은 오리지널 데이터 그 자체로 되어 있다.

[0248] 따라서, 입력 데이터 D1은, 오리지널의(원래의) N개의 오브젝트에 대한 메타데이터와 오디오 신호를 포함하는 데이터로 되어 있고, 입력 데이터 D1에는 프리렌더링에 의해 생성된 새로운 오브젝트의 메타데이터와 오디오 신호는 포함되어 있지 않다.

[0249] 또한, 입력 데이터 D2 및 입력 데이터 D3은, 오리지널 데이터에 대한 프리렌더링을 행함으로써 얻어진 데이터로 되어 있다.

[0250] 구체적으로는, 입력 데이터 D2는, 오리지널의 N개의 오브젝트 중 우선도가 높은 4개의 오브젝트의 메타데이터 및 오디오 신호와, 프리렌더링에 의해 생성된 새로운 12개의 오브젝트의 메타데이터 및 오디오 신호를 포함하는 데이터로 되어 있다.

[0251] 입력 데이터 D2에 포함되어 있는 오리지널이 아닌 12개의 오브젝트의 데이터는, 오리지널의 N개의 오브젝트 중, 입력 데이터 D2에 포함되어 있지 않은 (N-4)개의 오브젝트의 데이터에 기초하는 프리렌더링에 의해 생성된 것이다.

[0252] 또한, 입력 데이터 D2에서는, 4개의 오브젝트에 대해서는, 오리지널의 오브젝트의 메타데이터 및 오디오 신호가, 프리렌더링되지 않고, 그대로 입력 데이터 D2에 포함되어 있다.

[0253] 입력 데이터 D3은, 오리지널의 오브젝트의 데이터가 포함되어 있지 않은, 프리렌더링에 의해 생성된 새로운 10개의 오브젝트의 메타데이터 및 오디오 신호를 포함하는 데이터로 되어 있다.

[0254] 이들 10개의 오브젝트의 메타데이터 및 오디오 신호는, 오리지널의 N개의 오브젝트의 데이터에 기초하는 프리렌더링에 의해 생성된 것이다.

[0255] 이상과 같이, 오리지널의 오브젝트의 데이터에 기초하여 프리렌더링을 행하여, 새로운 오브젝트의 메타데이터 및 오디오 신호를 생성함으로써, 오브젝트수를 저감시킨 입력 데이터를 준비할 수 있다.

[0256] 또한, 여기서는 오리지널의 오브젝트의 데이터는, 입력 데이터 D1만으로 되어 있지만, OS의 인터럽트 등의 돌발성을 고려하여, 프리렌더링을 행하지 않은, 오리지널의 데이터를 복수의 입력 데이터로서 사용해도 된다. 즉, 예를 들어 입력 데이터 D1뿐만 아니라, 입력 데이터 D2도 오리지널 데이터로 해도 된다.

[0257] 그렇게 하면, 예를 들어 입력 데이터 D1을 입력으로 하는 인코더(201-1)에서, 돌발적으로 OS의 인터럽트 등이 발생하였다고 해도, 입력 데이터 D2를 입력으로 하는 인코더(201-2)에 있어서 OS의 인터럽트 등이 발생하지 않으면, 음질의 열화를 방지할 수 있다. 즉, 인코더(201-2)에서는, 값이 「1」인 Mute 정보가 포함되어 있지 않은 부호화 비트 스트림이 얻어질 가능성이 높다.

[0258] 그 밖에, 예를 들어 오리지널의 오브젝트의 데이터에 기초하는 프리렌더링에 의해, 도 12에 도시한 입력 데이터 D3보다, 더욱 오브젝트수가 적은 입력 데이터를 다수 준비하도록 해도 된다. 또한, 입력 데이터 D1, D2, D3의 각 오브젝트 신호(오디오 신호)와 오브젝트 메타데이터(메타데이터)의 개수는 유저측에서 설정되도록 해도 되고, 각 인코더(201)의 리소스 등에 따라서 동적으로 변경되는 것이어도 된다.

[0259] 이상과 같이, 제1 실시 형태 내지 제3 실시 형태에서 설명한 본 기술에 의하면, 실시간 처리에 있어서 모든 처리가 제한 시간 내에 완료되지 않는 경우에도, 오브젝트의 음성의 중요도가 높은 순으로 부호화 효율을 향상시키는 부가적인 비트 얼로케이션 처리를 행함으로써, 콘텐츠 전체의 부호화 효율을 향상시킬 수 있다.

[0260] <제4 실시 형태>

[0261] <언더플로우에 대해서>

[0262] 상술한 바와 같이, MPEG-H 3D Audio 규격 등에서 취급되는 3D Audio에서는, 음 소재(오브젝트)의 위치를 나타내는 수평 각도나 수직 각도, 거리, 오브젝트에 관한 게인 등과 같은 오브젝트마다의 메타데이터를 갖게 하여, 3차원적인 소리의 방향이나 거리, 확산 등을 재현할 수 있다.

[0263] 종래의 스테레오 재생에서는, 스튜디오에서 믹싱 엔지니어가 많은 음 소재로 구성되는 멀티트랙 데이터를 바탕으로, 믹스 다운이라고 불리는 개개의 음 소재를 좌우 채널에 패닝함으로써 스테레오의 오디오 신호가 얻어지고 있었다.

[0264] 이에 반해 3D Audio에서는, 오브젝트라고 불리는 개개의 음 소재가 3차원 공간 중에 배치되고, 그러한 오브젝트의 위치 정보가 상술한 메타데이터로서 기술된다. 그 때문에, 3D Audio에서는 믹스 다운되기 전의 다수의 오브젝트, 보다 상세하게는 오브젝트의 오브젝트 오디오 신호가 부호화되게 된다.

[0265] 그런데, 라이브 방송 등, 실시간으로 부호화를 행하는 경우, 다수의 오브젝트를 부호화할 때 송출 장치에 대해서 높은 처리 능력이 요구된다. 즉, 소정의 시간 이내에 1프레임의 데이터를 부호화할 수 없는 경우, 송출 장치에서 송출하는 데이터가 존재하지 않는 언더플로우의 상태로 되어 송출 처리가 파탄되게 된다.

[0266] 이러한 언더플로우를 피하기 위해서, 실시간성이 요구되는 부호화 장치에서는, 주로 많은 계산 자원을 필요로 하는 비트 얼로케이션이라고 불리는 처리에 관해서, 소정의 시간 내에 처리가 완료되도록 비트 얼로케이션 처리가 제어되고 있다.

[0267] 요즘의 부호화 장치에서는, 기술의 진화에 추종하고, 또한 비용을 삭감할 목적으로, 전용의 하드웨어를 사용한 부호화 장치가 아니라, PC(Personal Computer) 등의 범용 하드웨어에 Linux(등록 상표) 등의 OS(Operating System)를 탑재하고, 그런 상태에서 부호화 소프트웨어를 동작시키는 케이스가 많다.

[0268] 그러나, Linux(등록 상표) 등의 OS에서는, 다수의 부호화 이외의 시스템 처리가 병렬로 실행되고 있고, 또한 그 시스템 처리는 우선도가 높은 처리로서 실행되고 있기 때문에, 부호화 소프트웨어의 처리보다 우선되어 실행되는 경우가 많다. 이러한 케이스에서는, 최악의 경우, 부호화 시의 처리가 비트 얼로케이션 처리까지 도달하지 못해서 언더플로우가 되는 경우도 있다.

[0269] 이러한 언더플로우를 피하기 위해서, 출력하는 처리 데이터가 없을 경우, 무음 데이터(Mute 데이터)를 부호화해서 송출하는 방법이 취해지는 경우가 많다.

[0270] MPEG-D USAC나 MPEG-H 3D Audio 등의 부호화 규격에서는, 컨텍스트 베이스 산술 부호화 기술이 사용되고 있다.

[0271] 이 컨텍스트 베이스 산술 부호화 기술은, 전 프레임과 당해 프레임의 양자화 MDCT 계수가 컨텍스트로 되고, 그 컨텍스트에 의해 부호화하려고 하는 양자화 MDCT 계수의 출현 빈도 테이블이 자동적으로 선택되어 산술 부호화가 행해진다.

[0272] 여기서, 도 13을 참조하여 컨텍스트 베이스 산술 부호화에서의 컨텍스트의 계산 방법에 대해서 설명한다.

[0273] 또한, 도 13에서는, 세로 방향은 주파수를 나타내고 있고, 가로 방향은 시간, 즉 오브젝트 오디오 신호의 프레임을 나타내고 있다.

[0274] 또한, 각 사각형 또는 원은, 프레임마다의 각 주파수의 MDCT 계수 블록을 나타내고 있고, 각 MDCT 계수 블록에는, 2개의 MDCT 계수(양자화 MDCT 계수)가 포함되어 있다. 특히, 각 사각형은 부호화 완료된 MDCT 계수 블록을 나타내고 있고, 각 원은, 아직 부호화가 행해지지 않은 MDCT 계수 블록을 나타내고 있다.

[0275] 이 예에서는, MDCT 계수 블록 BLK11이 부호화 대상으로 되어 있다. 이때, MDCT 계수 블록 BLK11에 인접하는 4개의 MDCT 계수 블록 BLK12 내지 MDCT 계수 블록 BLK15가 컨텍스트로 된다.

[0276] 특히, MDCT 계수 블록 BLK12 내지 MDCT 계수 블록 BLK14는, 부호화 대상의 MDCT 계수 블록 BLK11의 프레임의 시간적으로 직전의 프레임에서의, MDCT 계수 블록 BLK11의 주파수와 동일한 또는 인접하는 주파수의 MDCT 계수 블록이다.

[0277] 또한, MDCT 계수 블록 BLK15는, 부호화 대상의 MDCT 계수 블록 BLK11의 프레임에서의, MDCT 계수 블록 BLK11의 주파수에 인접하는 주파수의 MDCT 계수 블록이다.

[0278] 이들 MDCT 계수 블록 BLK12 내지 MDCT 계수 블록 BLK15에 기초하여 컨텍스트값이 계산되고, 그 컨텍스트값에 기초하여, 부호화 대상의 MDCT 계수 블록 BLK11을 부호화하기 위한 출현 빈도 테이블(산술 부호 빈도 테이블)이 선택된다.

[0279] 복호 시에도 산술 부호, 즉 부호화된 양자화 MDCT 계수로부터, 부호화 시와 동일한 출현 빈도 테이블이 사용되어 가변 길이 복호가 행해지지 않으면 안된다. 그 때문에, 컨텍스트값의 계산으로서, 부호화 시와 복호 시에는 완전히 동일한 계산이 행해지지 않으면 안된다.

[0280] 또한, 컨텍스트 베이스 산술 부호화의 더욱 상세한 내용에 대해서는, 본 기술과 직접적으로 관계되지 않기 때문에, 여기서는 그 설명은 생략한다.

[0281] 그런데, 이상에서 설명한 무음 데이터(Mute 데이터)를 부호화해서 송출하는 방법에서는, 무음 데이터 그 자체를 부호화하는 연산이 생기기 때문에, 결과적으로 소정 시간 내에 1프레임의 데이터를 출력할 수 없게 되는 경우가 있다.

[0282] 그래서 본 기술에서는, Linux(등록 상표) 등의 OS를 사용한 소프트웨어 베이스의 부호화 장치에 있어서, 부호화 방식이 컨텍스트 베이스 산술 부호화 기술을 사용하는 MPEG-H 등의 경우라도, 언더플로우의 발생을 방지할 수 있도록 하였다.

[0283] 특히, 본 기술에서는, 예를 들어 OS 상에서 발생하는 다른 처리 부하에 의해 부호화 처리가 완료되지 않는 경우라도, 미리 준비한 부호화 Mute 데이터를 송출함으로써 언더플로우의 발생을 방지할 수 있다.

[0284] <인코더의 구성예>

[0285] 도 14는, 본 기술을 적용한 인코더의 다른 실시 형태의 구성예를 도시하는 도면이다. 또한, 도 14에서 도 1에서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있으며, 그 설명은 적절하게 생략한다.

[0286] 도 14에 도시하는 인코더(11)는, 예를 들어 OS를 사용한 소프트웨어 베이스의 부호화 장치 등으로 된다. 즉, 예를 들어 인코더(11)는, PC 등의 정보 처리 장치에 있어서 OS가 부호화 소프트웨어를 동작시킴으로써 실현된다.

[0287] 인코더(11)는, 초기화부(301), 오브젝트 메타데이터 부호화부(21), 오브젝트 오디오 부호화부(22) 및 패킹부(23)를 갖고 있다.

[0288] 초기화부(301)는, OS 등으로부터 공급된 초기화 정보에 기초하여, 인코더(11)의 기동 시 등에 행해지는 초기화를 행함과 함께, 초기화 정보에 기초하여 부호화 Mute 데이터를 생성하여, 오브젝트 오디오 부호화부(22)에 공급한다.

[0289] 부호화 Mute 데이터는, Mute 데이터의 양자화 값, 즉 MDCT 계수 「0」의 양자화 MDCT 계수를 부호화함으로써 얻어지는 데이터이다. 이러한 부호화 Mute 데이터는, 무음 데이터의 MDCT 계수의 양자화 값, 즉 무음의 오디오 신호의 MDCT 계수의 양자화 값을 부호화해서 얻어지는 부호화 무음 데이터라고 할 수 있다. 또한, 이하에서는 부호화로서, 컨텍스트 베이스의 산술 부호화가 행해지는 것으로서 설명을 행하지만, 이에 한정하지 않고 다른 부호화 방식으로 부호화가 행해져도 된다.

[0290] 오브젝트 오디오 부호화부(22)는, 공급된 각 오브젝트의 오디오 신호(이하, 오브젝트 오디오 신호라고도 칭함)를 MPEG-H 규격에 따라서 부호화하고, 그 결과 얻어진 부호화 오디오 신호를 패킹부(23)에 공급한다. 이때 오브젝트 오디오 부호화부(22)는, 적절하게, 초기화부(301)로부터 공급된 부호화 Mute 데이터를, 부호화 오디오 신호로서 사용한다.

[0291] 또한, 상술한 실시 형태와 마찬가지로, 오브젝트 오디오 부호화부(22)에서 각 오브젝트의 메타데이터에 기초하여 우선도 정보가 산출되고, 그 우선도 정보가 사용되어 MDCT 계수의 양자화 등이 행해지도록 해도 된다.

[0292] <오브젝트 오디오 부호화부의 구성예>

[0293] 또한, 도 14에 도시한 인코더(11)의 오브젝트 오디오 부호화부(22)는, 예를 들어 도 15에 도시하는 바와 같이 구성된다. 또한, 도 15에서 도 2에서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있으며, 그 설명은 적절하게 생략한다.

[0294] 도 15의 예에서는, 오브젝트 오디오 부호화부(22)는, 시간 주파수 변환부(52), 청각 심리 파라미터 계산부(53), 비트 얼로케이션부(54), 컨텍스트 처리부(331), 가변 길이 부호화부(332), 출력 버퍼(333), 처리 진척 감시부(334), 처리 완료 가부 판정부(335) 및 부호화 Mute 데이터 삽입부(336)를 갖고 있다.

[0295] 비트 얼로케이션부(54)는, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수 및 청각 심리 파라미터 계산부(53)로부터 공급된 청각 심리 파라미터에 기초하여, 비트 얼로케이션 처리를 행한다. 또한, 상술한 실시 형태와 마찬가지로, 비트 얼로케이션부(54)가 우선도 정보에 기초하여 비트 얼로케이션 처리를 행하도록 해도 된다.

[0296] 비트 얼로케이션부(54)는, 비트 얼로케이션 처리에 의해 얻어진 각 오브젝트의 스케일 팩터 밴드마다의 양자화 MDCT 계수를, 컨텍스트 처리부(331) 및 가변 길이 부호화부(332)에 공급한다.

[0297] 컨텍스트 처리부(331)는, 비트 얼로케이션부(54)로부터 공급된 양자화 MDCT 계수에 기초하여, 양자화 MDCT 계수의 부호화를 행할 때 필요해지는 출현 빈도 테이블을 결정(선택)한다.

[0298] 예를 들어 컨텍스트 처리부(331)는, 도 13을 참조하여 설명한 바와 같이, 주목하는 양자화 MDCT 계수(MDCT 계수 블록) 근방의 복수의 양자화 MDCT 계수의 대푯값으로부터, 그 주목하는 양자화 MDCT 계수의 부호화에 사용하는 출현 빈도 테이블을 결정한다.

[0299] 컨텍스트 처리부(331)는, 양자화 MDCT 계수마다, 보다 상세하게는 MDCT 계수 블록마다 결정한, 각 양자화 MDCT 계수의 출현 빈도 테이블을 나타내는 인덱스(이하, 출현 빈도 테이블 인덱스라고도 칭함)를 가변 길이 부호화부(332)에 공급한다.

[0300] 가변 길이 부호화부(332)는, 컨텍스트 처리부(331)로부터 공급된 출현 빈도 테이블 인덱스에 의해 나타내지는 출현 빈도 테이블을 참조하여, 비트 얼로케이션부(54)로부터 공급된 양자화 MDCT 계수를 가변 길이 부호화하여, 가역 압축을 행한다.

[0301] 구체적으로는, 가변 길이 부호화부(332)는, 가변 길이 부호화로서 컨텍스트 베이스의 산술 부호화를 행함으로써, 부호화 오디오 신호를 생성한다.

[0302] 또한, 상술한 비특허문헌 1 내지 비특허 문헌 3에 기재되는 부호화 규격에서는, 가변 길이 부호화 기술로서 산술 부호화가 사용되고 있다. 본 기술에서는 산술 부호화 기술 이외에도 예를 들어 허프만 코딩화 기술 등, 다른 가변 길이 부호화 기술을 적용하는 것이 가능하다.

[0303] 가변 길이 부호화부(332)는, 가변 길이 부호화에 의해 얻어진 부호화 오디오 신호를 출력 버퍼(333)에 공급하여, 보유시킨다.

[0304] 양자화 MDCT 계수의 부호화를 행하는 컨텍스트 처리부(331) 및 가변 길이 부호화부(332)가, 도 2에 도시한 오브젝트 오디오 부호화부(22)의 부호화부(55)에 대응한다.

[0305] 출력 버퍼(333)는, 가변 길이 부호화부(332)로부터 공급된 프레임마다의 부호화 오디오 신호를 포함하는 비트 스트림을 보유하고, 적절한 타이밍에 보유하고 있는 부호화 오디오 신호(비트 스트림)를 패킹부(23)에 공급한다.

[0306] 처리 진척 감시부(334)는, 시간 주파수 변환부(52) 내지 비트 얼로케이션부(54), 컨텍스트 처리부(331) 및 가변 길이 부호화부(332)에서 행해지는 각 처리의 진척을 감시하고, 그 감시 결과를 나타내는 진척 정보를 처리 완료 가부 판정부(335)에 공급한다.

[0307] 처리 진척 감시부(334)는, 처리 완료 가부 판정부(335)로부터 공급되는 판정 결과에 따라, 적절하게, 시간 주파수 변환부(52) 내지 비트 얼로케이션부(54), 컨텍스트 처리부(331), 가변 길이 부호화부(332)에 대해서, 실행하고 있는 처리의 중단 등을 지시한다.

[0308] 처리 완료 가부 판정부(335)는, 처리 진척 감시부(334)로부터 공급된 진척 정보에 기초하여, 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내에 완료되는지 여부를 판정하는 처리 완료 가부 판정을 행하고, 그 판정 결과를 처리 진척 감시부(334) 및 부호화 Mute 데이터 삽입부(336)에 공급한다. 또한, 보다 상세하게는, 부호화 Mute 데이터 삽입부(336)에의 판정 결과의 공급은, 소정의 시간 내에 처리가 완료되지 않는다고 판정된 경우에만 행해진다.

[0309] 부호화 Mute 데이터 삽입부(336)는, 처리 완료 가부 판정부(335)로부터 공급된 판정 결과에 따라, 미리 준비(생성)된 부호화 Mute 데이터를 출력 버퍼(333)에서의, 각 프레임의 부호화 오디오 신호를 포함하는 비트 스트림에 삽입한다.

[0310] 이 경우, 부호화 Mute 데이터는, 소정의 시간 내에 처리가 완료되지 않는다고 판정된 프레임의 부호화 오디오 신호로서 비트 스트림에 삽입된다.

[0311] 즉, 소정의 프레임에서 시간 내에 처리가 완료되지 않는다고 판정된 경우, 비트 얼로케이션 처리가 중단되기 때문에, 그 소정의 프레임에서의 부호화 오디오 신호를 얻을 수 없다. 그 때문에, 출력 버퍼(333)에는 소정의 프레임에서의 부호화 오디오 신호가 보유되어 있지 않은 상태가 된다. 그래서 제로 데이터, 즉 무음의 오디오 신호(무음 신호)를 부호화해서 얻어지는 부호화 무음 데이터인 부호화 Mute 데이터가, 소정의 프레임의 부호화 오디오 신호로서 비트 스트림에 삽입된다.

[0312] 예를 들어 부호화 Mute 데이터의 삽입은, 오브젝트(오브젝트 오디오 신호)마다 행해져도 되고, 비트 얼로케이션 처리가 중단되었을 경우에는, 전체 오브젝트의 부호화 오디오 신호가 부호화 Mute 데이터로 되어도 된다.

