

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6432180号
(P6432180)

(45) 発行日 平成30年12月5日 (2018. 12. 5)

(24) 登録日 平成30年11月16日 (2018. 11. 16)

(51) Int. Cl.

F I

H O 4 S 7/00 (2006. 01)

H O 4 S 7/00 3 0 0

G 1 0 L 19/008 (2013. 01)

G 1 0 L 19/008

H O 4 S 3/00 (2006. 01)

H O 4 S 3/00 2 0 0

請求項の数 13 (全 33 頁)

(21) 出願番号 特願2014-130898 (P2014-130898)
 (22) 出願日 平成26年6月26日 (2014. 6. 26)
 (65) 公開番号 特開2016-10090 (P2016-10090A)
 (43) 公開日 平成28年1月18日 (2016. 1. 18)
 審査請求日 平成29年2月3日 (2017. 2. 3)

(73) 特許権者 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100121131
 弁理士 西川 孝
 (74) 代理人 100082131
 弁理士 稲本 義雄
 (72) 発明者 山本 優樹
 東京都港区港南1丁目7番1号 ソニー株
 式会社内
 (72) 発明者 知念 徹
 東京都港区港南1丁目7番1号 ソニー株
 式会社内

最終頁に続く

(54) 【発明の名称】 復号装置および方法、並びにプログラム

(57) 【特許請求の範囲】

【請求項 1】

オーディオエレメントであるSCEとCPEの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、チャンネル音源グループの前記オーディオエレメント、またはオブジェクト音源グループの前記オーディオエレメントを選択することで、前記オーディオエレメントの1つの前記組み合わせを選択する選択部と、

選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成する生成部と

を備える復号装置。

10

【請求項 2】

前記選択部は、同じコンテンツについて予め用意された複数の前記組み合わせの中から1つの前記組み合わせを選択する

請求項 1 に記載の復号装置。

【請求項 3】

前記複数の前記組み合わせごとに用意された、前記組み合わせの前記オーディオエレメントから構成されるビットストリームのうちの、前記選択部により選択された前記組み合わせのビットストリームを受信する通信部をさらに備える

請求項 2 に記載の復号装置。

【請求項 4】

20

前記選択部は、ビットストリームを構成する複数の前記オーディオエレメントのうちのいくつかの前記オーディオエレメントを、1つの前記組み合わせとして選択する

請求項1または請求項2に記載の復号装置。

【請求項5】

前記選択部は、前記ビットストリームのメタデータに基づいて1つの前記組み合わせを選択する

請求項4に記載の復号装置。

【請求項6】

前記選択部は、前記メタデータとして、予め定められた複数の前記組み合わせを示す情報、および前記オーディオエレメントの優先度情報のうちの少なくとも何れか一方に基づいて、1つの前記組み合わせを選択する

10

請求項5に記載の復号装置。

【請求項7】

前記ビットストリームから、前記選択部により選択された前記組み合わせの前記オーディオエレメントを抽出する抽出部をさらに備える

請求項4乃至請求項6の何れか一項に記載の復号装置。

【請求項8】

前記選択部により選択された前記組み合わせの前記オーディオエレメントを受信する通信部をさらに備える

請求項4乃至請求項6の何れか一項に記載の復号装置。

20

【請求項9】

デコード対象として選択されなかった前記オーディオエレメントのサイズに基づいて、前記生成部によりデコードされる前記オーディオエレメントのバッファへの格納を制御するバッファ制御部をさらに備える

請求項5に記載の復号装置。

【請求項10】

前記選択部は、選択した前記組み合わせを構成する前記オーディオエレメントのなかから、デコード対象としない前記オーディオエレメントをさらに選択し、

前記バッファ制御部は、前記選択部により選択された前記デコード対象としない前記オーディオエレメントのサイズに基づいて、前記選択部により選択された前記組み合わせを構成する、前記デコード対象としない前記オーディオエレメント以外の前記オーディオエレメントの前記バッファへの格納を制御する

30

請求項9に記載の復号装置。

【請求項11】

前記選択部は、前記オーディオエレメントの優先度情報に基づいて、前記デコード対象としない前記オーディオエレメントを選択する

請求項10に記載の復号装置。

【請求項12】

オーディオエレメントであるSCEとCPEの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、チャンネル音源グループの前記オーディオエレメント、またはオブジェクト音源グループの前記オーディオエレメントを選択することで、前記オーディオエレメントの1つの前記組み合わせを選択し、

40

選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成する

ステップを含む復号方法。

【請求項13】

オーディオエレメントであるSCEとCPEの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、チャンネル音源グループの前記オーディオエレメント、またはオブジェクト音源グループの前記オーディオ

50

エレメントを選択することで、前記オーディオエレメントの1つの前記組み合わせを選択し、

選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成する

ステップを含む処理をコンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本技術は復号装置および方法、並びにプログラムに関し、特に、ハード規模の異なる機器でビットストリームをデコードできるようにした復号装置および方法、並びにプログラムに関する。

【背景技術】

【0002】

従来の5.1チャンネルサラウンド再生を超える、より高臨場感な再生や複数の音素材（オブジェクト）を伝送するための符号化技術として3D Audio規格が知られている（例えば、非特許文献1乃至3参照）。

【0003】

3D Audio規格では、デコーダで持つべき入力ビットストリームを格納しておくためのバッファのサイズの最小値がMinimum decoder input bufferサイズとして規定されている。例えば非特許文献3の4.5.3.1節では、Minimum decoder input buffer サイズ=6144×NCC (bits)と規定されている。

【0004】

ここで、NCCはNumber of Considered Channelの略であり、入力ビットストリームに含まれる全オーディオエレメントのうち、SCE(Single Channel Element)の数と、CPE(Channel Pair Element)の数の2倍との和を表している。

【0005】

また、SCEは1つのチャンネルのオーディオ信号が格納されるオーディオエレメントであり、CPEはペアとなる2つのチャンネルのオーディオ信号が格納されるオーディオエレメントである。したがって、例えば入力ビットストリームに含まれるSCEの数が5個でありCPEの数が3個の場合、 $NCC=5+2 \times 3=11$ となる。

【0006】

このように3D Audio規格では、デコーダが入力ビットストリームをデコードしようとするときには、規定されたサイズのバッファを最低限確保する必要がある。

【先行技術文献】

【非特許文献】

【0007】

【非特許文献1】ISO/IEC JTC1/SC29/WG11 N14459, April 2014, Valencia, Spain, "Text of ISO/IEC 23008-3/CD, 3D audio"

【非特許文献2】INTERNATIONAL STANDARD ISO/IEC 23003-3 First edition 2012-04-01 Information technology-coding of audio-visual objects-part3:Unified speech and audio coding

【非特許文献3】INTERNATIONAL STANDARD ISO/IEC 14496-3 Fourth edition 2009-09-01 Information technology-coding of audio-visual objects-part3:Audio

【発明の概要】

【発明が解決しようとする課題】

【0008】

ところが非特許文献1の3D Audio規格では、SCEの数およびCPEの数をほとんど任意に設定できるため、3D Audio規格で定めることのできるビットストリームの全てをデコードするためには、デコーダで持つべきMinimum decoder input bufferサイズが、非特許文献3の規格などに比べて非常に大きくなってしまふ。

10

20

30

40

50

【 0 0 0 9 】

具体的には、非特許文献 1 の 3D Audio 規格では、SCE の数と CPE の数を合計で最大 65805 個持つことができる。よって、Minimum decoder input buffer のサイズの最大値は、Minimum decoder input buffer サイズの最大値 = $6144 \times (0 + 65805 \times 2) = 808611840$ (bits) となり、約 100MByte となる。

【 0 0 1 0 】

このように最低限必要とされるバッファのサイズである Minimum decoder input buffer サイズが大きくなると、メモリサイズの小さいプラットフォームなどでは、この規定を満たすサイズのバッファを確保することができない場合もある。すなわち、機器のハード規模によっては、デコーダを実装できない場合がある。

10

【 0 0 1 1 】

本技術は、このような状況に鑑みてなされたものであり、ハード規模の異なる機器でビットストリームをデコードできるようにするものである。

【課題を解決するための手段】

【 0 0 1 2 】

本技術の一側面の復号装置は、オーディオエレメントである SCE と CPE の組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、チャンネル音源グループの前記オーディオエレメント、またはオブジェクト音源グループの前記オーディオエレメントを選択することで、前記オーディオエレメントの 1 つの前記組み合わせを選択する選択部と、選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成する生成部とを備える。

20

【 0 0 1 3 】

前記選択部には、同じコンテンツについて予め用意された複数の前記組み合わせの中から 1 つの前記組み合わせを選択させることができる。

【 0 0 1 4 】

復号装置には、前記複数の前記組み合わせごとに用意された、前記組み合わせの前記オーディオエレメントから構成されるビットストリームのうちの、前記選択部により選択された前記組み合わせのビットストリームを受信する通信部をさらに設けることができる。

【 0 0 1 5 】

前記選択部には、ビットストリームを構成する複数の前記オーディオエレメントのうちのいくつかの前記オーディオエレメントを、1 つの前記組み合わせとして選択させることができる。

30

【 0 0 1 6 】

前記選択部には、前記ビットストリームのメタデータに基づいて 1 つの前記組み合わせを選択させることができる。

【 0 0 1 7 】

前記選択部には、前記メタデータとして、予め定められた複数の前記組み合わせを示す情報、および前記オーディオエレメントの優先度情報のうちの少なくとも何れか一方に基づいて、1 つの前記組み合わせを選択させることができる。

40

【 0 0 1 8 】

復号装置には、前記ビットストリームから、前記選択部により選択された前記組み合わせの前記オーディオエレメントを抽出する抽出部をさらに設けることができる。

【 0 0 1 9 】

復号装置には、前記選択部により選択された前記組み合わせの前記オーディオエレメントを受信する通信部をさらに設けることができる。

【 0 0 2 0 】

復号装置には、デコード対象として選択されなかった前記オーディオエレメントのサイズに基づいて、前記生成部によりデコードされる前記オーディオエレメントのバッファへの格納を制御するバッファ制御部をさらに設けることができる。

【 0 0 2 1 】

50

前記選択部には、選択した前記組み合わせを構成する前記オーディオエレメントのなかから、デコード対象としない前記オーディオエレメントをさらに選択させ、前記バッファ制御部には、前記選択部により選択された前記デコード対象としない前記オーディオエレメントのサイズに基づいて、前記選択部により選択された前記組み合わせを構成する、前記デコード対象としない前記オーディオエレメント以外の前記オーディオエレメントの前記バッファへの格納を制御させることができる。

【0022】

前記選択部には、前記オーディオエレメントの優先度情報に基づいて、前記デコード対象としない前記オーディオエレメントを選択させることができる。

【0023】

本技術の一側面の復号方法またはプログラムは、オーディオエレメントであるSCEとCPEの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、チャンネル音源グループの前記オーディオエレメント、またはオブジェクト音源グループの前記オーディオエレメントを選択することで、前記オーディオエレメントの1つの前記組み合わせを選択し、選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成するステップを含む。

【0024】

本技術の一側面においては、オーディオエレメントであるSCEとCPEの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、チャンネル音源グループの前記オーディオエレメント、またはオブジェクト音源グループの前記オーディオエレメントを選択することで、前記オーディオエレメントの1つの前記組み合わせが選択され、選択された前記組み合わせの前記オーディオエレメントがデコードされてオーディオ信号が生成される。

【発明の効果】

【0025】

本技術の一側面によれば、ハード規模の異なる機器でビットストリームをデコードすることができる。

【0026】

なお、ここに記載された効果は必ずしも限定されるものではなく、本開示中に記載された何れかの効果であってもよい。

【図面の簡単な説明】

【0027】

【図1】入力ビットストリームの構成について説明する図である。

【図2】入力ビットストリームの配置例について説明する図である。

【図3】優先度情報について説明する図である。

【図4】転送ビットレートの調整について説明する図である。

【図5】転送ビットレートの調整について説明する図である。

【図6】転送ビットレートの調整について説明する図である。

【図7】サイズ情報について説明する図である。

【図8】コンテンツ配信システムの構成例を示す図である。

【図9】デコーダの構成例を示す図である。

【図10】復号処理を説明するフローチャートである。

【図11】デコーダの構成例を示す図である。

【図12】復号処理を説明するフローチャートである。

【図13】デコーダの構成例を示す図である。

【図14】復号処理を説明するフローチャートである。

【図15】デコーダの構成例を示す図である。

【図16】復号処理を説明するフローチャートである。

【図17】コンピュータの構成例を示す図である。

【発明を実施するための形態】

10

20

30

40

50

【0028】

以下、図面を参照して、本技術を適用した実施の形態について説明する。

【0029】

第1の実施の形態

本技術は、様々な許容メモリサイズのデコーダ、つまりハード規模の異なる様々な機器において、エンコード（符号化）されたマルチチャネルオーディオ信号が格納されている入力ビットストリームをデコード（復号）することができるようにするものである。

【0030】

本技術では、入力ビットストリーム内のオーディオエレメントの組み合わせを、入力ビットストリーム内で複数個定義し、オーディオエレメントの組み合わせごとに、デコーダで持つべき入力ビットストリームを格納しておくためのバッファのサイズの最小値を変化させることで、異なるハード規模でデコードができるようにされる。

【0031】

まず、本技術の概要について説明する。

【0032】

オーディオエレメントの組み合わせの定義の追加

本技術では3D Audio規格において、オーディオエレメントの組み合わせを複数個定義できるようにされている。ここでは、様々な許容メモリサイズのデコーダで入力ビットストリームをデコードできるように、複数個の組み合わせの定義が行われる。

【0033】

例えば1つのコンテンツを再生するための入力ビットストリームが、図1に示すオーディオエレメントで構成されたとする。なお、図中、1つの長方形は入力ビットストリームを構成する1つのオーディオエレメントを表している。また、 $SCE(i)$ （但し i は整数）と記されているオーディオエレメントは、 i 番目のSCEを表しており、 $CPE(i)$ （但し i は整数）と記されているオーディオエレメントは、 i 番目のCPEを表している。

【0034】

上述したようにSCEは、1チャンネル分のオーディオ信号をデコードするために必要なデータ、つまり1チャンネル分のオーディオ信号をエンコードして得られた符号化データが格納されているオーディオエレメントである。また、CPEはペアとなる2チャンネル分のオーディオ信号をデコードするために必要なデータである。

【0035】

図1では、 $CPE(1)$ は2チャンネル再生のための環境音が格納されたオーディオエレメントとなっている。以下では、 $CPE(1)$ からなるエレメントのグループをチャンネル音源グループ1とも称することとする。

【0036】

また、 $SCE(1)$ 、 $CPE(2)$ 、および $CPE(3)$ は5チャンネル再生のための環境音が格納されたオーディオエレメントとなっており、以下では、 $SCE(1)$ 、 $CPE(2)$ 、および $CPE(3)$ からなるエレメントのグループをチャンネル音源グループ2とも称する。

【0037】

$SCE(2)$ 乃至 $SCE(23)$ は22チャンネル再生のための環境音が格納されたオーディオエレメントとなっており、以下では、 $SCE(2)$ 乃至 $SCE(23)$ からなるエレメントのグループをチャンネル音源グループ3とも称する。

【0038】

$SCE(24)$ はオブジェクト（音素材）として所定言語、例えば日本語の対話音が格納されたオーディオエレメントとなっており、以下では、 $SCE(24)$ からなるエレメントのグループをオブジェクト音源グループ1とも称する。同様に $SCE(25)$ はオブジェクトとして韓国語の対話音が格納されたオーディオエレメントとなっており、以下では、 $SCE(25)$ からなるエレメントのグループをオブジェクト音源グループ2とも称する。

【0039】

さらに、 $SCE(26)$ 乃至 $SCE(30)$ は、オブジェクトとして車などの音が格納されたオーディ

10

20

30

40

50

オエレメントとなっており、以下では、SCE(26)乃至SCE(30)からなるエレメントのグループをオブジェクト音源グループ3とも称する。

【 0 0 4 0 】

入力ビットストリームをデコードしてコンテンツを再生する際には、デコーダではこれらのチャンネル音源グループ1乃至チャンネル音源グループ3やオブジェクト音源グループ1乃至オブジェクト音源グループ3を任意に組み合わせて再生することができる。

【 0 0 4 1 】

そのような場合、図1の例では、チャンネル音源グループやオブジェクト音源グループのオーディオエレメントの組み合わせは、以下の6つの組み合わせCM(1)乃至CM(6)となる。

【 0 0 4 2 】

組み合わせCM(1)

チャンネル音源グループ1、オブジェクト音源グループ1、オブジェクト音源グループ3

組み合わせCM(2)

チャンネル音源グループ1、オブジェクト音源グループ2、オブジェクト音源グループ3

組み合わせCM(3)

チャンネル音源グループ2、オブジェクト音源グループ1、オブジェクト音源グループ3

組み合わせCM(4)

チャンネル音源グループ2、オブジェクト音源グループ2、オブジェクト音源グループ3

組み合わせCM(5)

チャンネル音源グループ3、オブジェクト音源グループ1、オブジェクト音源グループ3

組み合わせCM(6)

チャンネル音源グループ3、オブジェクト音源グループ2、オブジェクト音源グループ3

【 0 0 4 3 】

これらの各組み合わせCM(1)乃至組み合わせCM(6)は、それぞれ2チャンネル日本語、2チャンネル韓国語、5チャンネル日本語、5チャンネル韓国語、22チャンネル日本語、および22チャンネル韓国語でコンテンツを再生するためのオーディオエレメントの組み合わせとなる。

【 0 0 4 4 】

また、この場合、それぞれの組み合わせで必要となるデコーダのメモリサイズの大小関係は、以下ようになる。

【 0 0 4 5 】

組み合わせCM(1) , CM(2) < 組み合わせCM(3) , CM(4) < 組み合わせCM(5) , CM(6)

【 0 0 4 6 】

このようなオーディオエレメントの組み合わせは、ビットストリームシンタックスとして定義することで実現することができる。

【 0 0 4 7 】

Minimum decoder input bufferの定義の修正

ところで3D Audio規格において、上述の組み合わせごとに、Minimum decoder input bufferのサイズが変えられるように、以下に示す現状の規定を修正することで、様々な許容メモリサイズのデコーダで入力ビットストリームをデコードできるようになる。

【 0 0 4 8 】

(現状の規定)

Minimum decoder input bufferサイズ=6144×NCC(bits)

【 0 0 4 9 】

上述したようにNCCは入力ビットストリームに含まれる全オーディオエレメントのうちのSCEの数と、CPEの数の2倍との和を表している。したがって、現状では、自身の許容メモリサイズ、つまり最大限確保可能なバッファサイズが、Minimum decoder input bufferサイズ(以下、必要バッファサイズとも称する)未満である機器では、所定の組み合わせに限っては十分なバッファサイズを確保できる場合であっても、入力ビットストリームをデコードすることができない。

【 0 0 5 0 】

10

20

30

40

50

そこで、本技術では、以下のような修正AM1または修正AM2を行うことで、各機器が自身のハード規模、つまり許容メモリサイズに応じて、自身に適したオーディオエレメントの組み合わせでコンテンツ（入力ビットストリーム）のデコードおよび再生を行うことができるようにした。

【 0 0 5 1 】

（修正AM1）

3D Audio規格で定められている規定において、NCCを入力ビットストリームに含まれる全オーディオエレメントのうちのSCEの数と、CPEの数の2倍との和としているのに代えて、NCCを入力ビットストリームに含まれるデコード対象となるオーディオエレメントの組み合わせに含まれる全オーディオエレメントのうちのSCEの数と、CPEの数の2倍との和とする

10

【 0 0 5 2 】

（修正AM2）

オーディオエレメントの組み合わせごとのMinimum decoder input bufferのサイズ（必要バッファサイズ）をビットストリームシンタックスとして定義する

【 0 0 5 3 】

このような修正AM1または修正AM2を行うことで、デコーダ側において、より許容メモリサイズが小さい機器においても入力ビットストリームのデコードを行うことができるようになるが、そのためにはデコーダ側とエンコーダ側において以下の修正が必要である。

【 0 0 5 4 】

20

（デコーダの信号処理の修正）

デコーダは、自身の許容メモリサイズと、入力ビットストリームのオーディオエレメントの組み合わせごとのサイズ（必要バッファサイズ）とを比較し、「自身の許容メモリサイズ 組み合わせごとのサイズ」という条件を満たすオーディオエレメントの組み合わせを特定し、その条件を満たす何れかの組み合わせのオーディオエレメントをデコードする

【 0 0 5 5 】

ここで、オーディオエレメントの組み合わせごとの必要バッファサイズを特定する方法は、修正AM1または修正AM2の何れかを適用すればよい。

【 0 0 5 6 】

30

すなわち、修正AM1を適用する場合には、例えばデコーダは、取得した入力ビットストリームに格納されている情報からオーディオエレメントの組み合わせを特定し、オーディオエレメントの各組み合わせについて必要バッファサイズを算出すればよい。また、修正AM2を適用する場合には、デコーダは入力ビットストリームからオーディオエレメントの組み合わせごとの必要バッファサイズを読み出せばよい。

【 0 0 5 7 】

なお、デコード対象とするオーディオエレメントの組み合わせは、必要バッファサイズが許容メモリサイズ以下となる組み合わせのうちの、ユーザ等により指定されたものとされてもよい。また、デコード対象とするオーディオエレメントの組み合わせは、必要バッファサイズが許容メモリサイズ以下となる組み合わせのうちの、定めた設定等により選択されたものとされてもよい。

40

【 0 0 5 8 】

さらに、以下では、オーディオエレメントの組み合わせについての必要バッファサイズが許容メモリサイズ以下となるという条件を、バッファサイズ条件とも称することとする。

【 0 0 5 9 】

デコード対象とするオーディオエレメントの組み合わせは、入力ビットストリームの取得前に選択されるようにしてもよいし、入力ビットストリームを取得してから選択されるようにしてもよい。つまり、本技術は、例えばテレビジョン放送などのプッシュ型のコンテンツ配信システムに適用することもできるし、MPEG(Moving Picture Experts Group)-D

50

ASH(Dynamic Adaptive Streaming over HTTP)に代表されるプル型のコンテンツ配信システムにも適用することができる。

【0060】

(エンコーダの運用規定の修正)

エンコーダは、オーディオエレメントの全ての組み合わせについて、修正規定したMinimum decoder input bufferサイズでデコードできるように、時間フレームごとのオーディオエレメント(符号化データ)のビット量の調整を行いエンコードする。

【0061】

すなわち、エンコーダは、デコーダがどのようなオーディオエレメントの組み合わせを選択したとしても、デコーダ側のバッファサイズが必要バッファサイズであるときにオーディオエレメントをデコードできるように、時間フレームごとに各チャネルの符号化データへの割り当てビット量を調整しながらエンコードを行う。ここで、オーディオエレメントをデコードできるとは、デコード対象とする組み合わせのオーディオエレメントを蓄積しておくバッファにおいて、オーバーフローもアンダーフローも生じさせることなくデコードを行うことができることをいう。