[0313] <초기화부의 구성예>

[0314] 또한, 도 14에 도시한 인코더(11)의 초기화부(301)는, 예를 들어 도 16에 도시하는 바와 같이 구성된다.

[0315] 초기화부(301)는, 초기화 처리부(361) 및 부호화 Mute 데이터 생성부(362)를 갖고 있다.

[0316] 초기화 처리부(361)에는, 초기화 정보가 공급된다. 예를 들어 초기화 정보에는, 지금부터 부호화하려고 하는 콘텐츠를 구성하는 오브젝트나 채널의 수, 즉 오브젝트수나 채널수를 나타내는 정보가 포함되어 있다.

[0317] 초기화 처리부(361)는, 공급된 초기화 정보에 기초하여 초기화를 행함과 함께, 초기화 정보에 의해 나타내지는 오브젝트수, 보다 상세하게는 오브젝트수를 나타내는 오브젝트수 정보를 부호화 Mute 데이터 생성부(362)에 공급한다.

[0318] 부호화 Mute 데이터 생성부(362)는, 초기화 처리부(361)로부터 공급된 오브젝트수 정보에 의해 나타내지는 오브젝트의 수만큼 부호화 Mute 데이터를 생성하여, 부호화 Mute 데이터 삽입부(336)에 공급한다. 즉, 부호화 Mute 데이터 생성부(362)에서는, 오브젝트마다 부호화 Mute 데이터가 생성된다. 또한, 각 오브젝트의 부호화 Mute 데이터는, 동일한 데이터로 되어 있다.

[0319] 또한, 인코더(11)에 있어서 각 채널의 오디오 신호의 부호화도 행해지는 경우에는, 부호화 Mute 데이터 생성부(362)는, 채널수를 나타내는 채널수 정보에 기초하여, 채널수분의 부호화 Mute 데이터도 생성한다.

[0320] <처리의 진척과 부호화 Mute 데이터에 대해서>

[0321] 계속해서, 인코더(11)의 각 부에서 행해지는 처리의 진척과 부호화 Mute 데이터에 대해서 설명한다.

[0322] 처리 진척 감시부(334)는, 프로세서나 OS로부터 공급되는 타이머에 의해 시각을 특정하고, 1프레임분의 오브젝트 오디오 신호가 입력되고 나서, 그 프레임의 부호화 오디오 신호가 생성될 때까지의 처리의 진척 정도를 나타내는 진척 정보를 생성한다.

[0323] 여기서, 도 17을 참조하여 진척 정보와 처리 완료 가부 판정의 구체적인 예에 대해서 설명한다. 또한, 도 17에서는 1프레임분의 오브젝트 오디오 신호가 1024 샘플을 포함하는 것으로 되어 있다.

[0324] 도 17에 도시하는 예에서는, 시각 t11은, 처리 대상이 되는 프레임의 오브젝트 오디오 신호가 시간 주파수 변환부(52)에 공급된 시각, 즉 처리 대상의 오브젝트 오디오 신호에 대한 시간 주파수 변환이 개시되는 시각을 나타내고 있다.

[0325] 또한, 시각 t12는 소정의 역치가 되는 시각이며, 시각 t12까지 오브젝트 오디오 신호의 양자화, 즉 양자화 MDCT 계수의 생성이 완료되었으면, 처리 대상의 프레임의 부호화 오디오 신호를 지연 없이 출력(송출)할 수 있다. 바꾸어 말하면, 시각 t12까지 양자화 MDCT 계수를 생성하는 처리가 완료되었으면, 언더플로우는 생기지 않는다.

[0326] 시각 t13은, 처리 대상의 프레임의 부호화 오디오 신호, 즉 부호화 비트 스트림의 출력을 개시하는 시각이다. 이 예에서는, 시각 t11부터 시각 t13까지의 시간이 21msec으로 되어 있다.

[0327] 또한, 해치(사선)가 실시된 직사각형의 부분은, 오브젝트 오디오 신호로부터 양자화 MDCT 계수를 얻을 때까지 행해지는 처리 중, 오브젝트 오디오 신호에 구애되지 않고, 필요한 계산량(연산량)이 대략 일정한 처리(이하, 불변 처리라고도 칭함)를 행하는데 필요해지는 시간을 나타내고 있다. 보다 상세하게는 해치가 실시된 직사각형의 부분은, 불변 처리가 완료될 때까지 필요해지는 시간을 나타내고 있다. 예를 들어, 시간 주파수 변환이나 청각 심리 파라미터의 계산이 불변 처리이다.

[0328] 이에 반해, 해치가 실시되어 있지 않은 직사각형의 부분은, 오브젝트 오디오 신호로부터 양자화 MDCT 계수를 얻을 때까지 행해지는 처리 중, 오브젝트 오디오 신호에 따라 필요한 계산량, 즉 처리 시간이 변화하는 처리(이하, 가변 처리라고도 칭함)를 행하는데 필요해지는 시간을 나타내고 있다. 예를 들어, 비트 얼로케이션 처리가 가변 처리이다.

[0329] 처리 진척 감시부(334)는, 시간 주파수 변환부(52) 내지 비트 얼로케이션부(54)에서의 처리의 진척 상황을 감시하거나, OS 등에서의 인터럽트 처리의 발생 상황을 감시하거나 함으로써, 불변 처리나 가변 처리가 완료될 때까지 필요한 시간을 특정한다. 또한, 불변 처리나 가변 처리가 완료될 때까지 필요한 시간은, OS에서의 인터럽트 처리의 발생 등에 따라 변화한다.

[0330] 예를 들어 처리 진척 감시부(334)는, 불변 처리가 완료될 때까지 필요한 시간과, 가변 처리가 완료될 때까지 필

요한 시간을 나타내는 정보를 진척 정보로서 생성하여, 처리 완료 가부 판정부(335)에 공급한다.

[0331] 예를 들어 화살표 Q11로 나타내는 예에서는, 불변 처리와 가변 처리가 역치로 되는 시각 t12까지 완료(종료)된다. 즉, 시각 t12까지 양자화 MDCT 계수를 얻을 수 있다.

[0332] 따라서 처리 완료 가부 판정부(335)는, 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내, 즉 부호화 오디오 신호의 출력을 개시해야 하는 시각까지 완료되는 취지의 판정 결과를 처리 진척 감시부(334)에 공급한다.

[0333] 또한, 예를 들어 화살표 Q12로 나타내는 예에서는, 불변 처리는 시각 t12까지 완료되지만, 가변 처리의 처리 시간이 길기 때문에, 가변 처리가 시각 t12까지 완료되지 않는다. 바꾸어 말하면, 가변 처리의 완료 시각이 시각 t12를 약간 지나 버린다.

[0334] 따라서, 처리 완료 가부 판정부(335)는, 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내에 완료되지 않는 취지의 판정 결과를 처리 진척 감시부(334)에 공급한다. 보다 상세하게는, 처리 완료 가부 판정부(335)는, 비트 얼로케이션 처리의 중단이 필요한 취지의 판정 결과를 처리 진척 감시부(334)에 공급한다.

[0335] 이 경우, 예를 들어 처리 진척 감시부(334)는, 처리 완료 가부 판정부(335)로부터 공급된 판정 결과에 따라서 비트 얼로케이션부(54)에, 비트 얼로케이션 처리, 보다 상세하게는 비트 얼로케이션 루프 처리의 중단을 지시한다.

[0336] 그러면, 비트 얼로케이션부(54)에서는, 비트 얼로케이션 루프 처리가 중단된다. 그러나, 비트 얼로케이션부(54)에서는, 적어도 필요 최소한의 양자화 처리는 행해지기 때문에, 품질의 저하는 생기지만 언더플로우를 생기게 하지 않고 양자화 MDCT 계수를 얻을 수 있다.

[0337] 또한, 예를 들어 화살표 Q13으로 나타내는 예에서는, OS에서의 인터럽트 처리가 생겼기 때문에, 불변 처리가 시각 t12까지 완료되지 않고, 언더플로우가 생겨 버린다.

[0338] 그래서 처리 완료 가부 판정부(335)는, 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내에 완료되지 않는 취지의 판정 결과를 처리 진척 감시부(334) 및 부호화 Mute 데이터 삽입부(336)에 공급한다. 보다 상세하게는, 처리 완료 가부 판정부(335)는, 부호화 Mute 데이터의 출력이 필요한 취지의 판정 결과를 처리 진척 감시부(334) 및 부호화 Mute 데이터 삽입부(336)에 공급한다.

[0339] 이 경우, 시간 주파수 변환부(52) 내지 가변 길이 부호화부(332)에서는, 행하고 있는 처리가 정지되고(중단되고), 부호화 Mute 데이터 삽입부(336)에 의한 부호화 Mute 데이터의 삽입이 행해진다.

[0340] 이어서, 부호화 Mute 데이터에 대해서 설명한다. 부호화 Mute 데이터의 설명을 함에 있어서, 먼저 부호화 오디오 신호에 대해서 설명한다.

[0341] 상술한 바와 같이, 가변 길이 부호화부(332)로부터 출력 버퍼(333)에는 프레임마다 부호화 오디오 신호가 공급되는데, 보다 상세하게는 부호화 오디오 신호를 포함하는 부호화 데이터가 공급된다. 또한, 여기서는, 예를 들어 MPEG-H 3D Audio 규격에 따라서 양자화 MDCT 계수의 가변 길이 부호화가 행해지는 것으로 한다.

[0342] 예를 들어 1프레임분의 부호화 데이터에는, 적어도 Indep 플래그(인디펜던시 플래그), 현 프레임의 부호화 오디오 신호(부호화된 양자화 MDCT 계수), 프리롤 프레임(PreRollFrame)에 관한 데이터의 유무를 나타내는 프리롤 프레임 플래그가 포함되어 있다.

[0343] Indep 플래그는, 현 프레임이 예측이나 차분이 사용되어 부호화되어 있는 프레임인지 여부를 나타내는 플래그 정보이다.

[0344] 예를 들어 Indep 플래그의 값 「1」, 즉 Indep=1은, 현 프레임이 예측이나 차분 등을 사용하지 않고 부호화되어 있는 프레임인 것을 나타내고 있다. 바꾸어 말하면, Indep=1은, 현 프레임의 부호화 오디오 신호가, 양자화 MDCT 계수의 절댓값, 즉 양자화 MDCT 계수를 그대로 부호화한 것인 것을 나타내고 있다.

[0345] 따라서, 디코더(81)측, 즉 재생 기기측에서는, 부호화 비트 스트림의 도중부터 재생을 행하는 경우, Indep=1의 프레임으로부터 처리(재생)를 개시하는 것이 가능하다. 바꾸어 말하면, Indep=1인 프레임은, 랜덤 액세스 가능한 프레임이다.

[0346] 이에 반해, Indep 플래그의 값 「0」, 즉 Indep=0은, 현 프레임이 예측이나 차분이 사용되어 부호화되어 있는 프레임인 것을 나타내고 있다. 바꾸어 말하면, Indep=0은, 현 프레임의 부호화 오디오 신호가, 현 프레임의 양

자화 MDCT 계수와, 현 프레임의 직전 프레임의 양자화 MDCT 계수의 차분값을 부호화한 것인 것을 나타내고 있다. 따라서, Indep=0인 프레임은, 랜덤 액세스를 할 수 없는, 즉 랜덤 액세스의 액세스처로 할 수 없는 프레임이다.

[0347] 또한, 프리롤 프레임 플래그는, 현 프레임의 부호화 데이터에, 프리롤 프레임의 부호화 오디오 신호가 포함되어 있는지 여부를 나타내는 플래그 정보이다.

[0348] 예를 들어 프리롤 프레임 플래그의 값이 「1」일 경우, 현 프레임의 부호화 데이터에 프리롤 프레임의 부호화 오디오 신호(부호화된 양자화 MDCT 계수)가 포함되어 있다.

[0349] 이 경우, 현 프레임의 부호화 데이터에는, Indep 플래그, 현 프레임의 부호화 오디오 신호, 프리롤 프레임 플래그 및 프리롤 프레임의 부호화 오디오 신호가 포함되어 있다.

[0350] 이에 반해, 프리롤 프레임 플래그의 값이 「0」일 경우, 현 프레임의 부호화 데이터에는, 프리롤 프레임의 부호화 오디오 신호는 포함되어 있지 않다.

[0351] 또한, 프리롤 프레임이란, 랜덤 액세스가 가능한 프레임, 즉 Indep=1인 프레임의 시간적으로 직전에 있는 프레임이다.

[0352] 여기서, 도 18을 참조하여, 복수의 각 프레임의 부호화 데이터(부호화 오디오 신호)를 포함하는 비트 스트림의 예에 대해서 설명한다.

[0353] 또한, 도 18에서 #x는, 오브젝트 오디오 신호의 프레임(시간 프레임)의 프레임 번호를 나타내고 있다. 또한, 문자 「Indep=1」이 기재되어 있지 않은 프레임은, Indep=0인 프레임으로 되어 있다.

[0354] 예를 들어 「#0」은 0오리진으로 0프레임째(0번째), 즉 선두의 프레임을 나타내고 있고, 「#25」는 25프레임째를 나타내고 있다. 이하에서는, 프레임 번호가 「#x」인 프레임을 프레임 #x라고도 기재하기로 한다.

[0355] 도 18에서는, 화살표 Q31로 나타내는 부분에는, 처리 완료 가부 판정부(335)에 의해, 처리가 소정의 시간 내에 완료된다고 판정된 경우에 행해지는, 통상의 부호화 프로세스에 의해 얻어지는 비트 스트림이 나타내져 있다.

[0356] 특히, 이 예에서는 화살표 W11에 의해 나타내지는 프레임 #0과, 화살표 W12에 의해 나타내지는 프레임 #25가 Indep=1인 프레임, 즉 랜덤 액세스 가능한 프레임으로 되어 있다.

[0357] 예를 들어 모든 프레임에서 Indep=1로 하면, 어느 프레임으로부터든 복호(재생)를 개시할 수 있지만, 부호화 효율이 현저하게 저하되어 버리기 때문에, 일반적으로는 수10프레임마다 Indep=1로서 부호화되어 있다. 그 때문에, 도 18에서는 25프레임마다 Indep=1로 하는 것으로서 설명을 행한다.

[0358] 또한, 프레임 #25의 부분에 기재되어 있는 문자 「PreRollFrame(=#24)」은, 프레임 #25에 대한 프리롤 프레임인 프레임 #24의 부호화 오디오 신호가, 프레임 #25의 부호화 데이터(비트 스트림)에 저장되어 있는 것을 나타내고 있다.

[0359] 예를 들어 프레임 #25부터 복호를 개시할 경우, MDCT의 성질상, 프레임 #25의 부호화 오디오 신호에는 신호(오브젝트 오디오 신호)의 기함수 성분밖에 포함되어 있지 않다. 그 때문에, 프레임 #25의 부호화 오디오 신호만을 사용하여 복호를 행하면, 프레임 #25를 완전한 데이터로서 재생할 수 없어 이음이 발생해 버린다.

[0360] 그래서, 그러한 이음의 발생을 방지하기 위해서, 프레임 #25의 부호화 데이터에는, 프리롤 프레임인 프레임 #24의 부호화 오디오 신호가 저장되어 있다.

[0361] 그리고 프레임 #25부터 복호를 개시할 경우에는, 프레임 #25의 부호화 데이터로부터 프레임 #24의 부호화 오디오 신호, 보다 상세하게는 부호화 오디오 신호의 우함수 성분이 추출되어(취출되어) 프레임 #25의 기함수 성분과 합성된다.

[0362] 이에 의해, 프레임 #25의 복호 결과로서, 완전한 오브젝트 오디오 신호를 얻을 수 있고, 재생 시의 이음의 발생을 방지할 수 있다.

[0363] 또한, 화살표 Q32로 나타내는 부분에는, 프레임 #24에 있어서, 처리 완료 가부 판정부(335)에 의해, 처리가 소정의 시간 내에 완료되지 않는다고 판정된 경우에 얻어지는 비트 스트림이 나타내져 있다. 즉, 화살표 Q32로 나타내는 부분에는, 프레임 #24에 있어서 부호화 Mute 데이터의 삽입이 행해지는 예가 나타내져 있다.

[0364] 또한, 이하, 부호화 Mute 데이터의 삽입이 행해지는 프레임을, 특히 뮤트 프레임이라고도 칭하기로 한다.

[0365] 이 예에서는, 화살표 W13에 의해 나타내지는 프레임 #24가 뮤트 프레임으로 되어 있고, 이 프레임 #24는, 랜덤 액세스 가능한 프레임 #25의 직전 프레임(프리롤 프레임)으로 되어 있다.

[0366] 뮤트 프레임인 프레임 #24에서는, 초기화 시에 오브젝트수에 기초하여 사전에 계산된 부호화 Mute 데이터가, 프레임 #24의 부호화 오디오 신호로서 비트 스트림에 삽입된다. 보다 상세하게는, 부호화 Mute 데이터를 포함하는 부호화 데이터가 비트 스트림에 삽입된다.

[0367] 부호화 Mute 데이터 생성부(362)에서는, 프레임 #24가 랜덤 액세스 가능한 프레임인, 즉 Indep=1인 것으로 해서 MDCT 계수 「0」의 양자화 MDCT 계수(Mute 데이터의 양자화 값)를 산술 부호화함으로써 부호화 Mute 데이터가 생성된다.

[0368] 특히, 부호화 Mute 데이터는, 처리 대상의 프레임에 상당하는 1프레임분의 양자화 MDCT 계수(무음 데이터)만이 사용되고, 그 처리 대상의 프레임의 직전 프레임에 상당하는 양자화 MDCT 계수는 사용되지 않고 생성된다. 즉, 부호화 Mute 데이터는, 직전 프레임과의 차분이나, 직전 프레임의 컨텍스트는 사용되지 않고 생성된다.

[0369] 이것은, 초기화 시, 즉 부호화 Mute 데이터의 생성 시에는, 프레임 #24의 직전 프레임 #23의 데이터(양자화 MDCT 계수)가 존재하지 않기 때문이다.

[0370] 이와 같이, 뮤트 프레임이 랜덤 액세스 가능한 프레임이 아닐 경우, 그 뮤트 프레임의 부호화 데이터로서, 값이 「1」인 Indep 플래그, 뮤트 프레임인 현 프레임의 부호화 오디오 신호로서의 부호화 Mute 데이터, 및 값이 「0」인 프리롤 프레임 플래그가 포함되는 부호화 데이터가 생성된다.

[0371] 이 경우, 뮤트 프레임에서는 Indep 플래그의 값이 「1」로 되어 있지만, 디코더(81)측에서는, 그 뮤트 프레임으로부터 복호가 개시되지 않도록 되어 있다.

[0372] 또한, 이 예에서는, 뮤트 프레임인 프레임 #24의 다음 프레임 #25가 랜덤 액세스 가능한 프레임, 즉 Indep=1인 프레임으로 되어 있다.

[0373] 따라서, 프레임 #25의 부호화 데이터에는, 프레임 #25의 프리롤 프레임인 프레임 #24의 부호화 Mute 데이터가, 프리롤 프레임의 부호화 오디오 신호로서 저장된다. 이 경우, 예를 들어 부호화 Mute 데이터 삽입부(336)가, 출력 버퍼(333)에 보유되어 있는 프레임 #25의 부호화 데이터에 프레임 #24의 부호화 Mute 데이터를 삽입(저장)한다.

[0374] 화살표 Q33으로 나타내는 부분에는, 랜덤 액세스가 가능한 프레임 #25가 뮤트 프레임으로 된 예가 나타내져 있다.

[0375] 뮤트 프레임인 프레임 #25에서는, 초기화 시에 오브젝트수에 기초하여 사전에 계산된 부호화 Mute 데이터를 포함하는 부호화 데이터가 비트 스트림에 삽입된다. 이 부호화 Mute 데이터는, 화살표 Q32로 나타낸 예와 마찬가지로, Indep=1인 것으로 해서 MDCT 계수 「0」의 양자화 MDCT 계수를 산술 부호화함으로써 얻어진 것이다.

[0376] 또한, 프레임 #25는 랜덤 액세스 가능한 프레임이기 때문에, 프레임 #25의 부호화 데이터에는, 프리롤 프레임의 부호화 오디오 신호도 저장된다. 이 경우, 부호화 Mute 데이터가 프리롤 프레임의 부호화 오디오 신호로 된다.

[0377] 따라서, 뮤트 프레임이 랜덤 액세스 가능한 프레임일 경우, 그 뮤트 프레임의 부호화 데이터로서, 값이 「1」인 Indep 플래그, 뮤트 프레임인 현 프레임의 부호화 오디오 신호로서의 부호화 Mute 데이터, 값이 「1」인 프리롤 프레임 플래그 및 프리롤 프레임의 부호화 오디오 신호로서의 부호화 Mute 데이터가 포함되는 부호화 데이터가 생성된다.

[0378] 이상과 같이, 부호화 Mute 데이터 삽입부(336)는, 뮤트 프레임으로 되는 현 프레임이 프리롤 프레임일지, 랜덤 액세스 가능한 프레임일지 등, 현 프레임의 종별에 따라, 부호화 Mute 데이터의 삽입을 행한다.

[0379] 본 기술에 의하면, Linux(등록 상표) 등의 OS를 사용한 소프트웨어 베이스의 부호화 장치에 있어서, 부호화 방식이 컨텍스트 베이스 산술 부호화 기술을 사용하는 MPEG-H 등의 경우라도, 언더플로우의 발생을 방지할 수 있다.

[0380] 특히, 본 기술에서는, 예를 들어 OS상에서 발생하는 다른 처리 부하에 의해 오브젝트 오디오 신호의 부호화가 완료되지 않는 경우라도 언더플로우의 발생을 방지할 수 있다.

[0381] <부호화 데이터의 구성예>

[0382] 계속해서, 부호화 오디오 신호가 저장되는 부호화 데이터의 구성예에 대해서 설명한다.

[0383] 도 19는, 부호화 데이터의 신택스 예를 나타내고 있다.

[0384] 이 예에서는, 「usacIndependencyFlag」는 Indep 플래그를 나타내고 있다.

[0385] 또한, 「mpegh3daSingleChannelElement(usacIndependencyFlag)」는, 오브젝트 오디오 신호, 보다 상세하게는 부호화 오디오 신호를 나타내고 있다. 이 부호화 오디오 신호는, 현 프레임의 데이터이다.

[0386] 또한 부호화 데이터에는, 「mpegh3daExtElement(usacIndependencyFlag)」에 의해 나타내지는 확장 데이터가 저장되어 있다.

[0387] 이 확장 데이터는, 예를 들어 도 20에 나타내는 구성으로 된다.

[0388] 도 20에 도시하는 예에서는, 확장 데이터에는, 적절하게 「usacExtElementSegmentData[i]」에 의해 나타내지는 세그먼트 데이터가 저장되어 있다.

[0389] 이 세그먼트 데이터에 저장되는 데이터와, 데이터가 저장되는 순번은, 예를 들어 도 21에 도시하는 바와 같이 config 데이터인 usacExtElementType에 의해 정해진다.

[0390] 도 21에 도시하는 예에서는, usacExtElementType이 「ID_EXT_ELE_AUDIOPREROLL」일 경우에는, 세그먼트 데이터에 「AudioPreRoll()」이 저장된다.

[0391] 이 「AudioPreRoll()」은, 예를 들어 도 22에 도시하는 구성의 데이터로 되어 있다.

[0392] 이 예에서는, 「numPreRollFrames」에 의해 나타내지는 수만큼 「AccessUnit()」에 의해 나타내지는 현 프레임보다 전의 프레임의 부호화 오디오 신호가 저장되어 있다.

[0393] 특히, 여기서는 「AccessUnit()」에 의해 나타내지는 1개의 부호화 오디오 신호가, 프리롤 프레임의 부호화 오디오 신호로 되어 있다. 또한, 「numPreRollFrames」에 의해 나타내지는 수를 증가시킴으로써, 시간적으로 더 앞(과거측)의 프레임의 부호화 오디오 신호를 저장할 수도 있게 되어 있다.

[0394] <초기화 처리의 설명>

[0395] 이어서 도 14에 도시한 인코더(11)의 동작에 대해서 설명한다.

[0396] 먼저, 도 23의 흐름도를 참조하여, 인코더(11)가 기동했을 때 등에 행해지는 초기화 처리에 대해서 설명한다.

[0397] 스텝 S201에서 초기화 처리부(361)는, 공급된 초기화 정보에 기초하여 초기화를 행한다. 예를 들어 초기화 처리부(361)는, 인코더(11)의 각 부에서 부호화 처리 시에 사용되는 파라미터를 리셋하거나, 출력 버퍼(333)를 리셋하거나 한다.

[0398] 또한, 초기화 처리부(361)는, 초기화 정보에 기초하여 오브젝트수 정보를 생성하여, 부호화 Mute 데이터 생성부(362)에 공급한다.

[0399] 스텝 S202에서 부호화 Mute 데이터 생성부(362)는, 초기화 처리부(361)로부터 공급된 오브젝트수 정보에 기초하여 부호화 Mute 데이터를 생성하여, 부호화 Mute 데이터 삽입부(336)에 공급한다.

[0400] 예를 들어 부호화 Mute 데이터 생성부(362)는, 도 18을 참조하여 설명한 바와 같이, Indep=1인 것으로 해서 MDCT 계수 「0」의 양자화 MDCT 계수를 산술 부호화함으로써, 부호화 Mute 데이터를 생성한다. 또한, 부호화 Mute 데이터는, 오브젝트수 정보에 의해 나타내지는 오브젝트의 수만큼 생성된다. 부호화 Mute 데이터가 생성되면, 초기화 처리는 종료된다.

[0401] 이상과 같이 해서 인코더(11)는 초기화를 행하여, 부호화 Mute 데이터를 생성한다. 부호화 전에 미리 부호화 Mute 데이터를 생성해 둠으로써, 오브젝트 오디오 신호의 부호화 시에는, 필요에 따라 부호화 Mute 데이터의 삽입을 행하여, 언더플로우의 발생을 방지할 수 있게 된다.

[0402] <부호화 처리의 설명>

[0403] 초기화 처리가 종료되면, 그 후, 인코더(11)는 임의의 타이밍에 부호화 처리와 부호화 Mute 데이터 삽입 처리를 병행해서 행한다. 먼저, 도 24의 흐름도를 참조하여, 인코더(11)에 의한 부호화 처리에 대해서 설명한다.

[0404] 또한, 스텝 S231 내지 스텝 S233의 처리는, 도 3의 스텝 S11, 스텝 S13 및 스텝 S14의 처리와 마찬가지이므로,

그 설명은 생략한다.

[0405] 스텝 S234에서 비트 얼로케이션부(54)는, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수 및 청각 심리 파라미터 계산부(53)로부터 공급된 청각 심리 파라미터에 기초하여 비트 얼로케이션 처리를 행한다.

[0406] 비트 얼로케이션 처리에서는, 임의의 순번으로 각 오브젝트에 대해서, 스케일 팩터 밴드마다의 MDCT 계수에 대해, 상술한 필요 최소한의 양자화 처리 및 부가적인 비트 얼로케이션 루프 처리가 행해진다.

[0407] 비트 얼로케이션부(54)는, 비트 얼로케이션 처리에 의해 얻어진 양자화 MDCT 계수를 컨텍스트 처리부(331) 및 가변 길이 부호화부(332)에 공급한다.

[0408] 스텝 S235에서 컨텍스트 처리부(331)는, 비트 얼로케이션부(54)로부터 공급된 양자화 MDCT 계수에 기초하여, 양자화 MDCT 계수의 부호화에 사용하는 출현 빈도 테이블을 선택한다.

[0409] 예를 들어 컨텍스트 처리부(331)는, 도 13을 참조하여 설명한 바와 같이, 현 프레임의 처리 대상의 양자화 MDCT 계수에 대해서, 현 프레임과, 그 현 프레임의 직전 프레임에서의, 처리 대상의 양자화 MDCT 계수의 주파수(스케일 팩터 밴드) 근방의 주파수의 양자화 MDCT 계수에 기초하여 컨텍스트값을 계산한다.

[0410] 그리고 컨텍스트 처리부(331)는, 컨텍스트값에 기초하여, 처리 대상의 양자화 MDCT 계수를 부호화하기 위한 출현 빈도 테이블을 선택하고, 그 선택 결과를 나타내는 출현 빈도 테이블 인덱스를 가변 길이 부호화부(332)에 공급한다.

[0411] 스텝 S236에서 가변 길이 부호화부(332)는, 컨텍스트 처리부(331)로부터 공급된 출현 빈도 테이블 인덱스에 의해 나타내지는 출현 빈도 테이블에 기초하여, 비트 얼로케이션부(54)로부터 공급된 양자화 MDCT 계수를 가변 길이 부호화한다.

[0412] 가변 길이 부호화부(332)는, 가변 길이 부호화에 의해 얻어진 부호화 오디오 신호, 보다 상세하게는 가변 길이 부호화에 의해 얻어진 현 프레임의 부호화 오디오 신호를 포함하는 부호화 데이터를 출력 버퍼(333)에 공급하여, 보유시킨다.

[0413] 즉, 가변 길이 부호화부(332)는, 도 18을 참조하여 설명한 바와 같이 적어도 Indep 플래그, 현 프레임의 부호화 오디오 신호, 및 프리롤 프레임 플래그를 포함하는 부호화 데이터를 생성하여, 출력 버퍼(333)에 보유시킨다. 상술한 바와 같이 부호화 데이터에는, 프리롤 프레임 플래그의 값에 따라, 적절하게 프리롤 프레임의 부호화 오디오 신호도 포함되어 있다.

[0414] 또한, 이상에서 설명한 스텝 S232 내지 스텝 S236의 각 처리는, 처리 완료 가부 판정부(335)에 의한 처리 완료 가부 판정의 결과에 따라, 오브젝트마다나 프레임마다 행해진다. 즉, 처리 완료 가부 판정의 결과에 따라, 그것들의 복수의 처리 중 일부 또는 전부의 처리가 실행되지 않거나, 처리의 실행이 도중에 정지되거나(중단되거나) 한다.

[0415] 또한, 후술하는 부호화 Mute 데이터 삽입 처리에 의해, 출력 버퍼(333)에 보유되어 있는 각 프레임의 오브젝트마다의 부호화 오디오 신호(부호화 데이터)를 포함하는 비트 스트림에 대해서, 적절하게 부호화 Mute 데이터가 삽입된다.

[0416] 출력 버퍼(333)는, 보유하고 있는 부호화 오디오 신호(부호화 데이터)를 적절한 타이밍에 패킹부(23)에 공급한다.

[0417] 출력 버퍼(333)로부터 패킹부(23)에 프레임마다 부호화 오디오 신호(부호화 데이터)가 공급되면, 그 후, 스텝 S237의 처리가 행해져서 부호화 처리는 종료되는데, 스텝 S237의 처리는 도 3의 스텝 S17의 처리와 마찬가지이므로, 그 설명은 생략한다. 또한, 보다 상세하게는, 스텝 S237에서는 부호화 메타 데이터와, 부호화 오디오 신호를 포함하는 부호화 데이터의 패킹이 행해지고, 그 결과 얻어진 부호화 비트 스트림이 출력된다.

[0418] 이상과 같이 해서 인코더(11)는, 가변 길이 부호화를 행하고, 그 결과 얻어진 부호화 오디오 신호와, 부호화 메타데이터를 패킹해서 부호화 비트 스트림을 출력한다. 이와 같이 함으로써, 효율적으로 오브젝트의 데이터를 전송할 수 있다.

[0419] <부호화 Mute 데이터 삽입 처리의 설명>

[0420] 이어서, 도 25의 흐름도를 참조하여, 인코더(11)에 있어서 부호화 처리와 동시에 행해지는 부호화 Mute 데이터 삽입 처리에 대해서 설명한다. 예를 들어 부호화 Mute 데이터 삽입 처리는, 오브젝트 오디오 신호의 프레임마

다 또는 오브젝트마다 행해진다.

[0421] 스텝 S251에서 처리 완료 가부 판정부(335)는 처리 완료 가부 판정을 행한다.

[0422] 예를 들어 상술한 부호화 처리가 개시되면, 처리 진척 감시부(334)는, 시간 주파수 변환부(52) 내지 비트 얼로케이션부(54), 컨텍스트 처리부(331) 및 가변 길이 부호화부(332)에서 행해지는 각 처리의 진척 감시를 개시하여, 진척 정보를 생성한다. 그리고 처리 진척 감시부(334)는, 생성한 진척 정보를 처리 완료 가부 판정부(335)에 공급한다.

[0423] 그러면, 처리 완료 가부 판정부(335)는, 처리 진척 감시부(334)로부터 공급된 진척 정보에 기초하여 처리 완료 가부 판정을 행하고, 그 판정 결과를 처리 진척 감시부(334)나 부호화 Mute 데이터 삽입부(336)에 공급한다.

[0424] 예를 들어, 비트 얼로케이션 처리로서 필요 최소한의 양자화 처리만을 행하였다고 해도, 패킹부(23)에서의 패킹을 개시해야 하는 시각까지, 가변 길이 부호화부(332)에서의 가변 길이 부호화가 완료되지 않는 경우에는, 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내에 완료되지 않는다고 판정된다. 그리고 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내에 완료되지 않는 취지의 판정 결과, 보다 상세하게는, 부호화 Mute 데이터의 출력이 필요한 취지의 판정 결과가 처리 진척 감시부(334) 및 부호화 Mute 데이터 삽입부(336)에 공급된다.

[0425] 또한, 예를 들어 비트 얼로케이션 처리에 있어서 필요 최소한의 양자화 처리만을 행하거나 또는 비트 얼로케이션 루프 처리를 도중에 중단하면, 패킹부(23)에서의 패킹을 개시해야 하는 시각까지, 가변 길이 부호화부(332)에서의 가변 길이 부호화를 완료시킬 수 있는 경우도 있다. 그러한 경우, 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내에 완료되지 않는다고 판정되지만, 그 판정 결과는 부호화 Mute 데이터 삽입부(336)에는 공급되지 않고, 처리 진척 감시부(334)에만 공급된다. 보다 상세하게는, 비트 얼로케이션 처리의 중단이 필요한 취지의 판정 결과가 처리 진척 감시부(334)에 공급된다.

[0426] 처리 진척 감시부(334)는, 처리 완료 가부 판정부(335)로부터 공급된 판정 결과에 따라, 적절하게 시간 주파수 변환부(52) 내지 비트 얼로케이션부(54)나 컨텍스트 처리부(331), 가변 길이 부호화부(332)에서 행해지는 처리의 실행을 제어한다.

[0427] 즉, 처리 진척 감시부(334)는, 예를 들어 도 17을 참조하여 설명한 바와 같이, 시간 주파수 변환부(52) 내지 가변 길이 부호화부(332)의 각 처리 블록에 대해, 적절하게, 지금부터 행하고자 하고 있는 처리의 실행 중지나, 실행 중인 처리의 중단 등을 지시한다.

[0428] 구체적으로는, 예를 들어 소정 프레임에 있어서 오브젝트 오디오 신호를 부호화하는 처리가 소정의 시간 내에 완료되지 않는 취지의 판정 결과, 보다 상세하게는, 부호화 Mute 데이터의 출력이 필요한 취지의 판정 결과가 처리 진척 감시부(334)에 공급되었다고 하자.

[0429] 그러한 경우, 처리 진척 감시부(334)는, 시간 주파수 변환부(52) 내지 가변 길이 부호화부(332)에 대해서, 그러한 시간 주파수 변환부(52) 내지 가변 길이 부호화부(332)에서 행해지는 소정 프레임에 관한 처리의 실행 중지 또는 실행 중인 처리의 중단을 지시한다. 그러면, 도 24를 참조하여 설명한 부호화 처리에서는, 스텝 S232 내지 스텝 S236의 처리가 중지되거나 또는 도중에 중단된다.

[0430] 그 때문에, 가변 길이 부호화부(332)에서 소정 프레임의 양자화 MDCT 계수의 가변 길이 부호화는 행해지지 않고, 가변 길이 부호화부(332)로부터 출력 버퍼(333)에는, 그 소정 프레임에 관한 부호화 오디오 신호(부호화 데이터)는 공급되지 않는다.

[0431] 또한, 예를 들어 소정 프레임에 있어서, 비트 얼로케이션 처리의 중단이 필요한 취지의 판정 결과가 처리 진척 감시부(334)에 공급되었다고 하자. 그러한 경우, 처리 진척 감시부(334)는, 비트 얼로케이션부(54)에 대해서, 필요 최소한의 양자화 처리만의 실행 또는 비트 얼로케이션 루프 처리의 중단을 지시한다.

[0432] 그러면, 도 24를 참조하여 설명한 부호화 처리에서는, 스텝 S234에서 처리 진척 감시부(334)의 지시에 따른 비트 얼로케이션 처리가 행해진다.

[0433] 스텝 S252에서 부호화 Mute 데이터 삽입부(336)는, 처리 완료 가부 판정부(335)로부터 공급된 판정 결과에 기초하여, 부호화 Mute 데이터를 삽입할지 여부, 바꾸어 말하면 처리 대상의 현 프레임이 뮤트 프레임인지 여부를 판정한다.

[0434] 예를 들어 스텝 S252에서는, 처리 완료 가부 판정의 결과로서, 오브젝트 오디오 신호를 부호화하는 처리가 소정

의 시간 내에 완료되지 않는 취지의 판정 결과, 보다 상세하게는 부호화 Mute 데이터의 출력이 필요한 취지의 판정 결과가 공급된 경우에, 부호화 Mute 데이터를 삽입한다고 판정된다.

[0435] 스텝 S252에서 부호화 Mute 데이터를 삽입하지 않는다고 판정된 경우, 스텝 S253의 처리는 행해지지 않고, 부호화 Mute 데이터 삽입 처리는 종료된다.

[0436] 예를 들어 비트 얼로케이션 처리의 중단이 필요한 취지의 판정 결과가 처리 진척 감시부(334)에 공급될 경우, 스텝 S252에서는 부호화 Mute 데이터를 삽입하지 않는다고 판정되므로, 부호화 Mute 데이터 삽입부(336)는 부호화 Mute 데이터의 삽입을 행하지 않는다.

[0437] 또한, 처리 대상의 현 프레임이 랜덤 액세스 가능한 프레임이며, 또한 현 프레임의 직전 프레임이 뮤트 프레임일 경우에는, 부호화 Mute 데이터 삽입부(336)는, 프리롤 프레임의 부호화 Mute 데이터의 삽입을 행한다.

[0438] 즉, 예를 들어 도 18의 화살표 Q32로 나타낸 바와 같이, 부호화 Mute 데이터 삽입부(336)는, 출력 버퍼(333)에 보유되어 있는 현 프레임의 부호화 데이터 내에, 프리롤 프레임의 부호화 오디오 신호로서 부호화 Mute 데이터를 삽입한다.

[0439] 스텝 S252에서 부호화 Mute 데이터를 삽입한다고 판정된 경우, 스텝 S253에서 부호화 Mute 데이터 삽입부(336)는, 처리 대상의 현 프레임의 종별에 따라, 현 프레임의 부호화 데이터에 부호화 Mute 데이터를 삽입한다.

[0440] 보다 상세하게는, 부호화 Mute 데이터 삽입부(336)는, 예를 들어 도 18을 참조하여 설명한 바와 같이, 값이 「1」인 Indep 플래그, 처리 대상의 현 프레임의 부호화 오디오 신호로서의 부호화 Mute 데이터 및 프리롤 프레임 플래그가 포함되는 현 프레임의 부호화 데이터를 생성한다.

[0441] 이때, 현 프레임이 랜덤 액세스 가능한 프레임일 경우에는, 부호화 Mute 데이터 삽입부(336)는, 처리 대상의 현 프레임의 부호화 데이터에, 프리롤 프레임의 부호화 오디오 신호로서의 부호화 Mute 데이터도 저장한다.

[0442] 그리고 부호화 Mute 데이터 삽입부(336)는, 출력 버퍼(333)에 보유되어 있는 각 프레임의 부호화 데이터를 포함하는 비트 스트림에서의, 현 프레임에 대응하는 부분에, 그 현 프레임의 부호화 데이터를 삽입한다.

[0443] 또한, 상술한 바와 같이 현 프레임이, 현 프레임의 다음(직후) 프레임의 프리롤 프레임일 경우에는, 다음 프레임의 부호화 데이터에는, 적절한 타이밍에 프리롤 프레임의 부호화 오디오 신호로서 부호화 Mute 데이터가 삽입된다.

[0444] 또한, 현 프레임이 뮤트 프레임일 경우, 가변 길이 부호화부(332)가, 부호화 오디오 신호가 저장되어 있지 않은 현 프레임의 부호화 데이터를 생성해서 출력 버퍼(333)에 공급하도록 해도 된다. 그러한 경우, 부호화 Mute 데이터 삽입부(336)는, 출력 버퍼(333)에 보유되어 있는 현 프레임의 부호화 데이터 내에, 현 프레임이나 프리롤 프레임의 부호화 오디오 신호로서 부호화 Mute 데이터를 삽입한다.

[0445] 출력 버퍼(333)에 보유되어 있는 비트 스트림에 부호화 Mute 데이터가 삽입되면, 부호화 Mute 데이터 삽입 처리는 종료된다.

[0446] 이상과 같이 해서 인코더(11)는, 적절하게 부호화 Mute 데이터를 삽입한다. 이와 같이 함으로써, 언더플로우의 발생을 방지할 수 있다.

[0447] 또한, 필요에 따라 부호화 Mute 데이터의 삽입이 행해지는 경우에도, 비트 얼로케이션부(54)에서 우선도 정보에 의해 나타내지는 순번으로 비트 얼로케이션 처리가 행해지도록 해도 된다. 그러한 경우, 비트 얼로케이션부(54)에서는, 도 4를 참조하여 설명한 비트 얼로케이션 처리와 마찬가지의 처리가 행해져서, 예를 들어 필요 최소한의 양자화 처리가 미완료인 오브젝트에 대해서 부호화 Mute 데이터의 삽입이 행해진다.

[0448] <디코더의 구성예>

[0449] 또한, 도 14에 도시한 인코더(11)에 의해 출력된 부호화 비트 스트림을 입력으로 하는 디코더(81)는, 예를 들어 도 6에 도시한 구성으로 된다.

[0450] 단, 디코더(81)에서의 언패킹/복호부(91)의 구성은, 예를 들어 도 26에 도시하는 구성으로 된다. 또한, 도 26에서 도 7에서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있으며, 그 설명은 적절하게 생략한다.