10

【0062】

以上のように、デコーダ側で、オーディオエレメントの組み合わせごとの必要バッファサイズに応じて、オーディオエレメントの組み合わせを適切に選択することにより、様々な許容メモリサイズのデコーダで、入力ビットストリームのデコードを行うことができる。すなわち、ハード規模の異なる様々な機器で入力ビットストリームをデコードすることができるようになる。

20

【0063】

オブジェクトの優先度情報を用いた転送ビットレートの削減

さらに、本技術をプル型のコンテンツ配信システムに適用する場合、メタデータ等に基づいて、必要なオーディオエレメントのみを選択して取得するようにすることで、入力ビットストリームの転送ビットレートを削減することができる。換言すれば、デコーダが不要なオーディオエレメントを取得しないようにすることで、入力ビットストリームの転送ビットレートを削減することができる。

【0064】

ここではMPEG-DASHに代表されるプル型のコンテンツ配信サービスについて考えることとする。そのような場合、3D Audioの入力ビットストリームが、サーバ上に例えば以下の配置パターン(1)または配置パターン(2)の2通りのうちの何れかで配置される。

30

【0065】

(配置パターン(1))

3D Audioの入力ビットストリーム全部を1本のストリームとして配置

【0066】

(配置パターン(2))

3D Audioの入力ビットストリームを、オーディオエレメントの組み合わせごとに分割して配置

【0067】

40

具体的には配置パターン(1)では、例えば図1に示したように全組み合わせのオーディオエレメント、つまり全てのチャネル音源グループおよびオブジェクト音源グループを構成するオーディオエレメントを含む1つの入力ビットストリームがサーバに配置される。

【0068】

この場合、デコーダは、例えば予めサーバ等から取得した情報や入力ビットストリームのヘッダ等に格納されている情報(メタデータ)などから、デコード対象とするオーディオエレメントの組み合わせを選択し、選択した組み合わせのオーディオエレメントのみをサーバから取得してデコードすることができる。また、デコーダが一旦、入力ビットストリームを取得しておき、その入力ビットストリームから必要なオーディオエレメントを選択してデコードすることもできる。

50

【 0 0 6 9 】

なお、配置パターン(1)の例において、入力ビットストリームの転送時の速度、つまり転送ビットレートごとに入力ビットストリームが用意され、サーバに配置されるようにしてもよい。

【 0 0 7 0 】

また、配置パターン(2)では図1に示した入力ビットストリームが、オーディオ要素の組み合わせごとに分割され、例えば図2に示すように分割により得られた各組み合わせのビットストリームがサーバに配置される。

【 0 0 7 1 】

なお、図2では、図1における場合と同様に、1つの長方形は1つのオーディオ要素、すなわちSCEまたはCPEを表している。

10

【 0 0 7 2 】

この例では、サーバ上には、矢印A11に示す組み合わせCM(1)の構成要素からなるビットストリーム、矢印A12に示す組み合わせCM(2)の構成要素からなるビットストリーム、および矢印A13に示す組み合わせCM(3)の構成要素からなるビットストリームが配置されている。

【 0 0 7 3 】

さらに、サーバ上には矢印A14に示す組み合わせCM(4)の構成要素からなるビットストリーム、矢印A15に示す組み合わせCM(5)の構成要素からなるビットストリーム、および矢印A16に示す組み合わせCM(6)の構成要素からなるビットストリームが配置されている。

20

【 0 0 7 4 】

この場合、デコーダはサーバ等から取得した情報などから、デコード対象とするオーディオ要素の組み合わせを選択し、選択した組み合わせのオーディオ要素をサーバから取得してデコードする。なお、配置パターン(2)の例においても分割された入力ビットストリームが転送ビットレートごとに用意され、サーバに配置されてもよい。

【 0 0 7 5 】

また、配置パターン(1)に示した1つの入力ビットストリームが、サーバからデコーダ側に送信されるときに分割されて、要求された組み合わせのオーディオ要素のみからなるビットストリームが送信されるようにしてもよい。

【 0 0 7 6 】

30

このようにデコード対象となるオーディオ要素の組み合わせのみを取得するようにすれば、転送ビットレートを削減することができる。

【 0 0 7 7 】

例えばデコード対象となるオーディオ要素の組み合わせのみデコーダ側により取得される場合には、入力ビットストリームに格納されているメタデータ等に基づいて、オーディオ要素の組み合わせが選択されるようにすることができる。ここで、オーディオ要素の組み合わせの選択は、例えば入力ビットストリームにメタデータとして格納されている、入力ビットストリームについて取得可能なオーディオ要素の各組み合わせを示す情報などに基づいて行われる。

【 0 0 7 8 】

40

これに加え、デコーダが、デコード対象となる組み合わせのオーディオ要素のなかの不要なオーディオ要素を取得しないようにすれば、さらに転送ビットレートを削減することができる。例えば、そのような不要なオーディオ要素は、ユーザにより指定されるようにしてもよいし、入力ビットストリームに格納されているメタデータ等に基づいて選択されてもよい。

【 0 0 7 9 】

特に、メタデータにより不要なオーディオ要素を選択する場合には、各オブジェクトの優先度(重要度)、つまりオーディオ要素の優先度を示す優先度情報に基づいて選択が行われるようにしてもよい。ここで、優先度情報は、その優先度情報の値が大きいほど、オーディオ要素の優先度が高く、重要な要素であることを示して

50

いる。

【 0 0 8 0 】

例えば3D Audio規格では、オブジェクト音源ごと、時間フレームごとに、オブジェクトの優先度情報 (object_priority) が入力ビットストリーム内、より詳細にはEXTエレメントの内部で定義されている。特に3D Audio規格では、EXTエレメントはSCEやCPEと同じシンタックスレイヤに定義されている。

【 0 0 8 1 】

そこで、コンテンツを再生するクライアント側、つまりデコーダ側は、このオブジェクトの優先度情報を読み取り、その値が、クライアント側で予め定められた閾値以下であるオブジェクトのオーディオエレメントについては、転送しないようにサーバに対して命令を出す。これにより、サーバから転送される入力ビットストリーム (データ) に、命令により指定したオブジェクト音源のオーディオエレメント (SCE) が含まれないようにすることができ、転送データのビットレートを削減することが可能となる。

10

【 0 0 8 2 】

このような優先度情報を利用した転送ビットレートの削減を実現するには、オブジェクトの優先度情報の先読みと、修正規定したMinimum decoder input bufferサイズでデコードを可能とするための転送ビットレート調整処理との2つの処理が必要となる。

【 0 0 8 3 】

(優先度情報の先読み)

クライアント (デコーダ) がサーバに特定のオブジェクトのオーディオエレメントの非転送を要求するためには、クライアントは、オブジェクト音源のオーディオエレメントが転送される前に、オブジェクトの優先度情報を読み取らなければならない。

20

【 0 0 8 4 】

上述したように3D Audio規格では、各オブジェクトの優先度情報はEXTエレメントに含まれている。したがって、オブジェクトの優先度情報を先読みするために、例えばEXTエレメントを以下の配置位置A(1)や配置位置A(2)のような配置とすればよい。なお、これらの例に限らず、優先度情報の先読みが可能であれば、EXTエレメント、すなわち優先度情報の配置位置はどのような位置とされてもよいし、どのようにして取得されるようにしてもよい。

【 0 0 8 5 】

(配置位置A(1))

EXTエレメントを1つのファイルとしてもち、クライアントはデコードの開始時に全フレームもしくは先読み数フレーム分のオブジェクトの優先度情報を読み取っていく

30

【 0 0 8 6 】

(配置位置A(2))

EXTエレメントをビットストリーム内の各フレームの先頭に配置し、クライアントは時間フレームごとにオブジェクトの優先度情報を読み取っていく

【 0 0 8 7 】

例えば配置位置A(1)では、例えば図3の矢印A21に示すようにコンテンツを構成する全オブジェクト、つまり全オブジェクトのオーディオエレメントの時間フレームごとの優先度情報が格納された1つのファイル (EXTエレメント) がサーバに記録されている。

40

【 0 0 8 8 】

図3では、文字「EXT(1)」が記された1つの長方形が、1つのEXTエレメントを表している。この例では、クライアント (デコーダ) は、デコード開始前の任意のタイミングでサーバからEXTエレメントを取得し、非転送とするオーディオエレメントを選択する。

【 0 0 8 9 】

また、例えば配置位置A(2)では、矢印A22に示すように、入力ビットストリームの各フレームの先頭にEXTエレメントが配置されてサーバに記録されている。ここで、EXTエレメント以下、つまり図中、下側に配置されている各長方形は、図1における場合と同様に1つのオーディオエレメント (SCEまたはCPE) を表している。

50

【 0 0 9 0 】

この例では、サーバに記録されている入力ビットストリームは、図 1 に示した構成の先頭にさらにEXTエレメントが配置されたものとなっている。

【 0 0 9 1 】

したがって、この場合には、クライアント（デコーダ）は、まず対象となる時間フレームについて、入力ビットストリームのEXTエレメントを受信して優先度情報を読み出す。そして、クライアントは、優先度情報に基づいて、非転送とするオーディオエレメントを選択し、そのオーディオエレメントを非転送とする旨の要求（命令）をサーバに行うことになる。

【 0 0 9 2 】

（転送ビットレートの調整処理）

続いて、修正規定したMinimum decoder input bufferサイズでデコードを可能とするための転送ビットレート調整処理について説明する。

【 0 0 9 3 】

例えばエンコーダでは、上述したようにサーバ上に配置される入力ビットストリームの各オーディオエレメントについて、修正規定したMinimum decoder input bufferサイズでデコードできるように、オーディオエレメント（符号化データ）のビット量の調整が行われる。

【 0 0 9 4 】

したがって、デコーダ側において、どの組み合わせのオーディオエレメントが選択されたときでも、例えば図 4 に示すように、必要バッファサイズのバッファに入力ビットストリームを順次、格納しながらデコードを行ってもアンダーフローおよびオーバーフローは発生しない。

【 0 0 9 5 】

なお、図 4 において縦軸はデコーダ側のバッファ内に格納されている各時刻における入力ビットストリームのデータ量を示しており、横軸は時間を示している。また、図中、折れ線の傾きは、入力ビットストリームの転送ビットレートを示しており、転送ビットレートは、例えば入力ビットストリームの伝送路の平均ビットレートなどとされる。

【 0 0 9 6 】

この例ではdata[1]乃至data[4]は、各時間フレーム分のオーディオエレメントがサーバから受信されてバッファに格納される期間を表しており、a1、b1、b2、c1、c2、d1、およびd2は、それぞれ所定期間内にバッファに格納されたデータ量を示している。また、縦軸におけるBFZは、Minimum decoder input bufferサイズを示している。

【 0 0 9 7 】

図 4 では、デコーダのバッファに、受信したオーディオエレメントがBFZ分だけ格納されると、最初の時間フレームのオーディオエレメントのデコードが開始され、その後、各時間フレームのオーディオエレメントのデコードが一定時間間隔で行われる。

【 0 0 9 8 】

例えば時刻t1では、a1分のデータ量である先頭時間フレームのデータ、つまり先頭時間フレームの各オーディオエレメントがバッファから読み出されてデコードされている。同様に、時刻t2乃至時刻t4のそれぞれにおいて、2番目乃至4番目の時間フレームの各オーディオエレメントがバッファから読み出されてデコードが行われている。

【 0 0 9 9 】

このとき、バッファ内に格納されているオーディオエレメントのデータ量は、どの時刻においても0以上、かつBFZ以下となっており、アンダーフローもオーバーフローも生じていない。したがって、コンテンツが時間的に連続して途切れることなく再生されることになる。

【 0 1 0 0 】

しかし、符号化データのビット量を調整しながらのエンコードは、どのオーディオエレメントの組み合わせが選択された場合でも、選択された組み合わせを構成する全オーディ

10

20

30

40

50

オエレメントがデコードされるという前提で行われたものとなっている。つまり、優先度情報等に基づいて選択された、組み合わせを構成する全オーディオエレメントのうちのいくつかをデコードしない場合については、考慮されていない。

【0101】

そのため、デコード対象とする組み合わせのオーディオエレメントのうちの、いくつかのオブジェクトのオーディオエレメントがデコードされない場合には、エンコーダ側での時間フレームごとのビット量の調整と、デコーダ側での各時間フレームでのデコードによるビットの消費量との整合が取れなくなってしまう。そうすると、場合によってはデコーダ側においてオーバーフローやアンダーフローが生じ、上述の修正規定したMinimum decoder input bufferサイズでのデコードができなくなってしまう。

10

【0102】

そこで、本技術では、エンコーダ側でのビット量の調整と、デコーダ側でのビット消費量との整合をとり、上述の修正規定したMinimum decoder input bufferサイズでのデコードを行うことができるようにするために、以下の転送ビットレート調整処理RMT(1)または転送ビットレート調整処理RMT(2)が行われる。

【0103】

(転送ビットレート調整処理RMT(1))

時間フレームごとに転送データに含めないオブジェクトのオーディオエレメントのサイズを読み取り、そのサイズから転送を停止する時間を算出し、その時間だけ転送を停止

【0104】

20

(転送ビットレート調整処理RMT(2))

時間フレームごとに転送データに含めないオブジェクトのオーディオエレメントのサイズを読み取り、そのサイズから、転送対象となる時間フレームの転送レートを調整

【0105】

転送ビットレート調整処理RMT(1)では、例えば図5に示すように所定の時間だけ入力ビットストリームの転送を停止させることにより、実質的に転送ビットレートを変更する。

【0106】

なお、図5において縦軸はデコーダ側のバッファ内に格納されている各時刻における入力ビットストリームのデータ量を示しており、横軸は時間を示している。また、図5において図4における場合と対応する部分には、同じ文字等が記されており、その説明は適宜省略する。

30

【0107】

この例では、図4においてa1、b1、b2、c1、d1、およびd2で表されていたデータ量が、それぞれa1'、b1'、b2'、c1'、d1'、およびd2'となっている。

【0108】

例えば図4では先頭の時間フレームのデコード対象のオーディオエレメントの合計データ量がa1であったものが、図5では、所定のオブジェクトのオーディオエレメントのデコードが行われないためにa1'となっている。

【0109】

そのため、先頭フレームでデコードしないとされた、つまり優先度情報等により選択されたオブジェクトのオーディオエレメントのサイズ(データ量)と、入力ビットストリームの転送ビットレート、つまり図中の折れ線の傾きとから定まる時間の期間T11だけ、入力ビットストリームの転送が停止されている。

40

【0110】

同様に、先頭時間フレーム以降の各時間フレームについても、それぞれ期間T12乃至期間T14で入力ビットストリームの転送が一時的に停止されている。

【0111】

このような転送ビットレート制御は、サーバ側で実現してもよいし、デコーダ側でバッファ制御を行うことで実現するようにしてもよい。

【0112】

50

サーバ側でビットレート制御を行うときには、例えばデコーダがサーバ側に、入力ビットストリームの一時的な転送停止を指示するようにしてもよいし、サーバが転送停止時間を算出し、入力ビットストリームの転送を一時的に停止するようにしてもよい。

【0113】

また、デコーダ側でのバッファ制御により転送ビットレート制御を行う場合、例えばデコーダは、受信した入力ビットストリームを蓄積しておくシステムバッファから、デコードのためのオーディオバッファへのオーディオエレメント転送時にオーディオエレメントの転送（格納）の一時的な停止を行う。

【0114】

ここで、システムバッファは、例えばコンテンツを構成する音声の入力ビットストリームだけでなく、コンテンツを構成する映像の入力ビットストリームなども蓄積されるバッファなどとされる。また、オーディオバッファは、Minimum decoder input bufferサイズ以上のバッファサイズの確保が必要となるデコード用のバッファである。

【0115】

一方、転送ビットレート調整処理RMT(2)では、例えば図6に示すように入力ビットストリームの転送ビットレートを可変させる。

【0116】

なお、図6において縦軸はデコーダ側のオーディオバッファ内に格納されている各時刻における入力ビットストリームのデータ量を示しており、横軸は時間を示している。また、図6において図4または図5における場合と対応する部分には、同じ文字等が記されており、その説明は適宜省略する。

【0117】

例えば図4では先頭の時間フレームのデコード対象のオーディオエレメントの合計データ量が $a1$ であったものが、図6では、所定のオブジェクトのオーディオエレメントのデコードが行われないために $a1'$ となっている。

【0118】

そのため、先頭フレーム分のオーディオエレメントの取得後、時刻 $t1$ までの期間において、先頭フレームでデコードしないとされた、つまり優先度情報等により選択されたオブジェクトのオーディオエレメントのサイズと、入力ビットストリームの転送ビットレートとから定まる新たな転送ビットレートで、オーディオエレメントの転送が行われている。

【0119】

同様に、それ以降の期間でも入力ビットストリームの転送が、新たに算出された転送ビットレートで行われている。例えば時刻 $t2$ から時刻 $t3$ までの期間では、時刻 $t3$ においてオーディオバッファ内に格納されているオーディオエレメントの合計データ量が、図5の例の時刻 $t3$ における場合と同じとなるように、新たな転送ビットレートを定めればよい。

【0120】

このような転送ビットレート制御は、サーバ側で実現してもよいし、デコーダ側でバッファ制御を行うことで実現するようにしてもよい。

【0121】

サーバ側でビットレート制御を行うときには、例えばデコーダがサーバ側に、入力ビットストリームの新たな転送ビットレートを指示するようにしてもよいし、サーバが新たな転送ビットレートを算出するようにしてもよい。

【0122】

また、デコーダ側でのバッファ制御により転送ビットレート制御を行う場合、例えばデコーダは、新たな転送ビットレートを算出し、その新たな転送ビットレートで、システムバッファからオーディオバッファへとオーディオエレメントを転送する。

【0123】

ここで、転送ビットレート調整処理RMT(1)や転送ビットレート調整処理RMT(2)を行う場合、デコード対象としないオブジェクトのオーディオエレメントのサイズを先読みする必要がある。そこで本技術では、各オーディオエレメントのサイズを示すサイズ情報が、例

10

20

30

40

50

えば以下のサイズ情報配置SIL(1)乃至サイズ情報配置SIL(3)に示す何れかの配置とされる。なお、サイズ情報の配置は、先読み可能な配置であれば、どのような配置とされてもよい。

【0124】

(サイズ情報配置SIL(1))

サイズ情報を1つのファイルとしてもち、クライアントはデコードの開始時に全フレームもしくは先読み数フレーム分の各オーディオエレメントのサイズを読み取っていく

【0125】

(サイズ情報配置SIL(2))

サイズ情報を入力ビットストリーム内の各フレームの先頭に配置し、クライアントは時間フレームごとにサイズ情報を読み取っていく

【0126】

(サイズ情報配置SIL(3))

各オーディオエレメントの先頭にサイズ情報を定義し、クライアントはオーディオエレメントごとにサイズ情報を読み取っていく

【0127】

サイズ情報配置SIL(1)では、例えば図7の矢印A31に示すようにコンテンツを構成する全オーディオエレメントの時間フレームごとのサイズ情報が格納された1つのファイルがサーバに記録されている。なお、図7において、文字「Size」が記された楕円がサイズ情報を表している。

【0128】

この例では、例えばクライアント(デコーダ)は、デコード開始前の任意のタイミングでサーバからサイズ情報を取得し、転送ビットレート調整処理RMT(1)や転送ビットレート調整処理RMT(2)を行う。

【0129】

また、例えばサイズ情報配置SIL(2)では、矢印A32に示すように、入力ビットストリームの各フレームの先頭にサイズ情報が配置されてサーバに記録されている。ここで、サイズ情報以下に配置されている各長方形は、図3における場合と同様に1つのオーディオエレメント(SCEまたはCPE)またはEXTエレメントを表している。

【0130】

この例では、サーバに記録されている入力ビットストリームは、図3の矢印A22に示した構成の先頭にさらにサイズ情報が配置されたものとなっている。

【0131】

したがって、この場合には、例えばクライアント(デコーダ)は、まず入力ビットストリームのサイズ情報やEXTエレメントを受信して、非転送とするオーディオエレメントを選択したり、その選択に応じて転送ビットレート調整処理RMT(1)や転送ビットレート調整処理RMT(2)を行ったりする。

【0132】

さらに、例えばサイズ情報配置SIL(3)では、矢印A33に示すように、各オーディオエレメント内の先頭部分にサイズ情報が配置されている。したがって、この場合には、例えばクライアント(デコーダ)は、各オーディオエレメントからサイズ情報を読み出して、転送ビットレート調整処理RMT(1)や転送ビットレート調整処理RMT(2)を行う。

【0133】

なお、以上においては、オブジェクトのオーディオエレメントを非転送とする例について説明したが、オブジェクトに限らず、各組み合わせを構成するどのオーディオエレメントを非転送とする場合でも、上述したオブジェクトの例と同様に、Minimum decoder input bufferサイズでのデコードが可能となる。

【0134】

以上のように、入力ビットストリームのなかのデコード対象としない不要なオーディオエレメントをメタデータ等に基づいて選択し、転送されないようにすることで、転送ビッ

10

20

30

40

50

トレートを削減することができる。

【 0 1 3 5 】

また、入力ビットストリームを構成する任意のオーディオエレメントをデコード対象としないようにする場合に、適切に転送ビットレートを調整することでMinimum decoder input bufferサイズでのデコードが可能となる。

【 0 1 3 6 】

コンテンツ配信システムの構成例

次に、以上において説明した本技術を適用した具体的な実施の形態について説明する。

【 0 1 3 7 】

以下では、本技術をMPEG-DASHに準ずるコンテンツ配信システムに適用した場合を例として説明する。そのような場合、本技術を適用したコンテンツ配信システムは、例えば図8に示すように構成される。

10

【 0 1 3 8 】

図8に示すコンテンツ配信システムは、サーバ11およびクライアント12から構成され、これらのサーバ11とクライアント12は、インターネットなどの有線や無線の通信網を介して相互に接続されている。

【 0 1 3 9 】

サーバ11には、例えば複数の転送ビットレートごとに、図1に示した入力ビットストリームや、図2に示した、入力ビットストリームをオーディオエレメントの組み合わせごとに分割して得られたビットストリームが記録されている。

20

【 0 1 4 0 】

また、サーバ11には、単独の1つのファイルとして、または各入力ビットストリームや分割された入力ビットストリームのフレームの先頭部分に配置されて、図3を参照して説明したEXTエレメントが記録されている。さらに、サーバ11には、単独の1つのファイルとして、各入力ビットストリームや分割された入力ビットストリームのフレームの先頭部分に配置されて、または各オーディオエレメント内の先頭部分に配置されて、図7を参照して説明したサイズ情報が記録されている。