[0451] 도 26에 도시하는 언패킹/복호부(91)는, 오브젝트 오디오 신호 취득부(122), 오브젝트 오디오 신호 복호부(123) 및 IMDCT부(126)를 갖고 있다.

[0452] 오브젝트 오디오 신호 취득부(122)는, 공급된 부호화 비트 스트림으로부터 각 오브젝트의 부호화 오디오 신호(부호화 데이터)를 취득하여, 오브젝트 오디오 신호 복호부(123)에 공급한다.

[0453] 또한, 오브젝트 오디오 신호 취득부(122)는, 공급된 부호화 비트 스트림으로부터 각 오브젝트의 부호화 메타데이터를 취득해서 복호하고, 그 결과 얻어진 메타데이터를 렌더링부(92)에 공급한다.

[0454] <복호 처리의 설명>

[0455] 이어서, 디코더(81)의 동작에 대해서 설명한다. 즉, 이하, 도 27의 흐름도를 참조하여, 디코더(81)에 의해 행해지는 복호 처리에 대해서 설명한다.

[0456] 스텝 S271에서, 언패킹/복호부(91)는, 인코더(11)로부터 송신되어 온 부호화 비트 스트림을 취득(수신)한다.

[0457] 스텝 S272에서, 언패킹/복호부(91)는 부호화 비트 스트림을 복호한다.

[0458] 즉, 언패킹/복호부(91)의 오브젝트 오디오 신호 취득부(122)는, 부호화 비트 스트림으로부터 각 오브젝트의 부호화 메타데이터를 취득해서 복호하고, 그 결과 얻어진 메타데이터를 렌더링부(92)에 공급한다.

[0459] 또한, 오브젝트 오디오 신호 취득부(122)는, 부호화 비트 스트림으로부터 각 오브젝트의 부호화 오디오 신호(부호화 데이터)를 취득하여, 오브젝트 오디오 신호 복호부(123)에 공급한다.

[0460] 그러면, 오브젝트 오디오 신호 복호부(123)는, 오브젝트 오디오 신호 취득부(122)로부터 공급된 부호화 오디오 신호를 복호하고, 그 결과 얻어진 MDCT 계수를 IMDCT부(126)에 공급한다.

[0461] 스텝 S273에서 IMDCT부(126)는, 오브젝트 오디오 신호 복호부(123)로부터 공급된 MDCT 계수에 기초하여 IMDCT를 행하여, 각 오브젝트의 오디오 신호를 생성하고, 렌더링부(92)에 공급한다.

[0462] IMDCT가 행해지면, 그 후, 스텝 S274 및 스텝 S275의 처리가 행해져서 복호 처리는 종료되는데, 이들 처리는 도 8의 스텝 S83 및 스텝 S84의 처리와 마찬가지이므로, 그 설명은 생략한다.

[0463] 이상과 같이 하여, 디코더(81)는, 부호화 비트 스트림을 복호하여, 음성을 재생시킨다. 이와 같이 함으로써, 언더플로우를 생기게 하지 않고, 즉 음성을 도중에 끊어지게 하지 않고 재생을 행할 수 있다.

[0464] <제5 실시 형태>

[0465] <인코더의 구성예>

[0466] 그런데, 콘텐츠를 구성하는 오브젝트 중에는, 다른 오브젝트로부터 마스크되고 싶지 않은 중요한 오브젝트가 있다. 또한, 1개의 오브젝트이어도, 오브젝트의 오디오 신호에 포함되는 복수의 주파수 성분 중에, 다른 오브젝트로부터 마스크되고 싶지 않은 중요한 주파수 성분도 있다.

[0467] 그래서, 다른 오브젝트로부터 마스크되고 싶지 않은 오브젝트나 주파수에 대해서, 오브젝트의 3차원 공간 상에 있는 다른 모든 오브젝트로부터의 소리에 관한 청각 마스킹양, 즉 마스킹 역치(공간 마스킹 역치)의 허용되는 상한값(이하, 허용 마스킹 역치라고도 칭함)이 설정되도록 해도 된다.

[0468] 마스킹 역치란, 마스킹에 의해 들리지 않게 되는 음압의 경계의 역치이며, 그 역치보다 작은 소리는 청감상 지각되지 않게 된다. 또한, 이하에서는, 주파수 마스킹을 단순히 마스킹으로 설명하지만, 주파수 마스킹 대신에 템포럴 마스킹을 사용해도 되고, 주파수 마스킹과 템포럴 마스킹 양쪽을 사용할 수도 있다. 주파수 마스킹이란, 복수의 주파수의 소리가 동시에 재생되었을 때, 어떤 주파수의 소리가 다른 주파수의 소리를 마스크해서 들리기 어렵게 하는 현상이다. 템포럴 마스킹이란, 어떤 소리가 재생되었을 때, 시간적으로 그 전후에 재생된 소리를 마스크해서 들리기 어렵게 하는 현상이다.

[0469] 이러한 상한값(허용 마스킹 역치)을 나타내는 설정 정보가 설정되는 경우, 설정 정보는, 비트 얼로케이션 처리, 보다 상세하게는 청각 심리 파라미터의 계산에 이용할 수 있다.

[0470] 설정 정보는, 다른 오브젝트로부터 마스크되고 싶지 않은 중요한 오브젝트나 주파수의 마스킹 역치에 관한 정보이다. 예를 들어 설정 정보에는, 허용 마스킹 역치, 즉 상한값이 설정되어 있는 오브젝트(오디오 신호)를 나타내는 오브젝트 ID나, 상한값이 설정되어 있는 주파수를 나타내는 정보, 설정되어 있는 상한값(허용 마스킹 역치) 등을 나타내는 정보가 포함되어 있다. 즉, 예를 들어 설정 정보에서는, 각 오브젝트에 대해서, 주파수마다 상한값(허용 마스킹 역치)이 설정되어 있다.

[0471] 설정 정보를 이용함으로써, 콘텐츠 제작자가 중요시하는 오브젝트나 주파수에 우선적으로 비트를 할당해서 다른

오브젝트나 주파수보다 음질비를 높게 하여, 콘텐츠 전체의 음질을 개선하거나 부호화 효율을 개선하거나 할 수 있다.

[0472] 도 28은, 설정 정보를 이용하는 경우의 인코더(11)의 구성예를 도시하는 도면이다. 또한, 도 28에서 도 1에서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있으며, 그 설명은 적절하게 생략한다.

[0473] 도 28에 도시하는 인코더(11)는, 오브젝트 메타데이터 부호화부(21), 오브젝트 오디오 부호화부(22) 및 패킹부(23)를 갖고 있다.

[0474] 이 예에서는, 도 1에 도시한 예와는 달리, 오브젝트 오디오 부호화부(22)에는, 오브젝트의 메타데이터에 포함되어 있는 Priority값은 공급되지 않는다.

[0475] 오브젝트 오디오 부호화부(22)는, 공급된 설정 정보에 기초하여, 공급된 N개의 각 오브젝트의 오디오 신호를 MPEG-H 규격 등에 따라서 부호화하고, 그 결과 얻어진 부호화 오디오 신호를 패킹부(23)에 공급한다.

[0476] 또한, 설정 정보에 의해 나타내지는 상한값은, 유저에 의해 설정(입력)된 것이어도 되고, 오브젝트 오디오 부호화부(22)에 의해 오디오 신호에 기초하여 설정되는 것이어도 된다.

[0477] 구체적으로는, 예를 들어 오브젝트 오디오 부호화부(22)가, 각 오브젝트의 오디오 신호에 기초하여 악곡 해석 등을 행하고, 그 결과로서 얻어진 콘텐츠의 장르나 멜로디 등의 해석 결과에 기초하여, 상한값을 설정하도록 해도 된다.

[0478] 예를 들어, Vocal(보컬)의 오브젝트에 대해서, 해석 결과에 기초하여 Vocal의 중요한 주파수 대역이 자동적으로 판별되고, 그 판별 결과에 기초하여 상한값이 설정되도록 할 수 있다.

[0479] 또한, 설정 정보에 의해 나타내지는 상한값(허용 마스킹 역치)은, 1개의 오브젝트에 대해 전체 주파수에서 공통의 값이 설정되도록 해도 되고, 1개의 오브젝트에 대해 주파수마다 설정되도록 해도 된다. 그 밖에, 복수의 오브젝트에 대해, 전체 주파수에서 공통의 상한값이나 주파수마다의 상한값이 설정되도록 해도 된다.

[0480] <오브젝트 오디오 부호화부의 구성예>

[0481] 또한, 도 28에 도시한 인코더(11)의 오브젝트 오디오 부호화부(22)는, 예를 들어 도 29에 도시하는 바와 같이 구성된다. 또한, 도 29에서 도 2에서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있으며, 그 설명은 적절하게 생략한다.

[0482] 도 29에 도시하는 예에서는, 오브젝트 오디오 부호화부(22)는, 시간 주파수 변환부(52), 청각 심리 파라미터 계산부(53), 비트 얼로케이션부(54) 및 부호화부(55)를 갖고 있다.

[0483] 시간 주파수 변환부(52)는, 공급된 각 오브젝트의 오디오 신호에 대해서 MDCT를 사용한 시간 주파수 변환을 행하고, 그 결과 얻어진 MDCT 계수를 청각 심리 파라미터 계산부(53) 및 비트 얼로케이션부(54)에 공급한다.

[0484] 청각 심리 파라미터 계산부(53)는, 공급된 설정 정보와, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수에 기초하여 청각 심리 파라미터를 계산하여, 비트 얼로케이션부(54)에 공급한다.

[0485] 또한, 여기서는 청각 심리 파라미터 계산부(53)에서, 설정 정보와 MDCT 계수에 기초하여 청각 심리 파라미터가 산출되는 예에 대해서 설명하지만, 청각 심리 파라미터가 설정 정보와 오디오 신호에 기초하여 산출되도록 해도 된다.

[0486] 비트 얼로케이션부(54)는, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수 및 청각 심리 파라미터 계산부(53)로부터 공급된 청각 심리 파라미터에 기초하여, 비트 얼로케이션 처리를 행한다.

[0487] 비트 얼로케이션 처리에서는, 각 스케일 팩터 밴드의 양자화 비트와 양자화 노이즈의 계산 및 평가를 행하는, 청각 심리 모델에 기초한 비트 얼로케이션이 행해진다. 그리고 그 비트 얼로케이션의 결과에 기초하여 스케일 팩터 밴드마다 MDCT 계수가 양자화되어, 양자화 MDCT 계수가 얻어진다(생성된다).

[0488] 비트 얼로케이션부(54)는, 이와 같이 하여 얻어진 각 오브젝트의 스케일 팩터 밴드마다의 양자화 MDCT 계수를, 각 오브젝트의 양자화 결과, 보다 상세하게는 각 오브젝트의 MDCT 계수의 양자화 결과로서 부호화부(55)에 공급한다.

[0489] 이상과 같은 비트 얼로케이션 처리에 의해, MDCT 계수의 양자화에서 발생해 버리는 양자화 노이즈가 마스크되어 지각되지 않는 스케일 팩터 밴드의 양자화 비트의 일부가, 양자화 노이즈가 지각되기 쉬운 스케일 팩터 밴드에

할당된다.

[0490] 이때, 설정 정보에 따라, 중요한 오브젝트나 주파수(스케일 팩터 밴드)에 대해서 우선적으로 비트가 할당된다. 바꾸어 말하면, 상한값이 설정되어 있는 오브젝트나 주파수에 대해서, 상한값에 따라서 적절하게 비트가 할당된다.

[0491] 이에 의해, 전체 음질의 열화, 특히 유저(콘텐츠 제작자)가 중요하다고 생각하는 오브젝트나 주파수의 음질 열화를 억제하여, 효율적인 양자화를 행할 수 있다. 즉, 부호화 효율을 향상시킬 수 있다.

[0492] 특히, 양자화 MDCT 계수의 산출에 있어서는, 청각 심리 파라미터 계산부(53)에서, 설정 정보에 기초하여, 각 오브젝트에 대해서 주파수마다 마스킹 역치(청각 심리 파라미터)가 계산된다. 그리고 비트 얼로케이션부(54)에서의 비트 얼로케이션 처리 시에는, 양자화 노이즈가 마스킹 역치를 초과하지 않도록 양자화 비트의 할당이 행해진다.

[0493] 예를 들어 청각 심리 파라미터의 계산 시에는, 설정 정보에 의해 상한값이 설정된 주파수에 대해서는, 허용되는 양자화 노이즈가 작아지는 파라미터 조정이 행해져서, 청각 심리 파라미터가 산출된다.

[0494] 또한, 설정 정보에 의해 나타내지는 허용 마스킹 역치, 즉 상한값에 따라서 파라미터 조정의 조정량이 변화하도록 해도 된다. 이에 의해, 해당 주파수에 많이 비트를 할당하도록 할 수 있다.

[0495] 부호화부(55)는, 비트 얼로케이션부(54)로부터 공급된 각 오브젝트의 스케일 팩터 밴드마다의 양자화 MDCT 계수를 부호화하고, 그 결과 얻어진 부호화 오디오 신호를 패킹부(23)에 공급한다.

[0496] <부호화 처리의 설명>

[0497] 계속해서, 도 28에 도시한 구성의 인코더(11)의 동작에 대해서 설명한다. 즉, 이하, 도 30의 흐름도를 참조하여, 도 28에 도시한 인코더(11)에 의한 부호화 처리에 대해서 설명한다.

[0498] 또한, 스텝 S301의 처리는, 도 3의 스텝 S11의 처리와 마찬가지이므로, 그 설명은 생략한다.

[0499] 스텝 S302에서 청각 심리 파라미터 계산부(53)는, 설정 정보를 취득한다.

[0500] 스텝 S303에서 시간 주파수 변환부(52)는, 공급된 각 오브젝트의 오디오 신호에 대해서 MDCT를 사용한 시간 주파수 변환을 행하여, 스케일 팩터 밴드마다의 MDCT 계수를 생성한다. 시간 주파수 변환부(52)는, 생성한 MDCT 계수를 청각 심리 파라미터 계산부(53) 및 비트 얼로케이션부(54)에 공급한다.

[0501] 스텝 S304에서 청각 심리 파라미터 계산부(53)는, 스텝 S302에서 취득한 설정 정보와, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수에 기초하여 청각 심리 파라미터를 계산하여, 비트 얼로케이션부(54)에 공급한다.

[0502] 이때 청각 심리 파라미터 계산부(53)는, 설정 정보에 의해 나타내지는 오브젝트나 주파수(스케일 팩터 밴드)에 대해서는, 허용되는 양자화 노이즈가 작아지도록, 설정 정보에 의해 나타내지는 상한값에 기초하여 청각 심리 파라미터를 산출한다.

[0503] 스텝 S305에서 비트 얼로케이션부(54)는, 시간 주파수 변환부(52)로부터 공급된 MDCT 계수 및 청각 심리 파라미터 계산부(53)로부터 공급된 청각 심리 파라미터에 기초하여, 비트 얼로케이션 처리를 행한다.

[0504] 비트 얼로케이션부(54)는, 비트 얼로케이션 처리에 의해 얻어진 양자화 MDCT 계수를 부호화부(55)에 공급한다.

[0505] 스텝 S306에서 부호화부(55)는, 비트 얼로케이션부(54)로부터 공급된 양자화 MDCT 계수를 부호화하고, 그 결과 얻어진 부호화 오디오 신호를 패킹부(23)에 공급한다.

[0506] 예를 들어 부호화부(55)에서는, 양자화 MDCT 계수에 대해서 컨텍스트 베이스의 산술 부호화가 행해지고, 부호화된 양자화 MDCT 계수가 부호화 오디오 신호로서 패킹부(23)에 출력된다. 또한, 부호화 방식은 산술 부호화에 한정되지 않고, 허프만 코딩화 방식이나 기타 부호화 방식 등, 다른 어떤 부호화 방식이어도 된다.

[0507] 스텝 S307에서 패킹부(23)는, 오브젝트 메타데이터 부호화부(21)로부터 공급된 부호화 메타 데이터와, 부호화부(55)로부터 공급된 부호화 오디오 신호를 패킹 하고, 그 결과 얻어진 부호화 비트 스트림을 출력한다. 패킹에 의해 얻어진 부호화 비트 스트림이 출력되면, 부호화 처리는 종료된다.

[0508] 이상과 같이 해서 인코더(11)는, 설정 정보에 기초하여 청각 심리 파라미터를 계산하여, 비트 얼로케이션 처리를 행한다. 이와 같이 함으로써, 콘텐츠 제작자가 우선하고 싶은 오브젝트나 주파수 대역의 소리에 대한 비트 할당을 증가시킬 수 있어, 부호화 효율을 개선할 수 있다.

[0509] 또한, 이 실시 형태에서는, 비트 얼로케이션 처리에 우선도 정보가 사용되지 않는 예에 대해서 설명하였다. 그러나, 이에 한정하지 않고, 비트 얼로케이션 처리에 우선도 정보를 이용하는 경우에도, 청각 심리 파라미터의 계산에 설정 정보가 이용되도록 해도 된다. 그러한 경우, 도 2에 도시한 오브젝트 오디오 부호화부(22)의 청각 심리 파라미터 계산부(53)에 설정 정보가 공급되고, 설정 정보가 사용되어 청각 심리 파라미터가 계산된다. 그 밖에, 도 15에 도시한 오브젝트 오디오 부호화부(22)의 청각 심리 파라미터 계산부(53)에 설정 정보가 공급되어, 청각 심리 파라미터의 계산에 설정 정보가 사용되도록 해도 된다.

[0510] <컴퓨터의 구성예>

[0511] 그런데, 상술한 일련의 처리는, 하드웨어에 의해 실행할 수도 있고, 소프트웨어에 의해 실행할 수도 있다. 일련의 처리를 소프트웨어에 의해 실행하는 경우에는, 그 소프트웨어를 구성하는 프로그램이 컴퓨터에 인스톨된다. 여기서, 컴퓨터에는, 전용의 하드웨어에 내장되어 있는 컴퓨터나, 각종 프로그램을 인스톨함으로써 각종 기능을 실행하는 것이 가능한, 예를 들어 범용의 퍼스널 컴퓨터 등이 포함된다.

[0512] 도 31은, 상술한 일련의 처리를 프로그램에 의해 실행하는 컴퓨터의 하드웨어의 구성예를 도시하는 블록도이다.

[0513] 컴퓨터에 있어서, CPU(Central Processing Unit)(501), ROM(Read Only Memory)(502), RAM(Random Access Memory)(503)은, 버스(504)에 의해 서로 접속되어 있다.

[0514] 버스(504)에는 또한, 입출력 인터페이스(505)가 접속되어 있다. 입출력 인터페이스(505)에는, 입력부(506), 출력부(507), 기록부(508), 통신부(509) 및 드라이브(510)가 접속되어 있다.

[0515] 입력부(506)는, 키보드, 마우스, 마이크로폰, 촬상 소자 등을 포함한다. 출력부(507)는, 디스플레이, 스피커 등을 포함한다. 기록부(508)는, 하드 디스크나 불휘발성의 메모리 등을 포함한다. 통신부(509)는, 네트워크 인터페이스 등을 포함한다. 드라이브(510)는, 자기 디스크, 광 디스크, 광자기 디스크, 또는 반도체 메모리 등의 리무버블 기록 매체(511)를 구동한다.

[0516] 이상과 같이 구성되는 컴퓨터에서는, CPU(501)가, 예를 들어 기록부(508)에 기록되어 있는 프로그램을, 입출력 인터페이스(505) 및 버스(504)를 통해서 RAM(503)에 로드해서 실행함으로써, 상술한 일련의 처리가 행해진다.

[0517] 컴퓨터(CPU501)가 실행하는 프로그램은, 예를 들어 패키지 미디어 등으로서의 리무버블 기록 매체(511)에 기록해서 제공할 수 있다. 또한, 프로그램은, 로컬 에어리어 네트워크, 인터넷, 디지털 위성 방송과 같은, 유선 또는 무선의 전송 매체를 통해서 제공할 수 있다.

[0518] 컴퓨터에서는, 프로그램은, 리무버블 기록 매체(511)를 드라이브(510)에 장착함으로써, 입출력 인터페이스(505)를 통해서 기록부(508)에 인스톨할 수 있다. 또한, 프로그램은, 유선 또는 무선의 전송 매체를 통해서 통신부(509)에서 수신하고, 기록부(508)에 인스톨할 수 있다. 그 밖에 프로그램은, ROM(502)이나 기록부(508)에 미리 인스톨해 둘 수 있다.

[0519] 또한, 컴퓨터가 실행하는 프로그램은, 본 명세서에서 설명하는 순서를 따라 시계열로 처리가 행해지는 프로그램이어도 되고, 병렬로, 혹은 호출이 행해졌을 때 등의 필요한 타이밍에 처리가 행해지는 프로그램이어도 된다.

[0520] 또한, 본 기술의 실시 형태는, 상술한 실시 형태에 한정되는 것이 아니라, 본 기술의 요지를 일탈하지 않는 범위에서 다양한 변경이 가능하다. 예를 들어, 본 기술의 실시 형태로서, 우선도가 높은 오브젝트부터 순서대로 양자화 처리가 행해지는 예에 대해서 설명하였지만, 유스 케이스에 따라서 우선도가 낮은 오브젝트부터 양자화 처리가 행해지도록 해도 된다.