【 0 1 4 1 】

サーバ11は、クライアント12からの要求に応じて、入力ビットストリームやEXTエレメント、サイズ情報などをクライアント12に送信する。

30

【 0 1 4 2 】

また、クライアント12は、サーバ11から入力ビットストリームを受信して、入力ビットストリームをデコードおよび再生することで、コンテンツをストリーミング再生する。

【 0 1 4 3 】

なお、入力ビットストリームの受信にあたっては、その入力ビットストリーム全部を受信するにしてもよいし、入力ビットストリームの分割された一部分のみを受信するにしてもよい。以下では、入力ビットストリームの全部と一部分とを特に区別する必要がない場合には、単に入力ビットストリームとも称することとする。

【 0 1 4 4 】

クライアント12は、ストリーミング制御部21、アクセス処理部22、およびデコーダ23を有している。

40

【 0 1 4 5 】

ストリーミング制御部21は、クライアント12全体の動作を制御する。例えばストリーミング制御部21は、サーバ11からEXTエレメント、サイズ情報、その他の制御情報を受信して、必要に応じてアクセス処理部22やデコーダ23に供給したり、受信した情報に基づいてストリーミング再生の制御を行ったりする。

【 0 1 4 6 】

アクセス処理部22は、デコーダ23等の要求に応じて、サーバ11に対して所定の転送ビットレートでの所定の組み合わせのオーディオエレメントの入力ビットストリームの

50

送信を要求したり、サーバ 11 から送信されてきた入力ビットストリームを受信してデコーダ 23 に供給したりする。デコーダ 23 は、必要に応じてストリーミング制御部 21 やアクセス処理部 22 と情報の授受を行いながら、アクセス処理部 22 から供給された入力ビットストリームをデコードし、図示せぬスピーカ等に出力する。

【0147】

デコーダの構成例

続いて、図 8 に示したデコーダ 23 のより詳細な構成について説明する。例えばデコーダ 23 は、より詳細には図 9 に示すように構成される。

【0148】

図 9 に示すデコーダ 23 は、取得部 71、バッファサイズ算出部 72、選択部 73、抽出部 74、オーディオバッファ 75、復号部 76、および出力部 77 を有している。

10

【0149】

この例では、アクセス処理部 22 から取得部 71 には、例えば図 1 に示した構成の、所定の転送ビットレートの入力ビットストリームが供給される。なお、アクセス処理部 22 がサーバ 11 から、どの転送ビットレートの入力ビットストリームを受信するかは、例えばアクセス処理部 22 等が通信網の状況等から、時間フレームごとに選択することができる。つまり、時間フレームごとに転送ビットレートを変更することができる。

【0150】

取得部 71 は、アクセス処理部 22 から入力ビットストリームを取得してバッファサイズ算出部 72 および抽出部 74 に供給する。バッファサイズ算出部 72 は、取得部 71 から供給された入力ビットストリームに基づいて、各オーディオエレメントの組み合わせごとに必要バッファサイズを算出し、選択部 73 に供給する。

20

【0151】

選択部 73 は、バッファサイズ算出部 72 から供給された各オーディオエレメントの組み合わせの必要バッファサイズと、デコーダ 23、すなわちオーディオバッファ 75 の許容メモリサイズとを比較して、デコード対象とするオーディオエレメントの組み合わせを選択し、その選択結果を抽出部 74 に供給する。

【0152】

抽出部 74 は、選択部 73 から供給された選択結果に基づいて、取得部 71 から供給された入力ビットストリームから、選択された組み合わせのオーディオエレメントを抽出し、オーディオバッファ 75 に供給する。

30

【0153】

オーディオバッファ 75 は、予め定められた所定の許容メモリサイズのバッファであり、抽出部 74 から供給されたデコード対象となるオーディオエレメントを一時的に保持し、復号部 76 に供給する。復号部 76 は、オーディオバッファ 75 から時間フレーム単位でオーディオエレメントを読み出してデコード（復号）するとともに、デコードにより得られたオーディオ信号に基づいて、所定チャネル構成のオーディオ信号を生成し、出力部 77 に供給する。出力部 77 は、復号部 76 から供給されたオーディオ信号を後段のスピーカ等に出力する。

【0154】

40

復号処理の説明

続いて、図 9 に示したデコーダ 23 により行われる復号処理について説明する。例えば復号処理は、時間フレームごとに行われる。

【0155】

ステップ S11 において、取得部 71 は、アクセス処理部 22 から入力ビットストリームを取得してバッファサイズ算出部 72 および抽出部 74 に供給する。

【0156】

ステップ S12 において、バッファサイズ算出部 72 は、取得部 71 から供給された入力ビットストリームに基づいて、オーディオエレメントの組み合わせごとに、必要バッファサイズを算出し、選択部 73 に供給する。

50

【 0 1 5 7 】

具体的にはバッファサイズ算出部 7 2 は、算出対象のオーディオエレメントの組み合わせについて、その組み合わせを構成するSCEの数と、CPEの数の2倍との和をNCCとし、NCCと6144との積を必要バッファサイズ (Minimum decoder input bufferサイズ) として算出する。

【 0 1 5 8 】

なお、入力ビットストリームに格納されているオーディオエレメントの選択可能な組み合わせは、メタデータ等を参照することで特定することができる。また、入力ビットストリームに各組み合わせについての必要バッファサイズを示す情報が格納されている場合には、バッファサイズ算出部 7 2 は、入力ビットストリームから、必要バッファサイズを示す情報を読み出して選択部 7 3 に供給する。

10

【 0 1 5 9 】

ステップ S 1 3 において、選択部 7 3 は、バッファサイズ算出部 7 2 から供給された必要バッファサイズに基づいて、オーディオエレメントの組み合わせを選択し、その選択結果を抽出部 7 4 に供給する。

【 0 1 6 0 】

すなわち、選択部 7 3 は、各オーディオエレメントの組み合わせの必要バッファサイズと、デコーダ 2 3、すなわちオーディオバッファ 7 5 の許容メモリサイズとを比較して、バッファサイズ条件を満たす組み合わせの1つをデコード対象として選択する。そして、選択部 7 3 は、その選択結果を抽出部 7 4 に供給する。

20

【 0 1 6 1 】

ステップ S 1 4 において、抽出部 7 4 は、取得部 7 1 から供給された入力ビットストリームから、選択部 7 3 から供給された選択結果により示される組み合わせのオーディオエレメントを抽出し、オーディオバッファ 7 5 に供給する。

【 0 1 6 2 】

ステップ S 1 5 において、復号部 7 6 は、オーディオバッファ 7 5 から1時間フレーム分のオーディオエレメントを読み出して、そのオーディオエレメント、すなわちオーディオエレメントに格納されている符号化データをデコードする。

【 0 1 6 3 】

また、復号部 7 6 は、デコードにより得られたオーディオ信号に基づいて、所定チャンネル構成のオーディオ信号を生成し、出力部 7 7 に供給する。例えば復号部 7 6 は、オブジェクトのオーディオ信号をスピーカに対応する各チャンネルに割り当てるなどして、目的とするチャンネル構成の各チャンネルのオーディオ信号を生成する。

30

【 0 1 6 4 】

ステップ S 1 6 において、出力部 7 7 は、復号部 7 6 から供給されたオーディオ信号を後段のスピーカ等に出し、復号処理は終了する。

【 0 1 6 5 】

以上のようにしてデコーダ 2 3 は、自身の許容メモリサイズと必要バッファサイズに応じて、オーディオエレメントの組み合わせを選択し、デコードを行う。これにより、ハード規模の異なる様々な機器で入力ビットストリームをデコードすることができる。

40

【 0 1 6 6 】

第 2 の実施の形態

デコーダの構成例

また、図 9 に示したデコーダ 2 3 では、オーディオエレメントの組み合わせを選択する例について説明したが、さらにデコーダ 2 3 において、優先度情報等のメタデータに基づいて、デコード対象としない不要なオーディオエレメントを選択するようにしてもよい。そのような場合、デコーダ 2 3 は、例えば図 1 1 に示すように構成される。なお、図 1 1 において図 9 における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

【 0 1 6 7 】

50

図 1 1 に示すデコーダ 2 3 は、取得部 7 1、バッファサイズ算出部 7 2、選択部 7 3、抽出部 7 4、システムバッファ 1 1 1、オーディオバッファ 7 5、復号部 7 6、および出力部 7 7 を有している。図 1 1 に示すデコーダ 2 3 の構成は、新たにシステムバッファ 1 1 1 が設けられている点で図 9 のデコーダ 2 3 と異なり、その他の点では図 9 のデコーダ 2 3 と同じ構成とされる。

【 0 1 6 8 】

図 1 1 に示すデコーダ 2 3 には、例えば図 1 に示した構成の、所定の転送ビットレートの入力ビットストリームが供給される。

【 0 1 6 9 】

また、取得部 7 1 は、サーバ 1 1 から EXT エlement およびサイズ情報を取得し、EXT エlement をバッファサイズ算出部 7 2 を介して選択部 7 3 に供給するとともに、サイズ情報を抽出部 7 4 を介してシステムバッファ 1 1 1 に供給する。

10

【 0 1 7 0 】

例えば図 3 の矢印 A21 に示したように EXT エlement が単独でサーバ 1 1 に記録されている場合、取得部 7 1 は、デコード開始前の任意のタイミングでストリーミング制御部 2 1 を介して、サーバ 1 1 から EXT エlement を取得する。

【 0 1 7 1 】

また、例えば図 3 の矢印 A22 に示したように EXT エlement が入力ビットストリームのフレーム先頭に配置されている場合には、取得部 7 1 は、その入力ビットストリームをバッファサイズ算出部 7 2 に供給する。そして、バッファサイズ算出部 7 2 は、入力ビットストリームから EXT エlement を読み出して選択部 7 3 に供給する。

20

【 0 1 7 2 】

なお、以下では、図 3 の矢印 A21 に示したように EXT エlement が単独でサーバ 1 1 に記録されており、予め EXT エlement が選択部 7 3 に供給されているものとして説明を続ける。

【 0 1 7 3 】

さらに、例えば図 7 の矢印 A31 に示したようにサイズ情報が単独でサーバ 1 1 に記録されている場合、取得部 7 1 は、デコード開始前の任意のタイミングでストリーミング制御部 2 1 を介して、サーバ 1 1 からサイズ情報を取得する。

【 0 1 7 4 】

30

また、例えば図 7 の矢印 A32 や矢印 A33 に示したように、サイズ情報が入力ビットストリームの各フレームの先頭や、各オーディオ Element 内の先頭に配置されている場合、取得部 7 1 は、その入力ビットストリームを抽出部 7 4 に供給する。そして、抽出部 7 4 は、入力ビットストリームからサイズ情報を読み出してシステムバッファ 1 1 1 に供給する。

【 0 1 7 5 】

なお、以下では、図 7 の矢印 A31 に示したようにサイズ情報が単独でサーバ 1 1 に記録されており、予めサイズ情報がシステムバッファ 1 1 1 に供給されているものとして説明を続ける。

【 0 1 7 6 】

40

選択部 7 3 は、バッファサイズ算出部 7 2 から供給された必要バッファサイズに基づいて、オーディオ Element の組み合わせを選択する。さらに選択部 7 3 は、バッファサイズ算出部 7 2 から供給された EXT Element に含まれる優先度情報に基づいて、選択した組み合わせを構成する各オーディオ Element のなかから、デコード対象としない不要なオーディオ Element、つまり非転送とするオーディオ Element を選択する。

【 0 1 7 7 】

なお、不要なオーディオ Element は、オブジェクトのオーディオ Element であってもよいし、それ以外のオーディオ Element であってもよい。

【 0 1 7 8 】

選択部 7 3 は、組み合わせの選択結果と、不要なオーディオ Element の選択結果とを

50

抽出部 7 4 に供給する。

【 0 1 7 9 】

抽出部 7 4 は、選択部 7 3 から供給された選択結果に基づいて、取得部 7 1 から供給された入力ビットストリームから、選択された組み合わせを構成し、かつ不要であるとされなかったオーディオエレメントを抽出し、システムバッファ 1 1 1 に供給する。

【 0 1 8 0 】

システムバッファ 1 1 1 は、予め抽出部 7 4 から供給されたサイズ情報に基づいて、上述した転送ビットレート調整処理RMT(1)または転送ビットレート調整処理RMT(2)によりバッファ制御を行って、抽出部 7 4 から供給されたオーディオエレメントをオーディオバッファ 7 5 に供給する。なお、以下では、転送ビットレート調整処理RMT(1)が行われるものとして説明を続ける。

10

【 0 1 8 1 】

復号処理の説明

次に図 1 2 のフローチャートを参照して、図 1 1 に示したデコーダ 2 3 により行われる復号処理について説明する。なお、ステップ S 4 1 およびステップ S 4 2 の処理は、図 1 0 のステップ S 1 1 およびステップ S 1 2 の処理と同様であるので、その説明は省略する。

【 0 1 8 2 】

ステップ S 4 3 において、選択部 7 3 は、バッファサイズ算出部 7 2 から供給された必要バッファサイズおよびEXTエレメントに含まれる優先度情報に基づいて、オーディオエレメントの組み合わせと、不要なオーディオエレメントとを選択する。

20

【 0 1 8 3 】

例えば選択部 7 3 は、図 1 0 のステップ S 1 3 と同様の処理を行って、オーディオエレメントの組み合わせを選択する。さらに、選択部 7 3 は、選択した組み合わせのオーディオエレメントのうち、優先度情報の値が所定の閾値以下であるオーディオエレメントをデコード対象としない不要なオーディオエレメントとして選択する。

【 0 1 8 4 】

選択部 7 3 は、組み合わせの選択結果と、不要なオーディオエレメントの選択結果とを抽出部 7 4 に供給する。

【 0 1 8 5 】

30

ステップ S 4 4 において、抽出部 7 4 は、選択部 7 3 から供給された選択結果に基づいて、取得部 7 1 から供給された入力ビットストリームから、選択された組み合わせを構成し、かつ不要であるとされなかったオーディオエレメントを抽出し、システムバッファ 1 1 1 に供給する。また、抽出部 7 4 は、選択部 7 3 により選択された、デコード対象としない不要なオーディオエレメントを示す情報をシステムバッファ 1 1 1 に供給する。

【 0 1 8 6 】

ステップ S 4 5 において、システムバッファ 1 1 1 は、予め抽出部 7 4 から供給されたサイズ情報、および抽出部 7 4 から供給された不要なオーディオエレメントを示す情報に基づいて、バッファ制御を行う。

【 0 1 8 7 】

40

具体的には、システムバッファ 1 1 1 は、抽出部 7 4 から供給された情報により示されるオーディオエレメントのサイズ情報に基づいて、転送を停止させるべき時間を算出する。そして、システムバッファ 1 1 1 は、適切なタイミングで、算出した時間だけオーディオエレメントのオーディオバッファ 7 5 への転送（格納）を停止させながら、抽出部 7 4 から供給されたオーディオエレメントをオーディオバッファ 7 5 に転送する。

【 0 1 8 8 】

バッファ制御が行われると、その後、ステップ S 4 6 およびステップ S 4 7 の処理が行われて復号処理は終了するが、これらの処理は図 1 0 のステップ S 1 5 およびステップ S 1 6 の処理と同様であるので、その説明は省略する。

【 0 1 8 9 】

50

以上のようにしてデコーダ２３は、オーディオエレメントの組み合わせを選択するとともに、優先度情報に基づいてデコード対象としないオーディオエレメントを選択する。これにより、ハード規模の異なる様々な機器で入力ビットストリームをデコードすることができる。また、バッファ制御により実質的な転送ビットレート制御を行うことで、Minimum decoder input bufferサイズでのデコードが可能となる。

【０１９０】

第３の実施の形態

デコーダの構成例

また、以上においては取得した入力ビットストリームからデコード対象とする組み合わせのオーディオエレメントを抽出する例について説明したが、選択した組み合わせのオーディオエレメントをサーバ１１から取得するようにしてもよい。そのような場合、デコーダ２３は、例えば図１３に示す構成とされる。なお、図１３において、図９における場合と対応する部分には同一の符号を付してあり、その説明は省略する。

10

【０１９１】

図１３に示すデコーダ２３は、通信部１４１、バッファサイズ算出部７２、選択部７３、要求部１４２、オーディオバッファ７５、復号部７６、および出力部７７を有している。

【０１９２】

図１３に示すデコーダ２３の構成は、取得部７１および抽出部７４が設けられておらず、新たに通信部１４１および要求部１４２が設けられている点で図９のデコーダ２３の構成と異なる。

20

【０１９３】

通信部１４１は、アクセス処理部２２やストリーミング制御部２１を介してサーバ１１との通信を行う。例えば通信部１４１は、サーバ１１から取得可能なオーディオエレメントの組み合わせを示す情報を受信してバッファサイズ算出部７２に供給したり、要求部１４２から供給された、分割された入力ビットストリームの一部分の送信要求をサーバ１１に送信したりする。また、通信部１４１は、送信要求に応じてサーバ１１から送信されてきた、分割された入力ビットストリームの一部分を受信してオーディオバッファ７５に供給する。

【０１９４】

ここで、サーバ１１から取得可能なオーディオエレメントの組み合わせを示す情報は、例えば入力ビットストリームのメタデータとして、入力ビットストリーム内に格納された状態で、または単独のファイルとしてサーバ１１に記録されている。なお、ここではサーバ１１から取得可能なオーディオエレメントの組み合わせを示す情報は、単独のファイルとしてサーバ１１に記録されているものとする。

30

【０１９５】

要求部１４２は、選択部７３から供給された、デコード対象とするオーディオエレメントの組み合わせの選択結果に基づいて、選択された組み合わせのオーディオエレメントからなるビットストリーム、つまり分割された入力ビットストリームの一部分の送信要求を通信部１４１に供給する。

40

【０１９６】

復号処理の説明

次に、図１４のフローチャートを参照して、図１３に示すデコーダ２３により行われる復号処理について説明する。

【０１９７】

ステップＳ７１において、通信部１４１は、サーバ１１から取得可能なオーディオエレメントの組み合わせを示す情報を受信してバッファサイズ算出部７２に供給する。

【０１９８】

すなわち、通信部１４１は、取得可能なオーディオエレメントの組み合わせを示す情報の送信要求を、ストリーミング制御部２１を介してサーバ１１に送信する。また、通信部

50

141は、その送信要求に応じてサーバ11から送信されてきたオーディオエレメントの組み合わせを示す情報を、ストリーミング制御部21を介して受信して、バッファサイズ算出部72に供給する。

【0199】

ステップS72において、バッファサイズ算出部72は、通信部141から供給された、サーバ11から取得可能なオーディオエレメントの組み合わせを示す情報に基づいて、その情報により示されるオーディオエレメントの組み合わせごとに、必要バッファサイズを算出し、選択部73に供給する。ステップS72では、図10のステップS12と同様の処理が行われる。

【0200】

ステップS73において、選択部73は、バッファサイズ算出部72から供給された必要バッファサイズに基づいて、オーディオエレメントの組み合わせを選択し、その選択結果を要求部142に供給する。ステップS73では、図10のステップS13と同様の処理が行われる。また、このとき選択部73において、転送ビットレートも選択されるようにしてもよい。

【0201】

さらに、オーディオエレメントの組み合わせが選択されると、要求部142は、選択部73から供給された選択結果により示される組み合わせのオーディオエレメントからなるビットストリームの送信要求を通信部141に供給する。この送信要求は、例えば図2の矢印A11乃至矢印A16のうちの何れかにより示されるビットストリームの送信を要求するものである。

【0202】

ステップS74において、通信部141は、要求部142から供給された、ビットストリームの送信要求を、アクセス処理部22を介してサーバ11に送信する。

【0203】

すると、サーバ11からは、送信要求に応じて、要求された組み合わせのオーディオエレメントからなるビットストリームが送信されてくる。

【0204】

ステップS75において、通信部141は、アクセス処理部22を介して、サーバ11からビットストリームを受信してオーディオバッファ75に供給する。

【0205】

ビットストリームが受信されると、その後、ステップS76およびステップS77の処理が行われて復号処理は終了するが、これらの処理は図10のステップS15およびステップS16の処理と同様であるので、その説明は省略する。

【0206】

以上のようにしてデコーダ23は、オーディオエレメントの組み合わせを選択し、選択した組み合わせのビットストリームをサーバ11から受信してデコードを行う。これにより、ハード規模の異なる様々な機器で入力ビットストリームをデコードすることができるとともに、入力ビットストリームの転送ビットレートを削減することができる。

【0207】

第4の実施の形態

デコーダの構成例

さらに、選択した組み合わせのオーディオエレメントをサーバ11から取得する場合に、その組み合わせのなかの不要なオーディオエレメントを非転送とするようにしてもよい。

【0208】

そのような場合、デコーダ23は、例えば図15に示すように構成される。なお、図15において、図11または図13における場合と対応する部分には同一の符号を付してあり、その説明は適宜省略する。

【0209】

図 15 に示すデコーダ 23 は、通信部 141、バッファサイズ算出部 72、選択部 73、要求部 142、システムバッファ 111、オーディオバッファ 75、復号部 76、および出力部 77 を有している。図 15 に示すデコーダ 23 の構成は、図 13 に示したデコーダ 23 の構成に、さらにシステムバッファ 111 を設けた構成とされている。

【0210】

図 15 に示すデコーダ 23 では、選択部 73 はオーディオエレメントの組み合わせと、その組み合わせを構成するオーディオエレメントのなかの非転送とする不要なオーディオエレメントとを選択し、それらの選択結果を要求部 142 に供給する。

【0211】

ここで、不要なオーディオエレメントの選択は、例えばEXTエレメントに含まれる優先度情報に基づいて行われるが、EXTエレメントはどのようにして取得されるようにしてもよい。

【0212】

すなわち、例えば図 3 の矢印A21に示したようにEXTエレメントが単独でサーバ 11 に記録されている場合、通信部 141 は、デコード開始前の任意のタイミングでストリーミング制御部 21 を介して、サーバ 11 からEXTエレメントを取得する。そして、通信部 141 は、バッファサイズ算出部 72 を介して選択部 73 にEXTエレメントを供給する。

【0213】

また、例えば図 3 の矢印A22に示したようにEXTエレメントが入力ビットストリームのフレーム先頭に配置されている場合には、通信部 141 は、まずは入力ビットストリームの先頭部分にあるEXTエレメントをサーバ 11 から受信してバッファサイズ算出部 72 に供給する。そして、バッファサイズ算出部 72 は、通信部 141 からのEXTエレメントを選択部 73 に供給する。

【0214】

なお、以下では、図 3 の矢印A21に示したようにEXTエレメントが単独でサーバ 11 に記録されているものとして説明を続ける。

【0215】

要求部 142 は、選択部 73 から供給された選択結果に基づいて、選択された組み合わせを構成する、非転送としなかったオーディオエレメントからなるビットストリームの送信要求を通信部 141 に供給する。