[0521] 예를 들어, 본 기술은, 1개의 기능을 네트워크를 통해서 복수의 장치에서 분담, 공동해서 처리하는 클라우드 컴퓨팅의 구성을 취할 수 있다.

[0522] 또한, 상술한 흐름도에서 설명한 각 스텝은, 1개의 장치에서 실행하는 것 외에, 복수의 장치에서 분담해서 실행할 수 있다.

[0523] 또한, 1개의 스텝에 복수의 처리가 포함되는 경우에는, 그 1개의 스텝에 포함되는 복수의 처리는, 1개의 장치에서 실행하는 것 외에, 복수의 장치에서 분담해서 실행할 수 있다.

[0524] 또한, 본 기술은, 이하가 구성으로 하는 것도 가능하다.

[0525] (1) 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하는 우선도 정보 생성부와,

[0526] 상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하는 시간 주파수 변환부와,

[0527] 복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는 비트 얼로케이션부

[0528] 를 구비하는 부호화 장치.

[0529] (2) 상기 비트 얼로케이션부는, 상기 복수의 상기 오디오 신호의 상기 MDCT 계수에 대해서 필요 최소한의 양자화 처리를 행함과 함께, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 필요 최소한의 양자화 처리의 결과에 기초하여 상기 MDCT 계수를 양자화하는 부가적인 양자화 처리를 행하는, (1)에 기재된 부호화 장치.

[0530] (3) 상기 비트 얼로케이션부는, 소정의 제한 시간 내에 모든 상기 오디오 신호에 대해서 상기 부가적인 양자화 처리를 행할 수 없었을 경우, 상기 부가적인 양자화 처리가 완료되지 않은 상기 오디오 신호의 양자화 결과로서, 상기 필요 최소한의 양자화 처리의 결과를 출력하는, (2)에 기재된 부호화 장치.

[0531] (4) 상기 비트 얼로케이션부는, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 필요 최소한의 양자화 처리를 행하는, (3)에 기재된 부호화 장치.

[0532] (5) 상기 비트 얼로케이션부는, 상기 제한 시간 내에 모든 상기 오디오 신호에 대해서 상기 필요 최소한의 양자화 처리를 행할 수 없었을 경우, 상기 필요 최소한의 양자화 처리가 완료되지 않은 상기 오디오 신호의 양자화 결과로서, 제로 데이터의 양자화 값을 출력하는, (4)에 기재된 부호화 장치.

[0533] (6) 상기 비트 얼로케이션부는, 상기 오디오 신호의 양자화 결과가 상기 제로 데이터의 양자화 값인지를 나타내는 뮤트 정보를 또한 출력하는, (5)에 기재된 부호화 장치.

[0534] (7) 상기 비트 얼로케이션부는, 상기 비트 얼로케이션부의 후단에 있어서 필요해지는 처리 시간에 기초하여 상기 제한 시간을 결정하는, (3) 내지 (6) 중 어느 한 항에 기재된 부호화 장치.

[0535] (8) 상기 비트 얼로케이션부는, 지금까지 행한 상기 필요 최소한의 양자화 처리의 결과, 또는 상기 부가적인 양자화 처리의 결과에 기초하여, 상기 제한 시간을 동적으로 변경하는, (7)에 기재된 부호화 장치.

[0536] (9) 상기 우선도 정보 생성부는, 상기 오디오 신호의 음압, 상기 오디오 신호의 스펙트럼 형상, 또는 복수의 상기 오디오 신호간의 상기 스펙트럼 형상의 상관에 기초하여, 상기 우선도 정보를 생성하는, (2) 내지 (8) 중 어느 한 항에 기재된 부호화 장치.

[0537] (10) 상기 메타데이터에는, 미리 생성된 상기 오디오 신호의 우선도를 나타내는 Priority값이 포함되어 있는, (2) 내지 (9) 중 어느 한 항에 기재된 부호화 장치.

[0538] (11) 상기 메타데이터에는, 상기 오디오 신호에 기초하는 소리의 음원 위치를 나타내는 위치 정보가 포함되어 있고,

[0539] 상기 우선도 정보 생성부는, 적어도 상기 위치 정보와, 유저의 청취 위치를 나타내는 청취 위치 정보에 기초하여 상기 우선도 정보를 생성하는, (2) 내지 (10) 중 어느 한 항에 기재된 부호화 장치.

[0540] (12) 상기 복수의 상기 오디오 신호에는, 오브젝트의 상기 오디오 신호 및 채널의 상기 오디오 신호 중 적어도 어느 한쪽이 포함되어 있는, (2) 내지 (11) 중 어느 한 항에 기재된 부호화 장치.

[0541] (13) 상기 오디오 신호에 기초하여 청각 심리 파라미터를 계산하는 청각 심리 파라미터 계산부를 더 구비하고,

[0542] 상기 비트 얼로케이션부는, 상기 청각 심리 파라미터에 기초하여, 상기 필요 최소한의 양자화 처리 및 상기 부가적인 양자화 처리를 행하는, (2) 내지 (12) 중 어느 한 항에 기재된 부호화 장치.

[0543] (14) 상기 비트 얼로케이션부로부터 출력된, 상기 오디오 신호의 양자화 결과를 부호화하는 부호화부를 더 구비하는, (2) 내지 (13) 중 어느 한 항에 기재된 부호화 장치.

[0544] (15) 상기 청각 심리 파라미터 계산부는, 상기 오디오 신호와, 상기 오디오 신호에 관한 마스킹 역치에 관한 설정 정보에 기초하여 상기 청각 심리 파라미터를 계산하는, (13)에 기재된 부호화 장치.

[0545] (16) 부호화 장치가,

[0546] 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를

나타내는 우선도 정보를 생성하고,

[0547] 상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

[0548] 복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는, 부호화 방법.

[0549] (17) 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여, 상기 오디오 신호의 우선도를 나타내는 우선도 정보를 생성하고,

[0550] 상기 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

[0551] 복수의 상기 오디오 신호에 대해서, 상기 우선도 정보에 의해 나타내지는 상기 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 상기 MDCT 계수의 양자화를 행하는 처리를 컴퓨터에 실행시키는 프로그램.

[0552] (18) 복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고, 상기 부호화 오디오 신호를 복호하는 복호부를 구비하는, 복호 장치.

[0553] (19) 상기 복호부는, 상기 오디오 신호의 양자화 결과가 제로 데이터의 양자화 값인지를 나타내는 뮤트 정보를 또한 취득하고, 상기 뮤트 정보에 따라, 상기 복호에 의해 얻어진 상기 MDCT 계수에 기초하여 상기 오디오 신호를 생성하거나 또는 상기 MDCT 계수를 0으로 해서 상기 오디오 신호를 생성하는, (18)에 기재된 복호 장치.

[0554] (20) 복호 장치가,

[0555] 복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고,

[0556] 상기 부호화 오디오 신호를 복호하는, 복호 방법.

[0557] (21) 복수의 오디오 신호에 대해서, 상기 오디오 신호, 및 상기 오디오 신호의 메타데이터 중 적어도 어느 것에 기초하여 생성된 우선도 정보에 의해 나타내지는 우선도가 높은 상기 오디오 신호로부터 차례로, 상기 오디오 신호의 MDCT 계수의 양자화를 행함으로써 얻어진 부호화 오디오 신호를 취득하고,

[0558] 상기 부호화 오디오 신호를 복호하는 처리를 컴퓨터에 실행시키는 프로그램.

[0559] (22) 오디오 신호를 부호화하여, 부호화 오디오 신호를 생성하는 부호화부와,

[0560] 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 보유하는 버퍼와,

[0561] 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는 삽입부

[0562] 를 구비하는 부호화 장치.

[0563] (23) 상기 오디오 신호의 MDCT 계수의 양자화를 행하는 비트 얼로케이션부를 더 구비하고, 상기 부호화부는, 상기 MDCT 계수의 양자화 결과를 부호화하는, (22)에 기재된 부호화 장치.

[0564] (24) 상기 부호화 무음 데이터를 생성하는 생성부를 더 구비하는, (23)에 기재된 부호화 장치.

[0565] (25) 상기 생성부는, 무음 데이터의 MDCT 계수의 양자화 값을 부호화함으로써 상기 부호화 무음 데이터를 생성하는, (24)에 기재된 부호화 장치.

[0566] (26) 상기 생성부는, 1프레임분의 상기 무음 데이터에만 기초하여 상기 부호화 무음 데이터를 생성하는, (24) 또는 (25)에 기재된 부호화 장치.

[0567] (27) 상기 오디오 신호는, 채널 또는 오브젝트의 오디오 신호이며,

[0568] 상기 생성부는, 채널수 및 오브젝트수의 적어도 어느 것에 기초하여, 상기 부호화 무음 데이터를 생성하는, (24) 내지 (26) 중 어느 한 항에 기재된 부호화 장치.

[0569] (28) 상기 삽입부는, 상기 처리 대상의 프레임의 종별에 따라서 상기 부호화 무음 데이터의 삽입을 행하는, (22) 내지 (27) 중 어느 한 항에 기재된 부호화 장치.

[0570] (29) 상기 삽입부는, 상기 처리 대상의 프레임이 랜덤 액세스 가능한 프레임의 프리롤 프레임일 경우, 상기 랜덤 액세스 가능한 프레임에 관한 상기 프리롤 프레임의 상기 부호화 오디오 신호로서 상기 부호화 무음 데이터를 상기 비트 스트림에 삽입하는, (28)에 기재된 부호화 장치.

[0571] (30) 상기 삽입부는, 상기 처리 대상의 프레임이 랜덤 액세스 가능한 프레임일 경우, 상기 처리 대상의 프레임에 관한 프리롤 프레임의 상기 부호화 오디오 신호로서 상기 부호화 무음 데이터를 상기 비트 스트림에 삽입하는, (28) 또는 (29)에 기재된 부호화 장치.

[0572] (31) 상기 삽입부는, 상기 비트 얼로케이션부에서, 상기 MDCT 계수에 대해서 필요 최소한의 양자화 처리만을 행하거나 또는 상기 MDCT 계수에 대해서 상기 필요 최소한의 양자화 처리 후에 행해지는 부가적인 양자화 처리를 도중에 중단하면, 상기 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료될 경우, 상기 부호화 무음 데이터의 삽입을 행하지 않는, (23) 내지 (27) 중 어느 한 항에 기재된 부호화 장치.

[0573] (32) 상기 부호화부는, 상기 오디오 신호에 대해서 가변 길이 부호화를 행하는, (22) 내지 (31) 중 어느 한 항에 기재된 부호화 장치.

[0574] (33) 상기 가변 길이 부호화는, 컨텍스트 베이스의 산술 부호화인, (32)에 기재된 부호화 장치.

[0575] (34) 부호화 장치가,

[0576] 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고,

[0577] 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 버퍼에 보유하고,

[0578] 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는, 부호화 방법.

[0579] (35) 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고,

[0580] 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림을 버퍼에 보유하고,

[0581] 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 상기 비트 스트림에 삽입하는 처리를 컴퓨터에 실행시키는 프로그램.

[0582] (36) 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는 복호부를 구비하는, 복호 장치.

[0583] (37) 복호 장치가,

[0584] 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는, 복호 방법.

[0585] (38) 오디오 신호를 부호화해서 부호화 오디오 신호를 생성하고, 처리 대상의 프레임에 대해서, 소정의 시간 내에 상기 오디오 신호를 부호화하는 처리가 완료되지 않는 경우, 프레임마다의 상기 부호화 오디오 신호를 포함하는 비트 스트림에 상기 처리 대상의 프레임의 상기 부호화 오디오 신호로서, 미리 생성된 부호화 무음 데이터를 삽입함으로써 얻어진 상기 비트 스트림을 취득하고, 상기 부호화 오디오 신호를 복호하는 처리를 컴퓨터에 실행시키는 프로그램.

[0586] (39) 오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하는 시간 주파수 변환부와,

[0587] 상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산

하는 청각 심리 파라미터 계산부와,

[0588] 상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는 비트 얼로케이션부

[0589] 를 구비하는 부호화 장치.

[0590] (40) 상기 설정 정보에는, 주파수마다 설정된 상기 마스킹 역치의 상한값을 나타내는 정보가 포함되어 있는, (39)에 기재된 부호화 장치.

[0591] (41) 상기 설정 정보에는, 1개 또는 복수의 상기 오브젝트마다 설정된 상기 마스킹 역치의 상한값을 나타내는 정보가 포함되어 있는, (39) 또는 (40)에 기재된 부호화 장치.

[0592] (42) 부호화 장치가,

[0593] 오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

[0594] 상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산하고,

[0595] 상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는, 부호화 방법.

[0596] (43) 오브젝트의 오디오 신호에 대한 시간 주파수 변환을 행하여, MDCT 계수를 생성하고,

[0597] 상기 MDCT 계수와, 상기 오브젝트에 관한 마스킹 역치에 관한 설정 정보에 기초하여 청각 심리 파라미터를 계산하고,

[0598] 상기 청각 심리 파라미터 및 상기 MDCT 계수에 기초하여 비트 얼로케이션 처리를 행하여, 양자화 MDCT 계수를 생성하는 스텝을 포함하는 처리를 컴퓨터에 실행시키는 프로그램.

부호의 설명

[0599] 11: 인코더 21: 오브젝트 메타데이터 부호화부

22: 오브젝트 오디오 부호화부 23: 패킹부

51: 우선도 정보 생성부 52: 시간 주파수 변환부

53: 청각 심리 파라미터 계산부 54: 비트 얼로케이션부

55: 부호화부 81: 디코더

91: 언패킹/복호부 92: 렌더링부

331: 컨텍스트 처리부 332: 가변 길이 부호화부

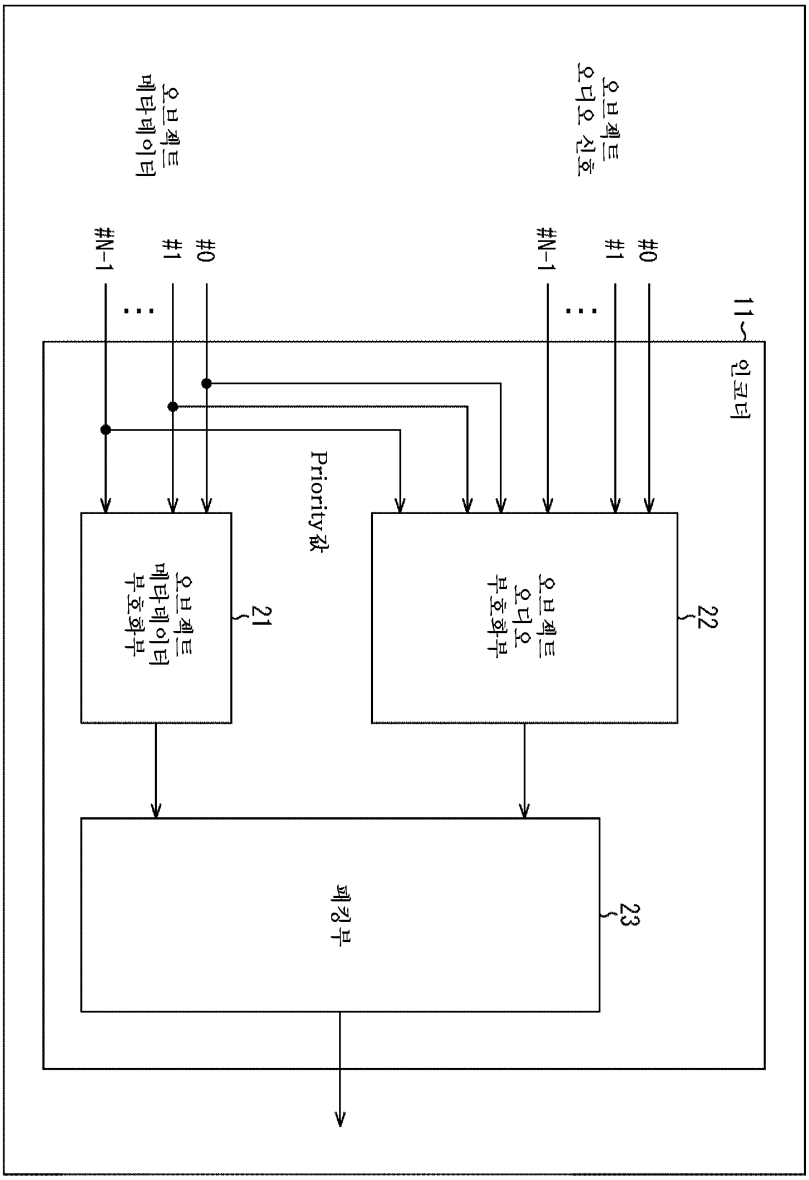
333: 출력 버퍼 334: 처리 진척 감시부

335: 처리 완료 가부 판정부 336: 부호화 Mute 데이터 삽입부

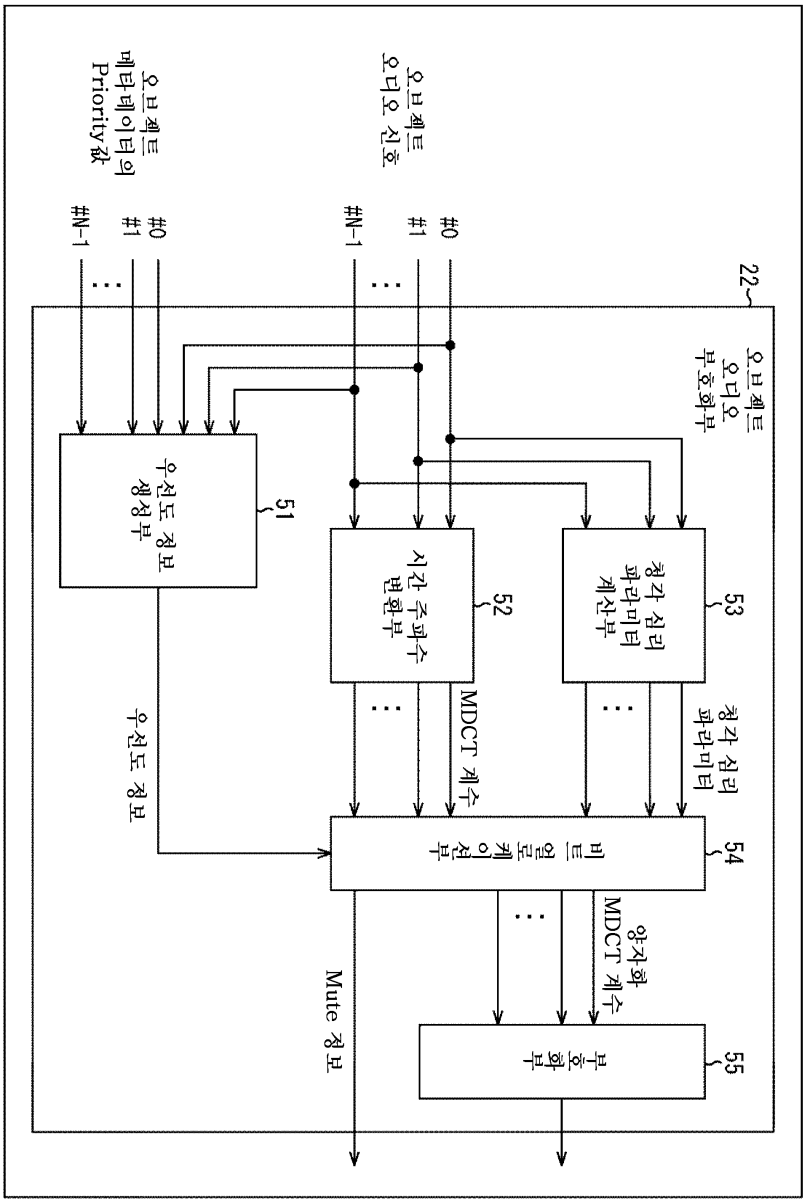
362: 부호화 Mute 데이터 생성부

도면

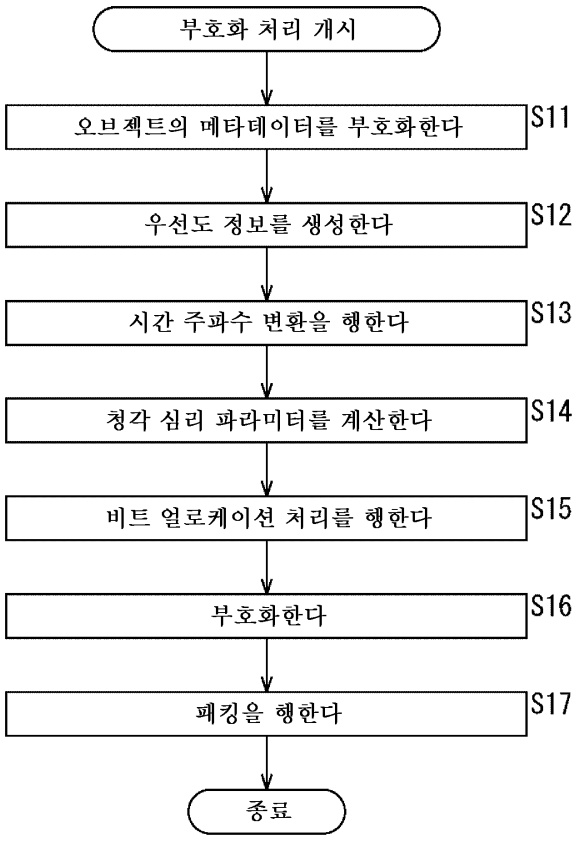
도면1


오브젝트 메타데이터
오브젝트 오디오 신호
#N-1
#1
#0
#N-1
#1
#0
11
인코더
Priority값
오브젝트 메타데이터 부호화부
21
오브젝트 오디오 부호화부
22
패킹부
23