【0216】

システムバッファ 111 は、通信部 141 からサイズ情報の供給を受ける。

【0217】

例えば図 7 の矢印A31に示したようにサイズ情報が単独でサーバ 11 に記録されている場合、通信部 141 は、デコード開始前の任意のタイミングでストリーミング制御部 21 を介して、サーバ 11 からサイズ情報を取得し、システムバッファ 111 に供給する。

【0218】

また、例えば図 7 の矢印A32や矢印A33に示したように、サイズ情報が入力ビットストリームの各フレームの先頭や、各オーディオエレメント内の先頭に配置されている場合、通信部 141 は、サーバ 11 から受信した入力ビットストリーム、より詳細には入力ビットストリームの分割された一部分をシステムバッファ 111 に供給する。

【0219】

なお、図 7 の矢印A33に示したようにサイズ情報が各オーディオエレメント内の先頭に配置されている場合には、選択部 73 により選択された組み合わせの非転送とされたオーディオエレメントについては、サイズ情報だけがビットストリームに含まれるようにされる。

【0220】

システムバッファ 111 は、サイズ情報に基づいて、上述した転送ビットレート調整処理RMT(1)または転送ビットレート調整処理RMT(2)によりバッファ制御を行って、通信部 141 から供給されたオーディオエレメントをオーディオバッファ 75 に供給する。なお、

10

20

30

40

50

以下では、転送ビットレート調整処理RMT(1)が行われるものとして説明を続ける。

【0221】

復号処理の説明

次に図16のフローチャートを参照して、図15に示したデコーダ23により行われる復号処理について説明する。

【0222】

ステップS101において、通信部141は、サーバ11から取得可能なオーディオエレメントの組み合わせを示す情報、およびEXTエレメントを受信してバッファサイズ算出部72に供給する。

【0223】

すなわち、通信部141は、取得可能なオーディオエレメントの組み合わせを示す情報、およびEXTエレメントの送信要求を、ストリーミング制御部21を介してサーバ11に送信する。また、通信部141は、その送信要求に応じてサーバ11から送信されてきたオーディオエレメントの組み合わせを示す情報、およびEXTエレメントを、ストリーミング制御部21を介して受信して、バッファサイズ算出部72に供給する。さらに、バッファサイズ算出部72は、通信部141からのEXTエレメントを選択部73に供給する。

【0224】

オーディオエレメントの組み合わせを示す情報が取得されると、ステップS102およびステップS103の処理が行われて転送を要求するオーディオエレメントが選択されるが、これらの処理は図12のステップS42およびステップS43の処理と同様であるので、その説明は省略する。

【0225】

但し、ステップS102では、オーディオエレメントの組み合わせを示す情報に基づいて必要バッファサイズが算出され、ステップS103では、選択部73による選択結果は要求部142に供給される。

【0226】

また、要求部142は、選択部73から供給された選択結果に基づいて、選択された組み合わせを構成する、非転送としなかったオーディオエレメントからなるビットストリームの送信要求を通信部141に供給する。換言すれば、選択された組み合わせのオーディオエレメントの送信が要求されるとともに、その組み合わせのなかのデコード対象とされないものとして選択された不要なオーディオエレメントの非転送が要求される。

【0227】

ステップS104において、通信部141は、要求部142から供給された、選択された組み合わせを構成する、非転送としなかったオーディオエレメントからなるビットストリームの送信要求を、アクセス処理部22を介してサーバ11に送信する。

【0228】

すると、サーバ11からは、ビットストリームの送信要求に応じて、要求された組み合わせを構成する、非転送とされなかったオーディオエレメントからなるビットストリームが送信されてくる。

【0229】

ステップS105において、通信部141は、アクセス処理部22を介して、サーバ11からビットストリームを受信してシステムバッファ111に供給する。

【0230】

ビットストリームが受信されると、その後、ステップS106乃至ステップS108の処理が行われて復号処理は終了するが、これらの処理は図12のステップS45乃至ステップS47の処理と同様であるので、その説明は省略する。

【0231】

以上のようにしてデコーダ23は、オーディオエレメントの組み合わせを選択するとともに、優先度情報に基づいてデコード対象としない不要なオーディオエレメントを選択する。これにより、ハード規模の異なる様々な機器で入力ビットストリームをデコードする

10

20

30

40

50

ことができるとともに入力ビットストリームの転送ビットレートを削減することができる。また、バッファ制御を行うことで、Minimum decoder input bufferサイズでのデコードが可能となる。

【0232】

ところで、上述した一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、コンピュータにインストールされる。ここで、コンピュータには、専用のハードウェアに組み込まれているコンピュータや、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどが含まれる。

10

【0233】

図17は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

【0234】

コンピュータにおいて、CPU (Central Processing Unit) 501, ROM (Read Only Memory) 502, RAM (Random Access Memory) 503は、バス504により相互に接続されている。

【0235】

バス504には、さらに、入出力インターフェース505が接続されている。入出力インターフェース505には、入力部506、出力部507、記録部508、通信部509、及びドライブ510が接続されている。

20

【0236】

入力部506は、キーボード、マウス、マイクロホン、撮像素子などよりなる。出力部507は、ディスプレイ、スピーカなどよりなる。記録部508は、ハードディスクや不揮発性のメモリなどよりなる。通信部509は、ネットワークインターフェースなどよりなる。ドライブ510は、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどのリムーバブルメディア511を駆動する。

【0237】

以上のように構成されるコンピュータでは、CPU501が、例えば、記録部508に記録されているプログラムを、入出力インターフェース505及びバス504を介して、RAM503にロードして実行することにより、上述した一連の処理が行われる。

30

【0238】

コンピュータ (CPU501) が実行するプログラムは、例えば、パッケージメディア等としてのリムーバブルメディア511に記録して提供することができる。また、プログラムは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提供することができる。

【0239】

コンピュータでは、プログラムは、リムーバブルメディア511をドライブ510に装着することにより、入出力インターフェース505を介して、記録部508にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部509で受信し、記録部508にインストールすることができる。その他、プログラムは、ROM502や記録部508に、あらかじめインストールしておくことができる。

40

【0240】

なお、コンピュータが実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであっても良いし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであっても良い。

【0241】

また、本技術の実施の形態は、上述した実施の形態に限定されるものではなく、本技術の要旨を逸脱しない範囲において種々の変更が可能である。

【0242】

50

例えば、本技術は、１つの機能をネットワークを介して複数の装置で分担、共同して処理するクラウドコンピューティングの構成をとることができる。

【０２４３】

また、上述のフローチャートで説明した各ステップは、１つの装置で実行する他、複数の装置で分担して実行することができる。

【０２４４】

さらに、１つのステップに複数の処理が含まれる場合には、その１つのステップに含まれる複数の処理は、１つの装置で実行する他、複数の装置で分担して実行することができる。

【０２４５】

さらに、本技術は、以下の構成とすることも可能である。

【０２４６】

[１]

オーディオエレメントの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、前記オーディオエレメントの１つの前記組み合わせを選択する選択部と、

選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成する生成部と

を備える復号装置。

[２]

前記選択部は、同じコンテンツについて予め用意された複数の前記組み合わせの中から１つの前記組み合わせを選択する

[１] に記載の復号装置。

[３]

前記複数の前記組み合わせごとに用意された、前記組み合わせの前記オーディオエレメントから構成されるビットストリームのうちの、前記選択部により選択された前記組み合わせのビットストリームを受信する通信部をさらに備える

[２] に記載の復号装置。

[４]

前記選択部は、ビットストリームを構成する複数の前記オーディオエレメントのうちのいくつかの前記オーディオエレメントを、１つの前記組み合わせとして選択する

[１] または [２] に記載の復号装置。

[５]

前記選択部は、前記ビットストリームのメタデータに基づいて１つの前記組み合わせを選択する

[４] に記載の復号装置。

[６]

前記選択部は、前記メタデータとして、予め定められた複数の前記組み合わせを示す情報、および前記オーディオエレメントの優先度情報のうちの少なくとも何れか一方に基づいて、１つの前記組み合わせを選択する

[５] に記載の復号装置。

[７]

前記ビットストリームから、前記選択部により選択された前記組み合わせの前記オーディオエレメントを抽出する抽出部をさらに備える

[４] 乃至 [６] の何れか一項に記載の復号装置。

[８]

前記選択部により選択された前記組み合わせの前記オーディオエレメントを受信する通信部をさらに備える

[４] 乃至 [６] の何れか一項に記載の復号装置。

[９]

10

20

30

40

50

デコード対象として選択されなかった前記オーディオエレメントのサイズに基づいて、前記生成部によりデコードされる前記オーディオエレメントのバッファへの格納を制御するバッファ制御部をさらに備える

[5] に記載の復号装置。

[1 0]

前記選択部は、選択した前記組み合わせを構成する前記オーディオエレメントのなかから、デコード対象としない前記オーディオエレメントをさらに選択し、

前記バッファ制御部は、前記選択部により選択された前記デコード対象としない前記オーディオエレメントのサイズに基づいて、前記選択部により選択された前記組み合わせを構成する、前記デコード対象としない前記オーディオエレメント以外の前記オーディオエレメントの前記バッファへの格納を制御する

10

[9] に記載の復号装置。

[1 1]

前記選択部は、前記オーディオエレメントの優先度情報に基づいて、前記デコード対象としない前記オーディオエレメントを選択する

[1 0] に記載の復号装置。

[1 2]

オーディオエレメントの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、前記オーディオエレメントの 1 つの前記組み合わせを選択し、

20

選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成する

ステップを含む復号方法。

[1 3]

オーディオエレメントの組み合わせごとに定まる、前記組み合わせの前記オーディオエレメントのデコードに必要なバッファサイズに基づいて、前記オーディオエレメントの 1 つの前記組み合わせを選択し、

選択された前記組み合わせの前記オーディオエレメントをデコードしてオーディオ信号を生成する

ステップを含む処理をコンピュータに実行させるプログラム。

30

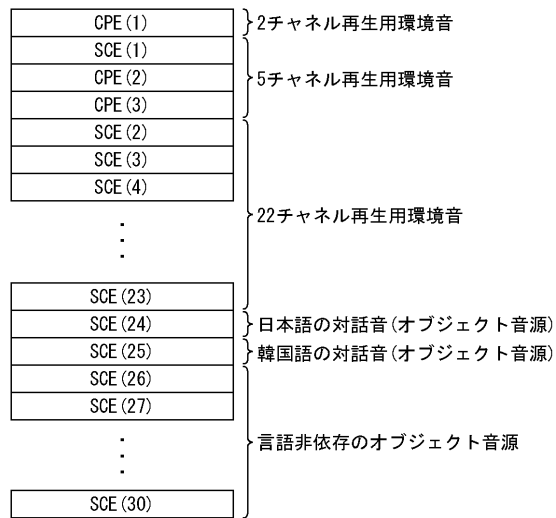
【符号の説明】

【 0 2 4 7 】

2 3 デコーダ, 7 1 取得部, 7 2 バッファサイズ算出部, 7 3 選択部,
7 4 抽出部, 7 5 オーディオバッファ, 7 6 復号部, 1 1 1 システムバッファ,
1 4 1 通信部, 1 4 2 要求部