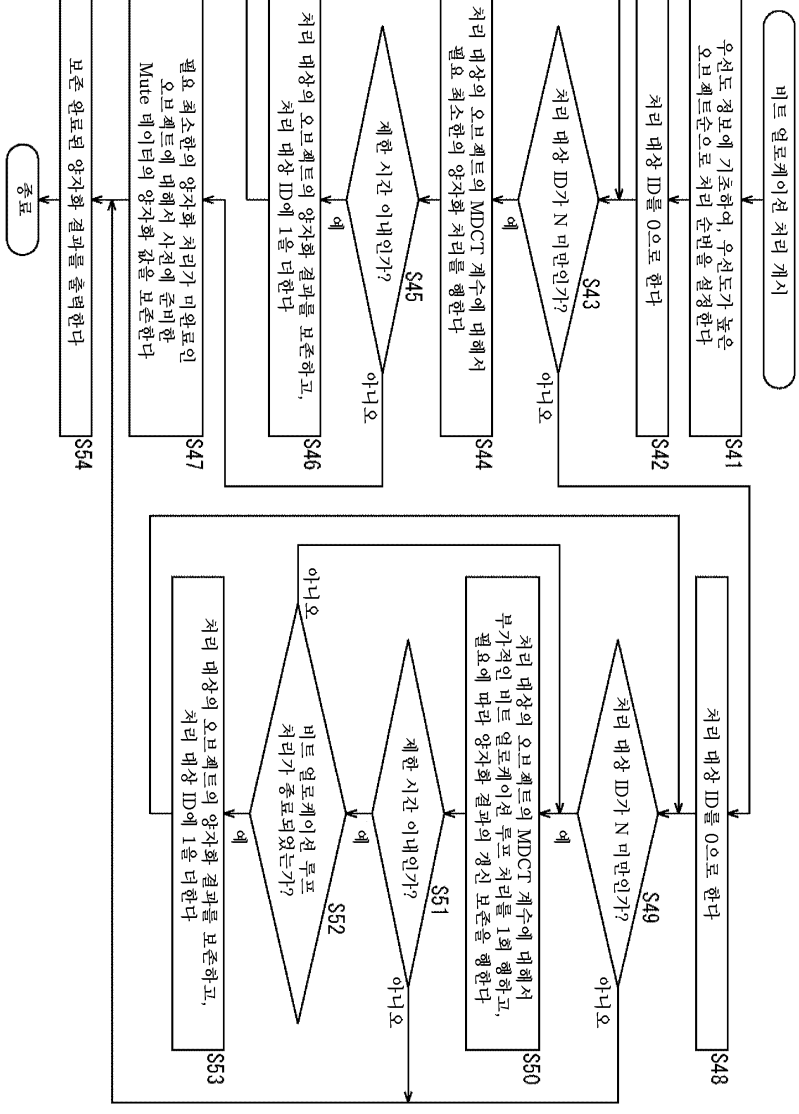
도면2


오브젝트 메타데이터의 Priority값
오브젝트 오디오 신호
#0
#1
#N-1
22
오브젝트 오디오 부호화부
우선도 정보 생성부
51
시간 주파수 변환부
52
청각 심리 파라미터 계산부
53
비트 할당부
54
부호화부
55
우선도 정보
MDCT 계수
청각 심리 파라미터
양자화 MDCT 계수
Mute 정보

도면3


부호화 처리 개시
오브젝트의 메타데이터를 부호화한다
S11
우선도 정보를 생성한다
S12
시간 주파수 변환을 행한다
S13
청각 심리 파라미터를 계산한다
S14
비트 얼로케이션 처리를 행한다
S15
부호화한다
S16
패킹을 행한다
S17
종료

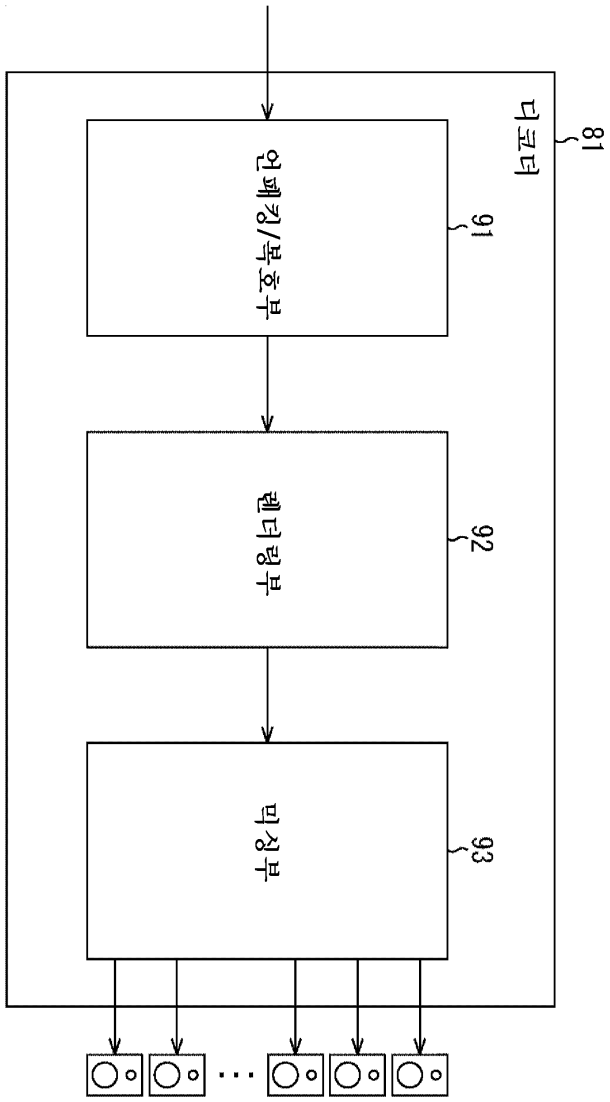
도면4


비트 얼로케이션 처리 개시
우선도 정보에 기초하여, 우선도가 높은 오브젝트순으로 처리 순번을 설정한다
S41
처리 대상 ID를 0으로 한다
S42
처리 대상 ID가 N 미만인가?
S43
아니오
예
처리 대상의 오브젝트의 MDCT 계수에 대해서 필요 최소한의 양자화 처리를 행한다
S44
제한 시간 이내인가?
S45
아니오
예
처리 대상의 오브젝트의 양자화 결과를 보존하고, 처리 대상 ID에 1을 더한다
S46
필요 최소한의 양자화 처리가 미완료인 오브젝트에 대해서 사전에 준비한 Mute 데이터의 양자화 값을 보존한다
S47
보존 완료된 양자화 결과를 출력한다
S54
종료
처리 대상 ID를 0으로 한다
S48
처리 대상 ID가 N 미만인가?
S49
아니오
예
처리 대상의 오브젝트의 MDCT 계수에 대해서 부가적인 비트 얼로케이션 루프 처리를 1회 행하고, 필요에 따라 양자화 결과의 갱신 보존을 행한다
S50
제한 시간 이내인가?
S51
아니오
예
비트 얼로케이션 루프 처리가 종료되었는가?
S52
아니오
예
처리 대상의 오브젝트의 양자화 결과를 보존하고, 처리 대상 ID에 1을 더한다
S53

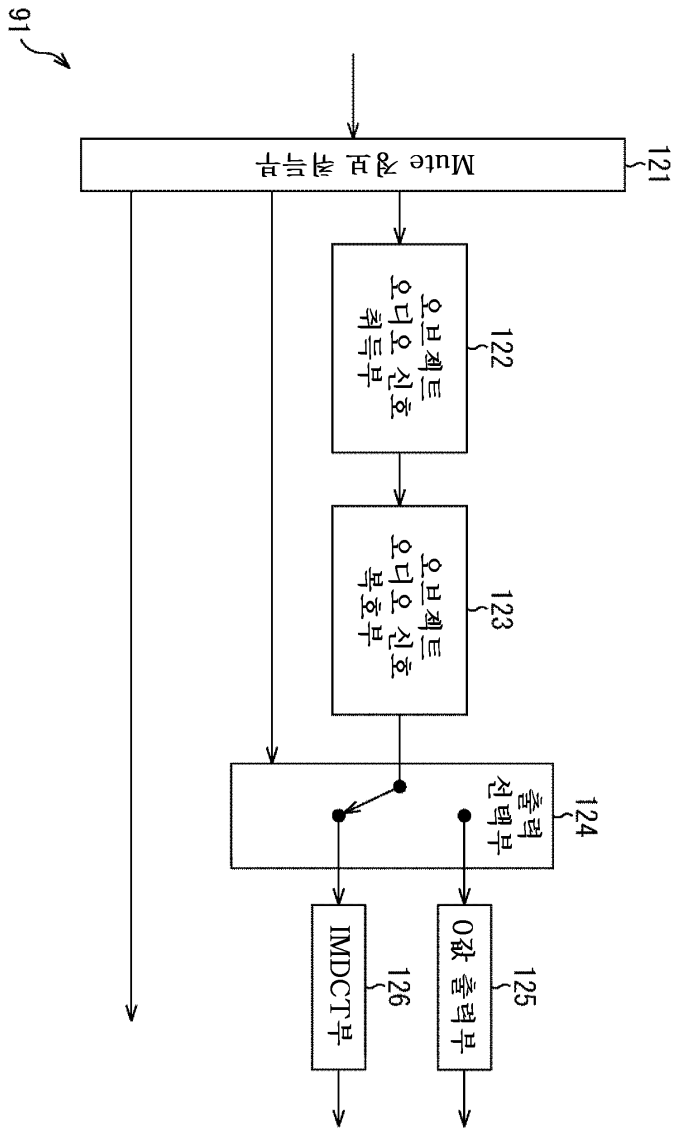
도면5

Syntax	No. of bits	Mnemonic
ObjectMetadataConfig()		
{		
lowDelayMetadataCoding;	1	bslbf
hasCoreLength;	1	bslbf
if (!hasCoreLength) {		
frameLength;	6	uimsbf
OAMFrameLength = (frameLength +1)<<6;		
} else {		
OAMFrameLength = outputFrameLength;		
}		
hasScreenRelativeObjects;	1	bslbf
if(hasScreenRelativeObjects) {		
for (o = 0; o < num_objects; o++) {		
isScreenRelativeObject[o];	1	bslbf
}		
}		
hasDynamicObjectPriority;	1	bslbf
hasUniformSpread;	1	bslbf
for (o = 0; o < num_objects; o++) {		
mutedObjectFlag[o];	1	bslbf
}		
}		

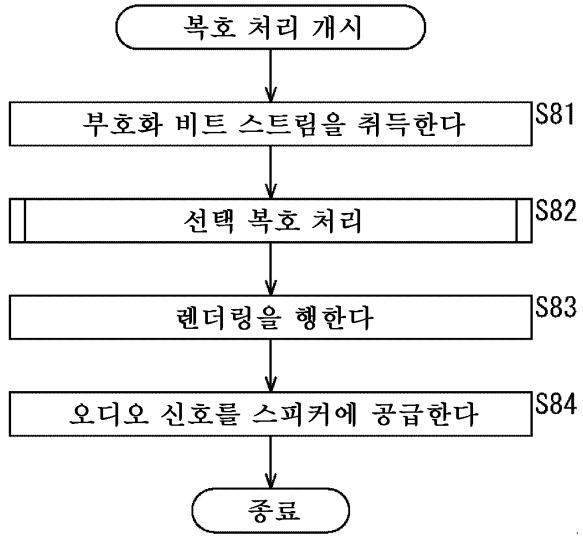
도면6



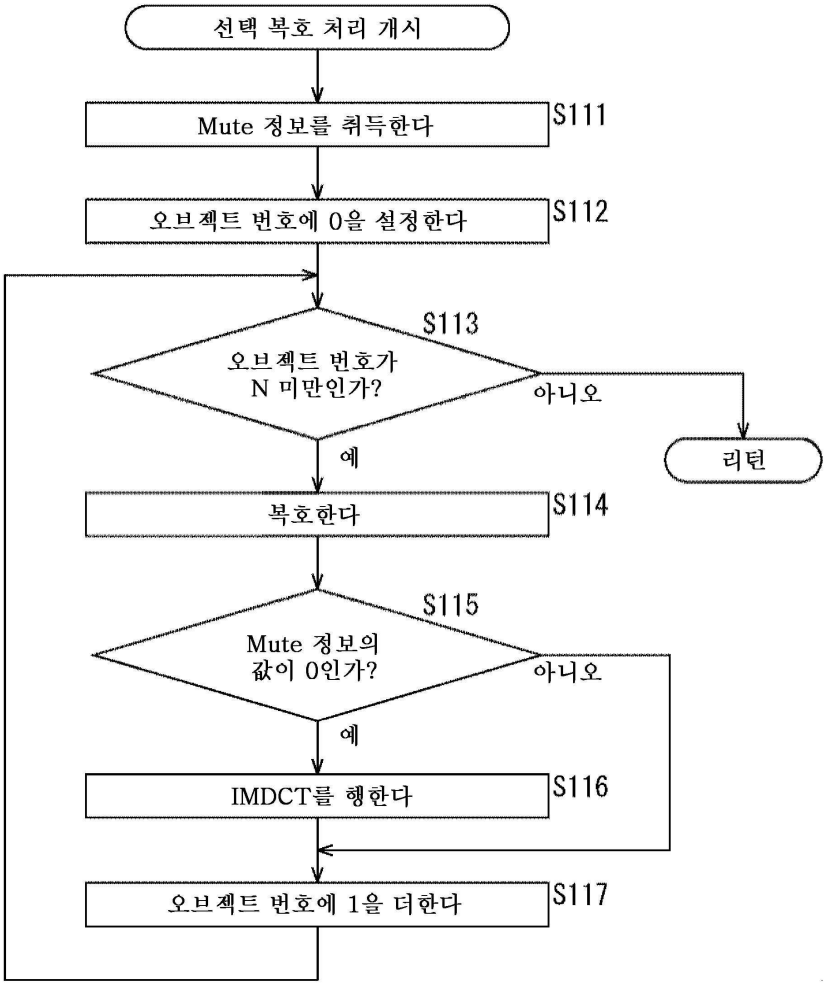
도면7


91
Mute 정보 취득부
121
오브젝트 오디오 신호 취득부
122
오브젝트 오디오 신호 복호부
123
출력 선택부
124
IMDCT부
126
0값 출력부
125

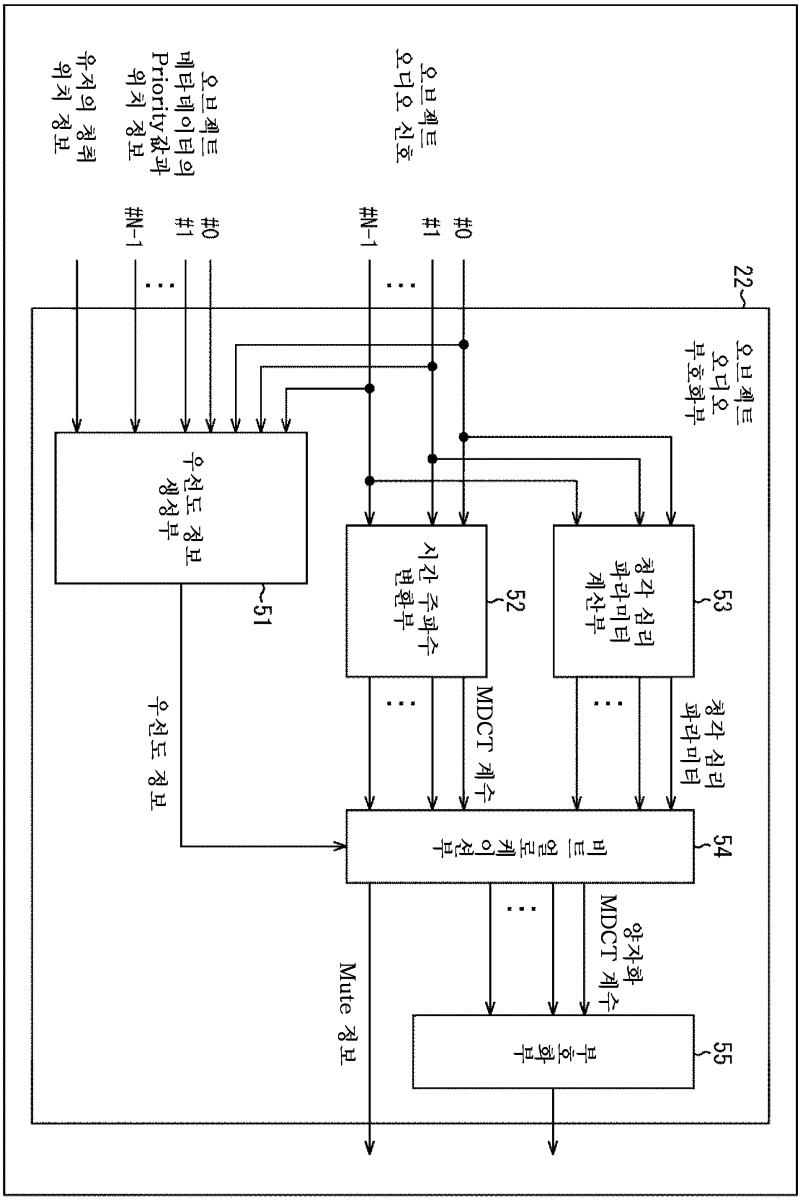
도면8


복호 처리 개시
부호화 비트 스트림을 취득한다
S81
선택 복호 처리
S82
렌더링을 행한다
S83
오디오 신호를 스피커에 공급한다
S84
종료

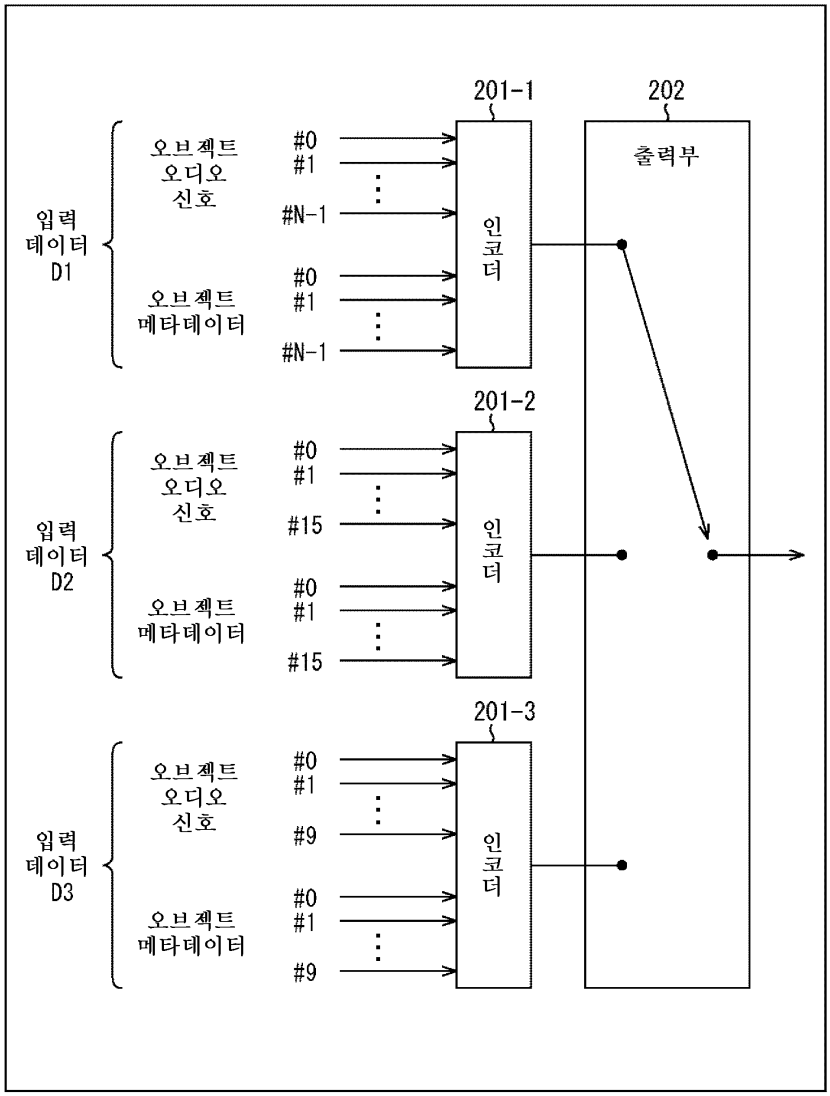
도면9


선택 복호 처리 개시
Mute 정보를 취득한다
S111
오브젝트 번호에 0을 설정한다
S112
S113
오브젝트 번호가
N 미만인가?
아니오
예
리턴
복호한다
S114
S115
Mute 정보의
값이 0인가?
아니오
예
IMDCT를 행한다
S116
오브젝트 번호에 1을 더한다
S117

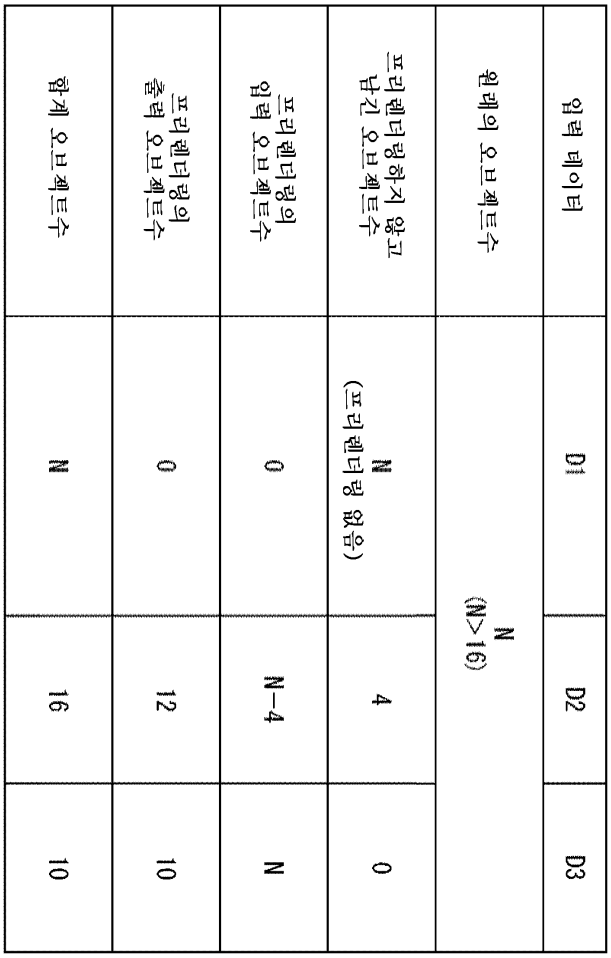
도면10


22
오브젝트 오디오 부호화부
오브젝트 오디오 신호
#0
#1
#N-1
오브젝트 메타데이터의 Priority값과 위치 정보
#0
#1
#N-1
유저의 청취 위치 정보
우선도 정보 생성부
51
시간 주파수 변환부
52
청각 심리 파라미터 계산부
53
MDCT 계수
청각 심리 파라미터
우선도 정보
비트 할당부
54
양자화 MDCT 계수
Mute 정보
부호화부
55

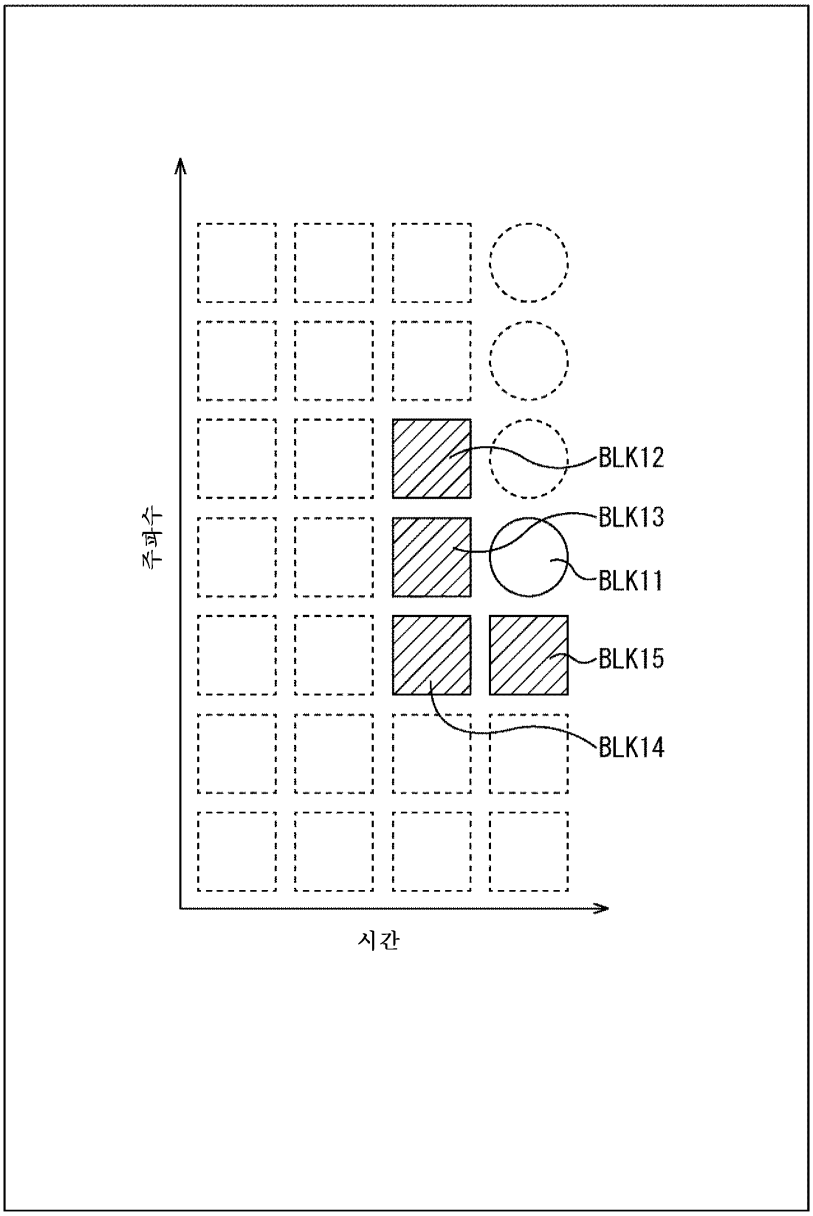
도면11


201-1
202
출력부
입력
데이터
D1
오브젝트
오디오
신호
오브젝트
메타데이터
#0
#1
#N-1
#0
#1
#N-1
인
코
더
201-2
입력
데이터
D2
오브젝트
오디오
신호
오브젝트
메타데이터
#0
#1
#15
#0
#1
#15
인
코
더
201-3
입력
데이터
D3
오브젝트
오디오
신호
오브젝트
메타데이터
#0
#1
#9
#0
#1
#9
인
코
더

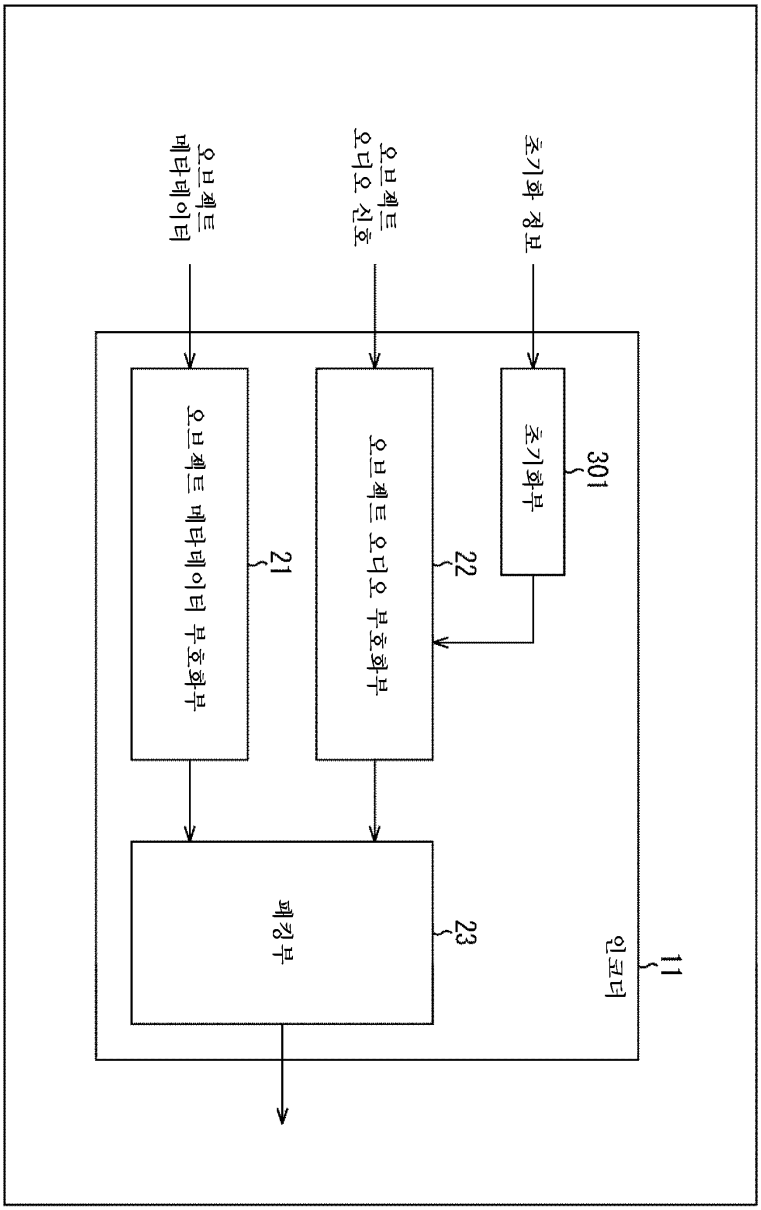
도면12


입력 데이터
D1
D2
D3
원래의 오브젝트수
N
(N>16)
프리렌더링하지 않고
남긴 오브젝트수
N
(프리렌더링 없음)
4
0
프리렌더링의
입력 오브젝트수
0
N-4
N
프리렌더링의
출력 오브젝트수
0
12
10
합계 오브젝트수
N
16
10

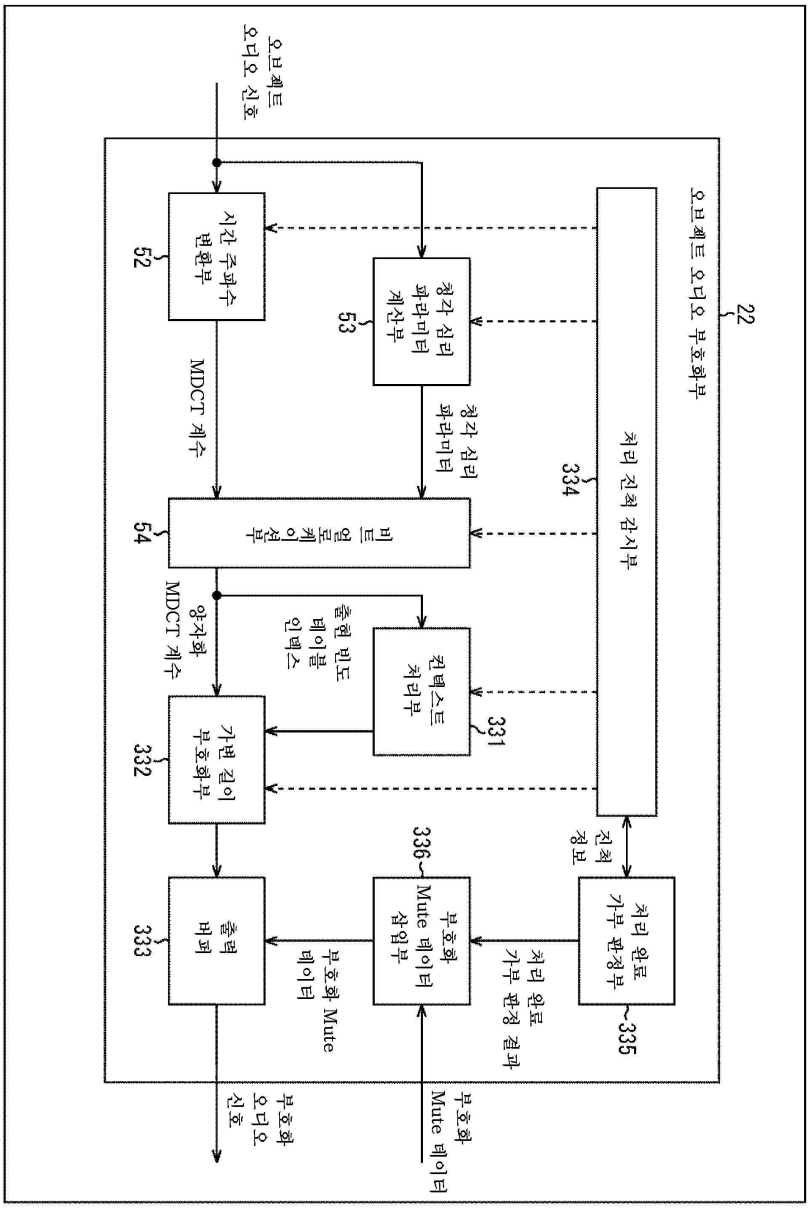
도면13



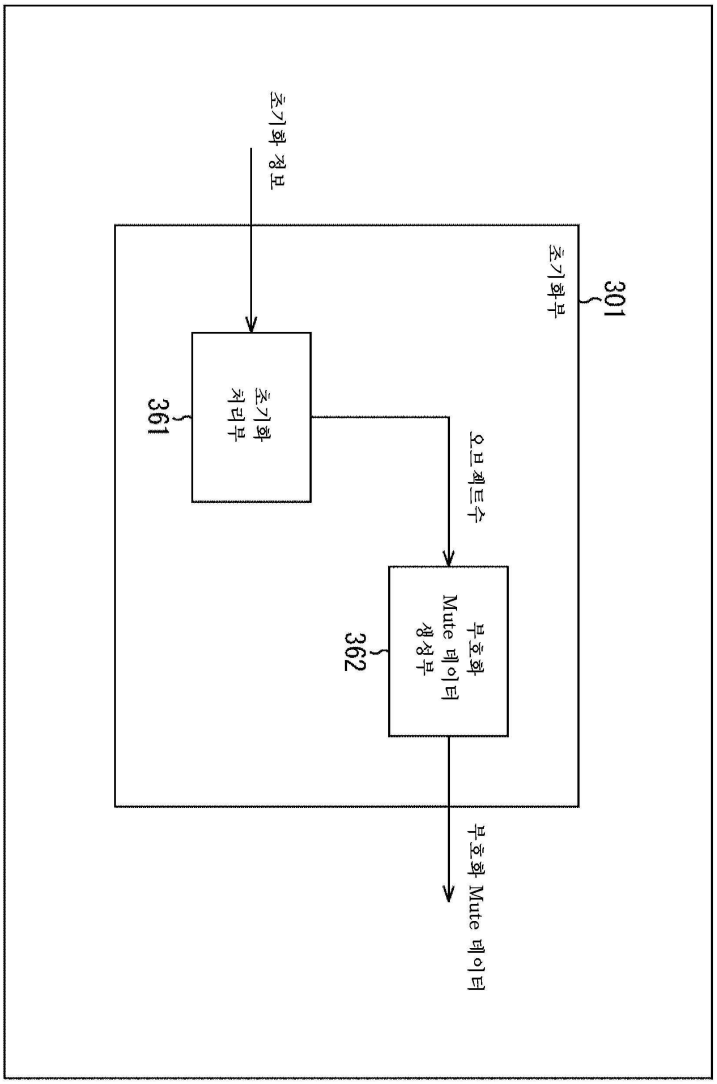
도면14


초기화 정보
오브젝트 오디오 신호
오브젝트 메타데이터
301
초기화부
22
오브젝트 오디오 부호화부
21
오브젝트 메타데이터 부호화부
23
패킹부
11
인코더

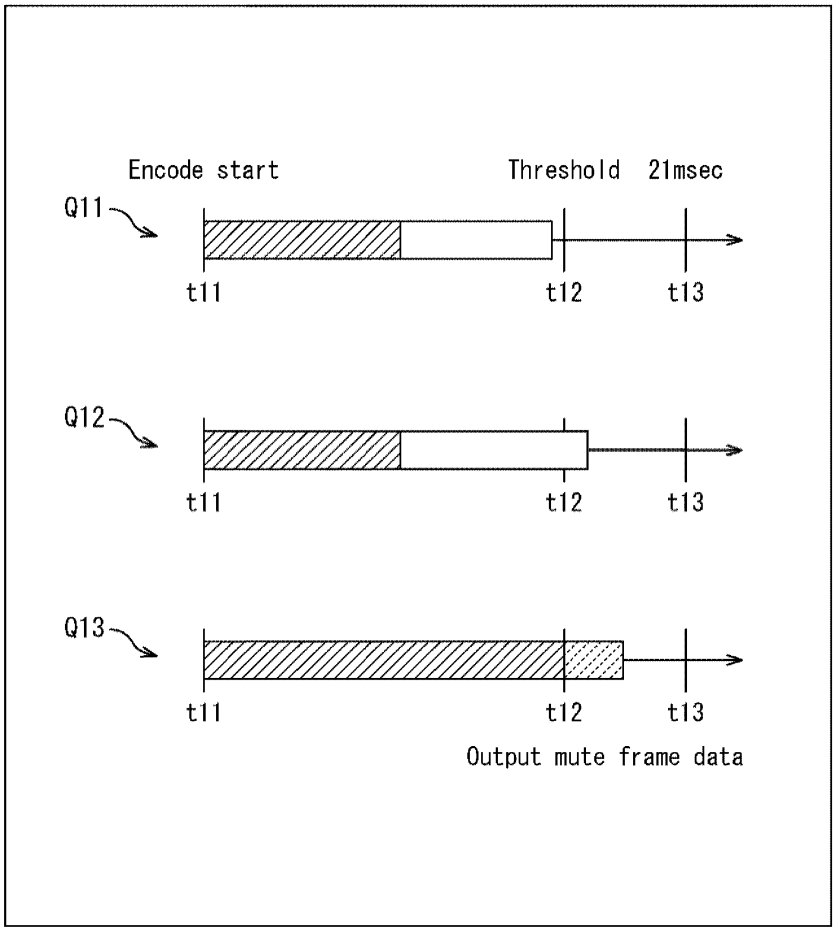
도면15


오브젝트 오디오 신호
오브젝트 오디오 부호화부
22
시간 주파수 변환부
52
청각 심리 파라미터 계산부
53
처리 진척 감시부
334
MDCT 계수
청각 심리 파라미터
비트 할당 시퀀스부
54
양자화 MDCT 계수
출현 빈도 테이블 인덱스
컨텍스트 처리부
331
가변 길이 부호화부
332
진척 정보
처리 완료 가부 판정부
335
336
부호화 Mute 데이터 삽입부
출력 버퍼
333
부호화 Mute 데이터
처리 완료 가부 판정 결과
부호화 오디오 신호
부호화 Mute 데이터

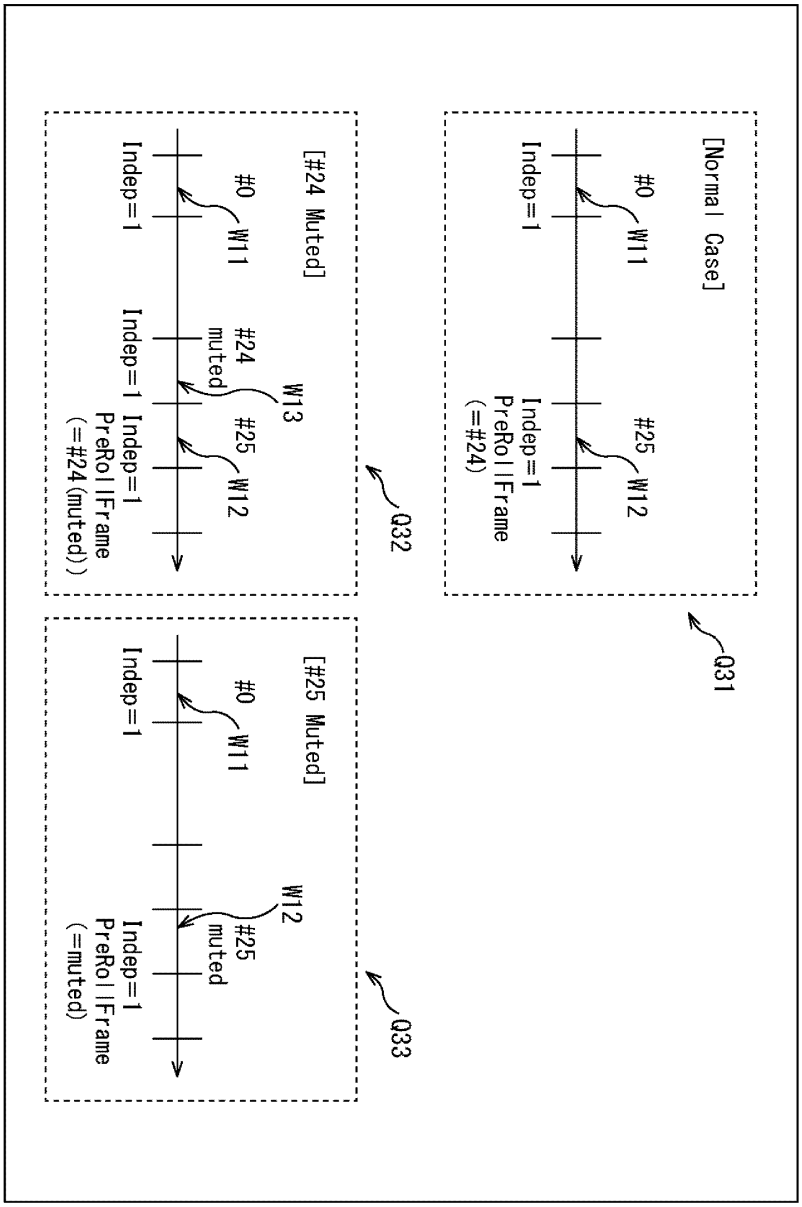
도면16


초기화 정보
초기화부
301
초기화 처리부
361
오브젝트수
부호화 Mute 데이터 생성부
362
부호화 Mute 데이터

도면17


Encode start
Threshold
21msec
Q11
t11
t12
t13
Q12
t11
t12
t13
Q13
t11
t12
t13
Output mute frame data

도면18


[Normal Case]
#0
W11
#25
W12
Indep=1
Indep=1
PreRollFrame
(=#24)
Q31
[#24 Muted]
#0
W11
#24
muted
W13
#25
W12
Indep=1
Indep=1
Indep=1
PreRollFrame
(=#24(muted))
Q32
[#25 Muted]
#0
W11
W12
#25
muted
Indep=1
Indep=1
PreRollFrame
(=muted)
Q33

도면19

Syntax	No. of bits	Mnemonic

mpegh3daFrame()
{
usacIndependencyFlag;
for(elemIdx=0;elemIdx<numElements;++elemIdx){
if((usacElementType[elemIdx] != ID_USAC_EXT)&&
(elementLengthPresent == 1)){
elementLength
}
switch(usacElementType[elemIdx]){
case ID_USAC_SCE:
mpegh3daSingleChannelElement(usacIndependencyFlag);
break;
case ID_USAC_CPE:
mpegh3daChannelPairElement(usacIndependencyFlag);
break;
case ID_USAC_LFE:
mpegh3dalfeElement(usacIndependencyFlag);
break;
case ID_USAC_EXT:
mpegh3daExtElement(usacIndependencyFlag);
break;
}
}
}
}
	1
	
	16
	
	elementLength
	
	elementLength
	
	elementLength
	
	elementLength
		uimsbf
		
		uimsbf

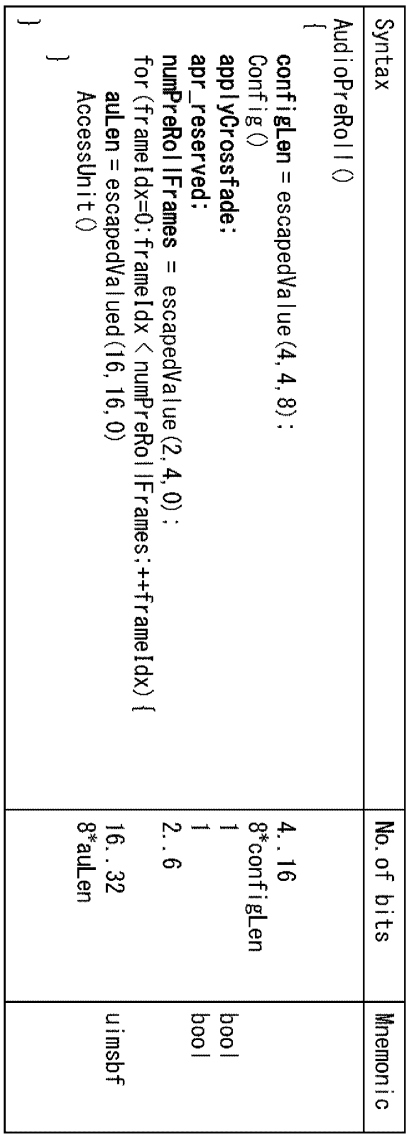
도면20

Syntax	No. of bits	Mnemonic
mpegh3daExtElement(indepFlag)		
{		
usacExtElementPresent	1	uimsbf
if(usacExtElementPresent==1) {		
usacExtElementUseDefaultLength	1	uimsbf
if(usacExtElementUseDefaultLength) {		
usacExtElementPayloadLength = usacExtElementDefaultLength;		
} else {		
usacExtElementPayloadLength;	8	uimsbf
if (usacExtElementPayloadLength == 255)		
valueAdd	16	uimsbf
usacExtElementPayloadLength += valueAdd - 2;		
}		
}		
if(usacExtElementPayloadLength>0) {		
for(i=0;i<usacExtElementPayloadLength;i++) {		
usacExtElementSegmentData[i];	8	uimsbf
}		
}		
}		
}		

도면21

usacExtElementType	The concatenated usacExtElementSegmentData shall be:
ID_EXT_ELE_FILL	Series of fill_byte
ID_EXT_ELE_MPEGS	SpatialFrame() as defined in ISO/IEC 23003-1
ID_EXT_ELE_SAOC	SAOCFrame() as defined in ISO/IEC 23003-2
ID_EXT_ELE_AUDIOPREROLL	AudioPreRoll()
ID_EXT_ELE_UNI_DRC	uniDrcGain() as defined in ISO/IEC 23003-4
ID_EXT_ELE_OBJ_METADATA	objectMetadataFrame()
<omitted>	

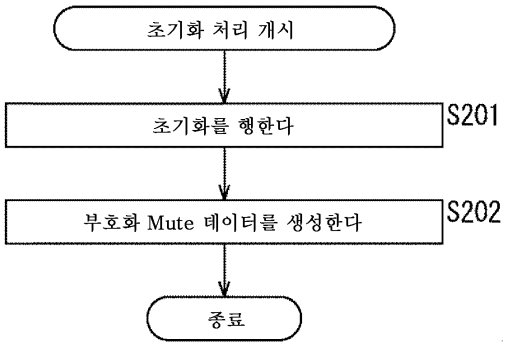
도면22



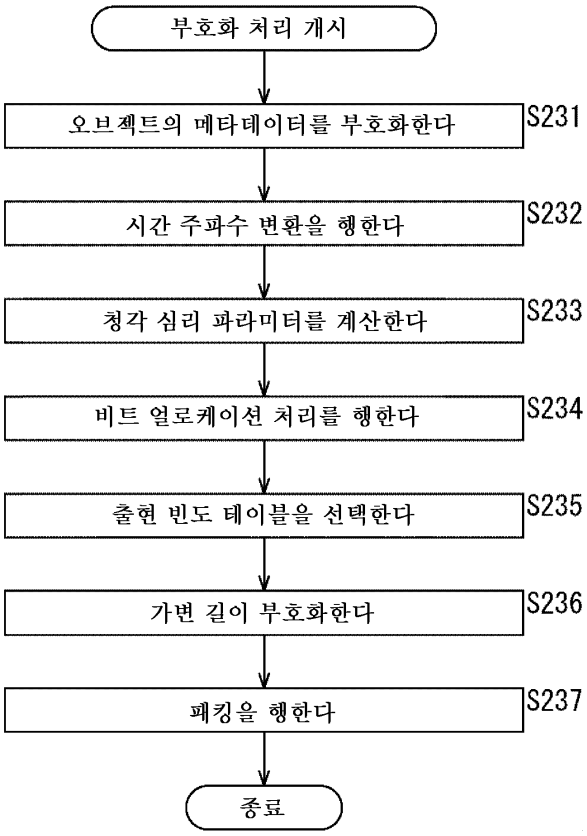
Syntax	No. of bits	Mnemonic

AudioPreRoll()
{
configLen = escapedValue(4, 4, 8);
Config()
applyCrossfade;
apr_reserved;
numPreRollFrames = escapedValue(2, 4, 0);
for (frameIdx=0; frameIdx < numPreRollFrames; ++frameIdx) {
auLen = escapedValued(16, 16, 0)
AccessUnit()
}
}
	4..16
	8*configLen
	1
	1
	2..6
	
	16..32
	8*auLen
		
		bool
		bool
		
		uimsbf

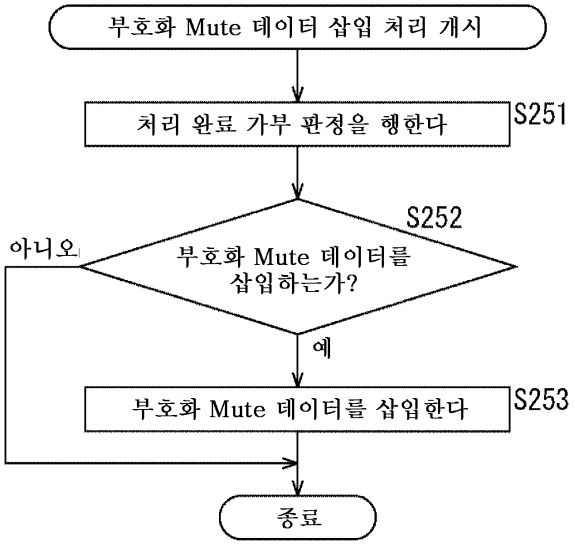
도면23



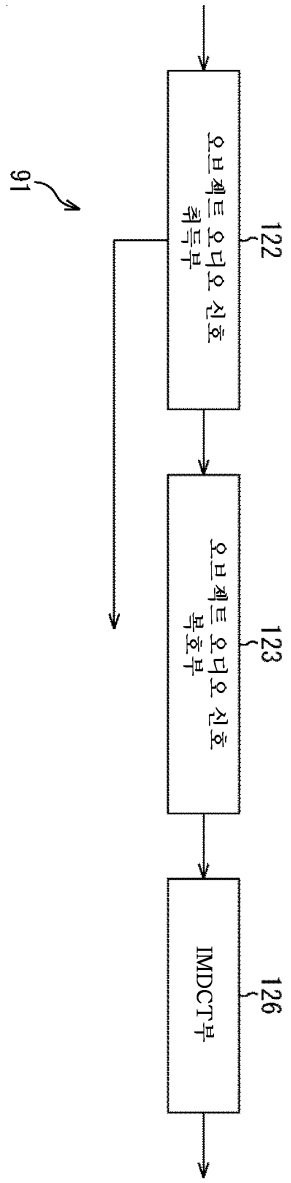
도면24


부호화 처리 개시
오브젝트의 메타데이터를 부호화한다
S231
시간 주파수 변환을 행한다
S232
청각 심리 파라미터를 계산한다
S233
비트 얼로케이션 처리를 행한다
S234
출현 빈도 테이블을 선택한다
S235
가변 길이 부호화한다
S236
패킹을 행한다
S237
종료

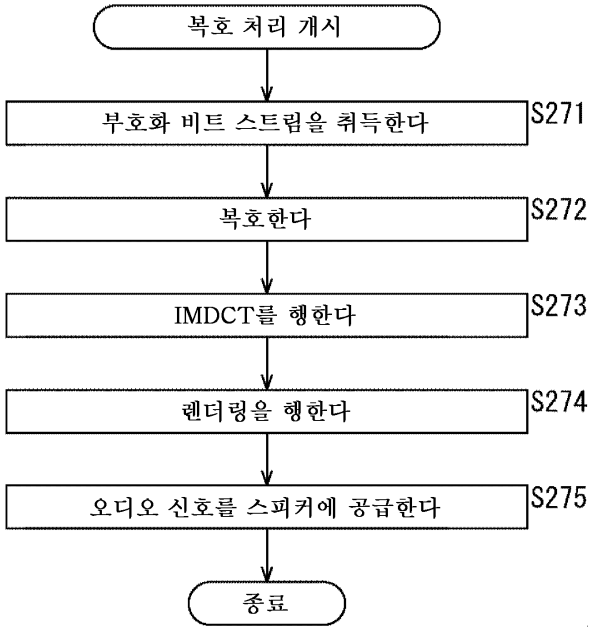
도면25


부호화 Mute 데이터 삽입 처리 개시
처리 완료 가부 판정을 행한다
S251
S252
부호화 Mute 데이터를
삽입하는가?
아니오
예
부호화 Mute 데이터를 삽입한다
S253
종료

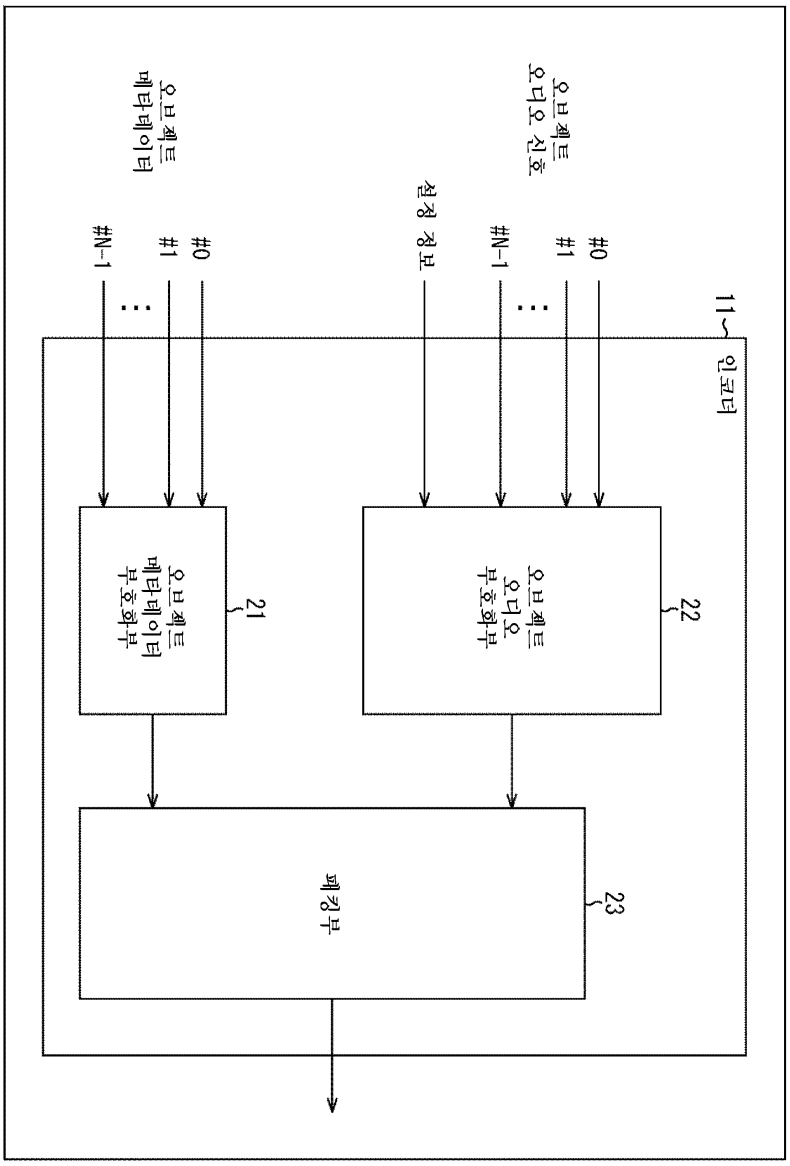
도면26



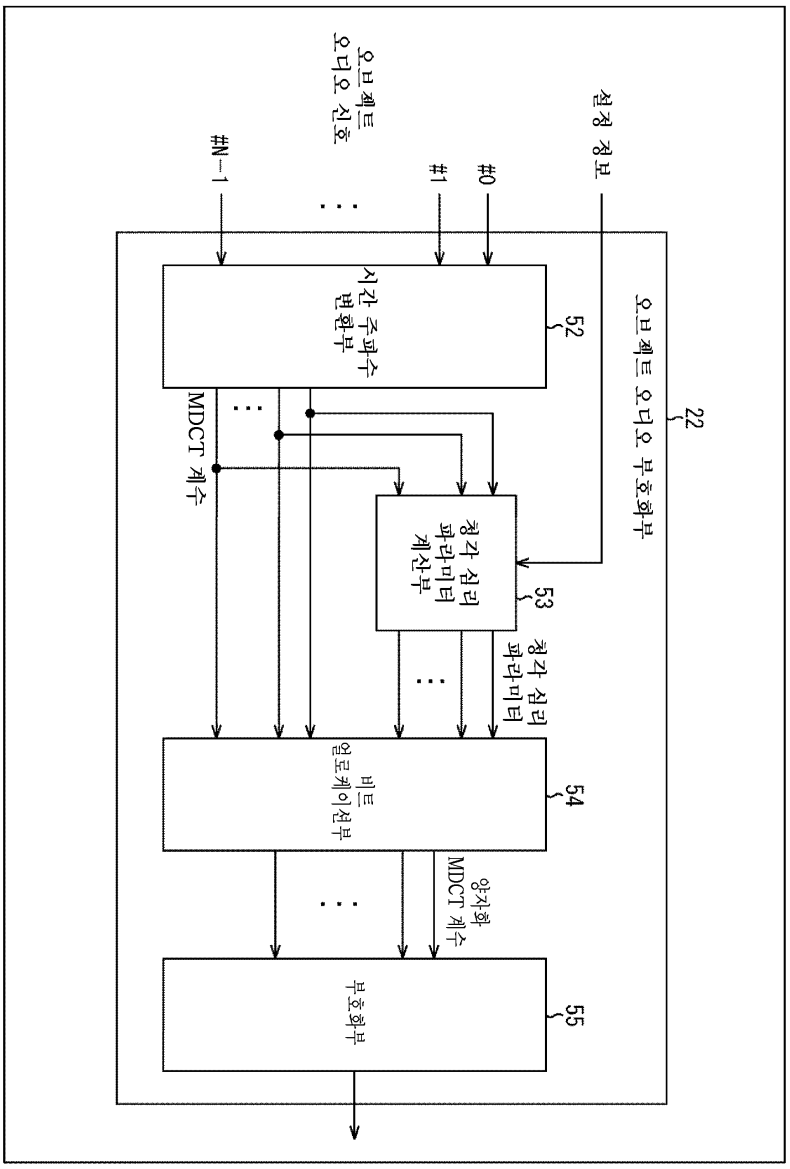
도면27


복호 처리 개시
부호화 비트 스트림을 취득한다
S271
복호한다
S272
IMDCT를 행한다
S273
렌더링을 행한다
S274
오디오 신호를 스피커에 공급한다
S275
종료

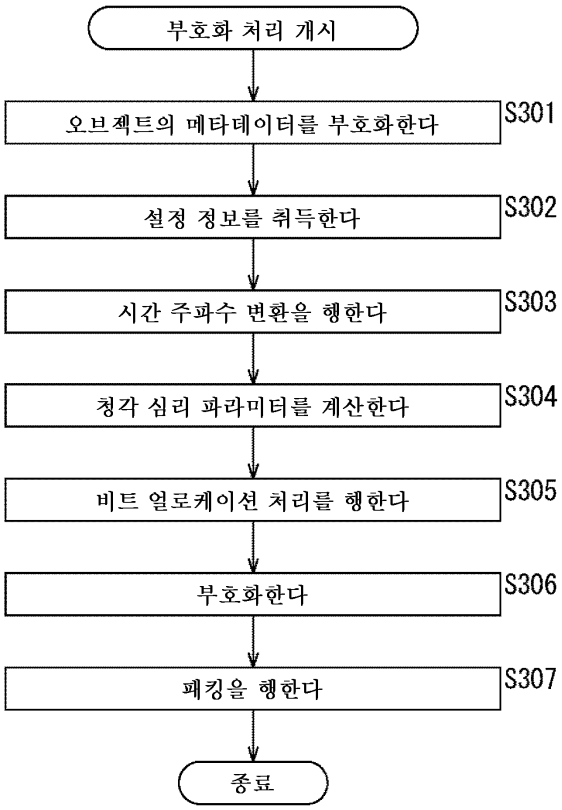
도면28


오브젝트 메타데이터
오브젝트 오디오 신호
설정 정보
#N-1
#1
#0
#N-1
#1
#0
11
인코더
오브젝트 메타데이터 부호화부
21
오브젝트 오디오 부호화부
22
패킹부
23

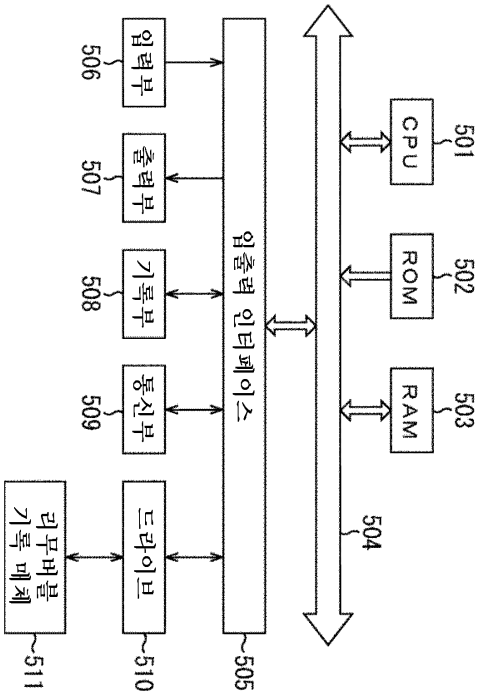
도면29


오브젝트 오디오 신호
#N-1
#1
#0
설정 정보
오브젝트 오디오 부호화부
22
시간 주파수 변환부
52
MDCT 계수
청각 심리 파라미터 계산부
53
청각 심리 파라미터
비트 얼로케이션부
54
양자화 MDCT 계수
부호화부
55

도면30


부호화 처리 개시
오브젝트의 메타데이터를 부호화한다
S301
설정 정보를 취득한다
S302
시간 주파수 변환을 행한다
S303
청각 심리 파라미터를 계산한다
S304
비트 얼로케이션 처리를 행한다
S305
부호화한다
S306
패킹을 행한다
S307
종료

도면31


입력부
506
출력부
507
기록부
508
통신부
509
리무버블 기록 매체
511
드라이브
510
입출력 인터페이스
505
CPU
501
ROM
502
RAM
503
504