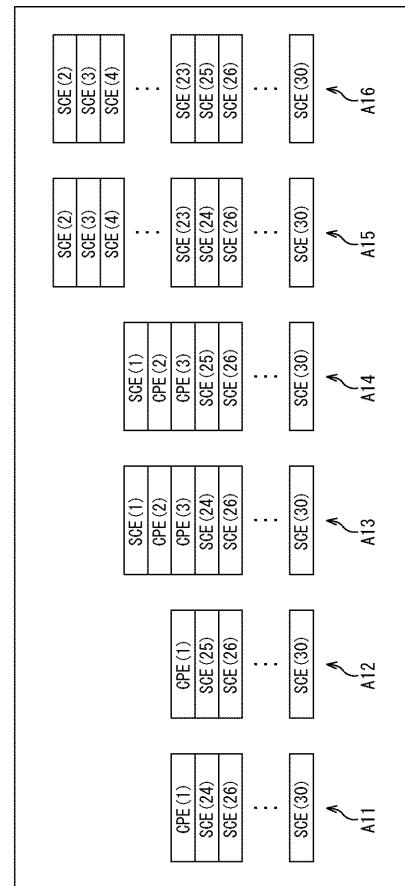
【図 1】

図1



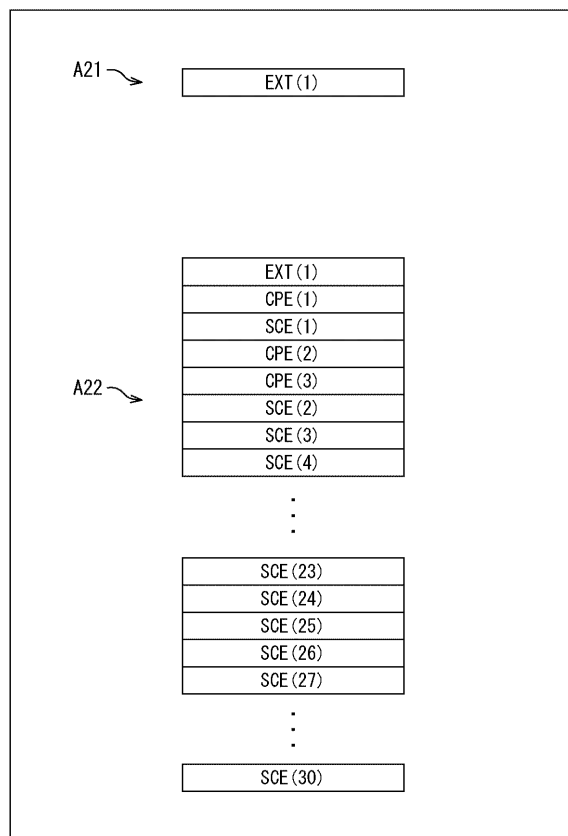
【図 2】

図2



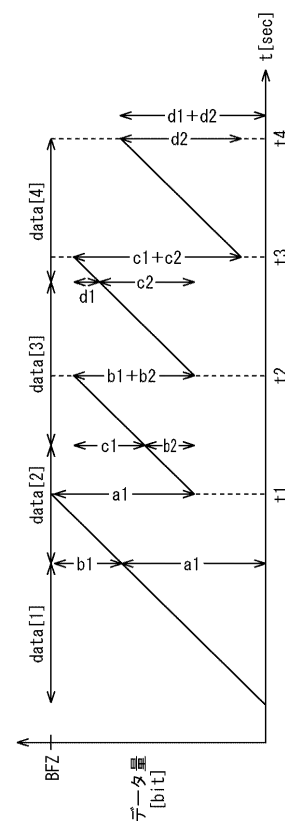
【図 3】

図3



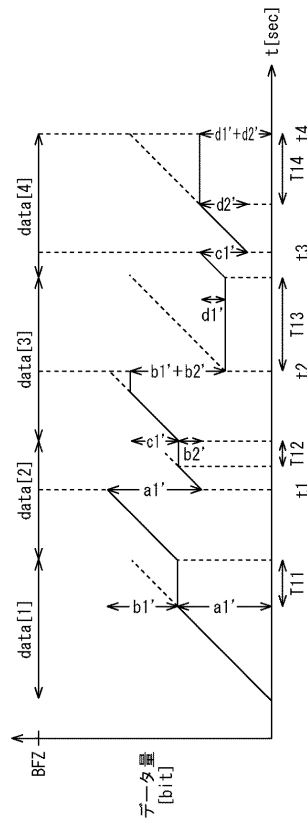
【図 4】

図4



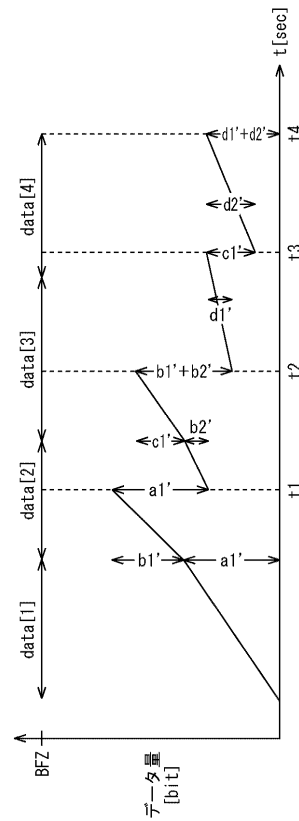
【図 5】

図5



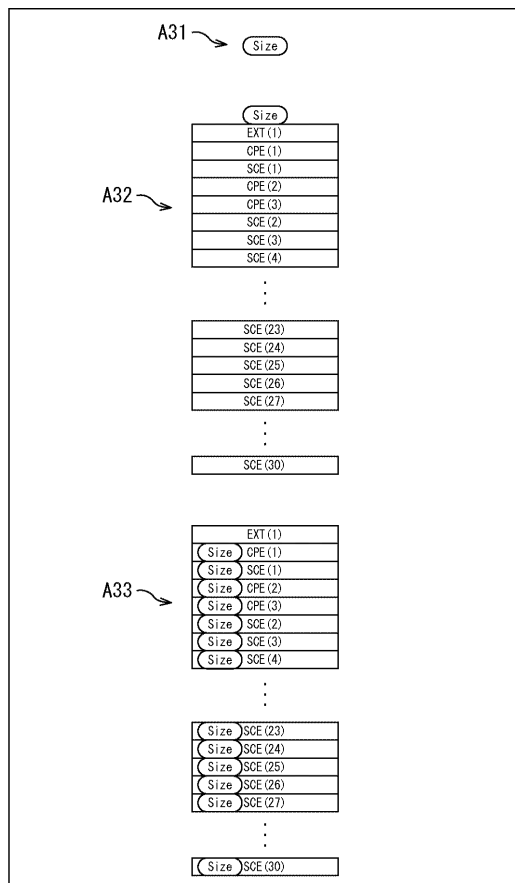
【図 6】

図6



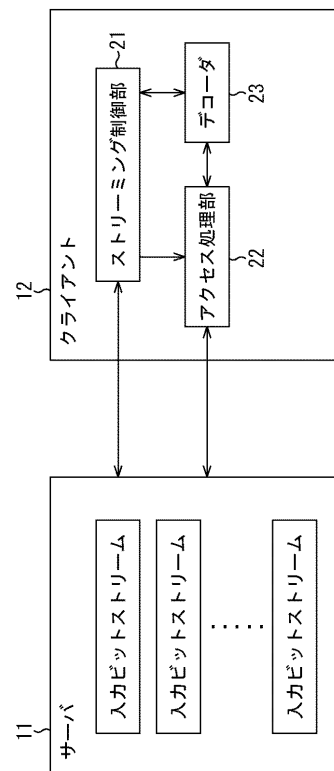
【図 7】

図7



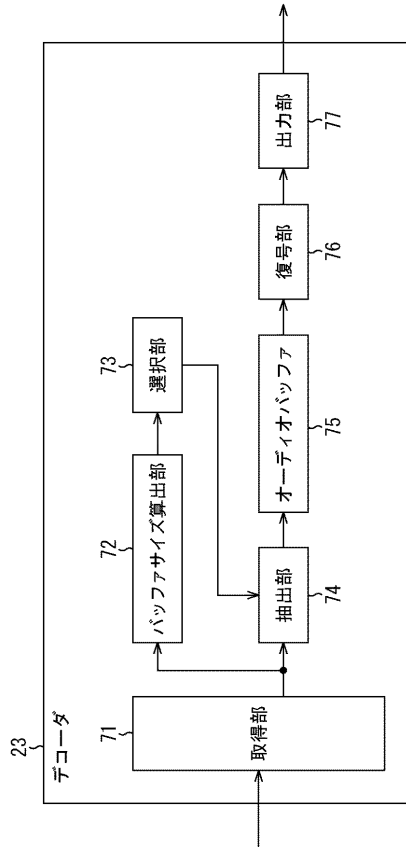
【図 8】

図8



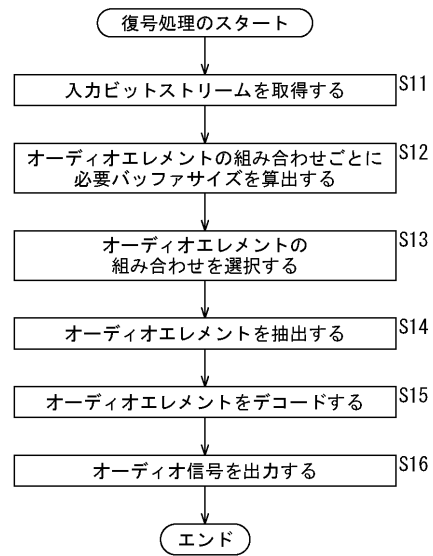
【図 9】

図9



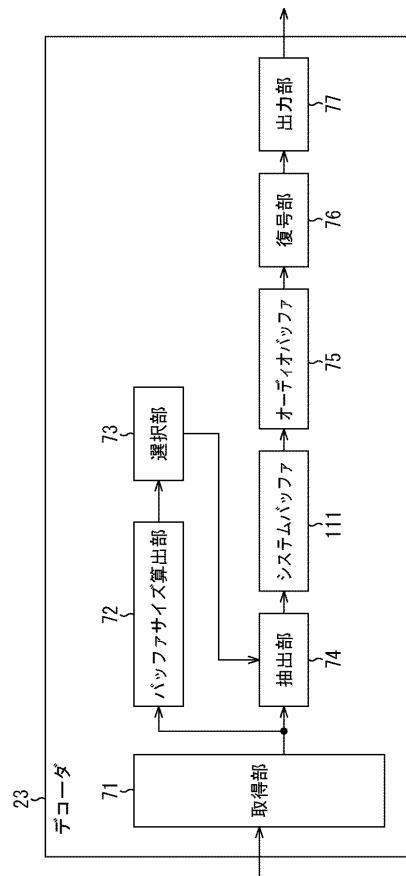
【図 10】

図10



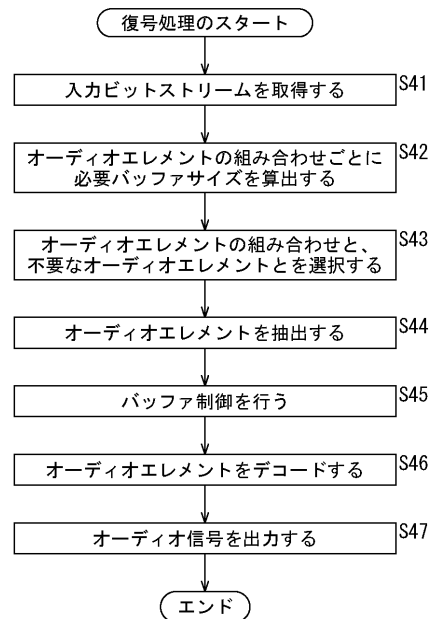
【図 11】

図11



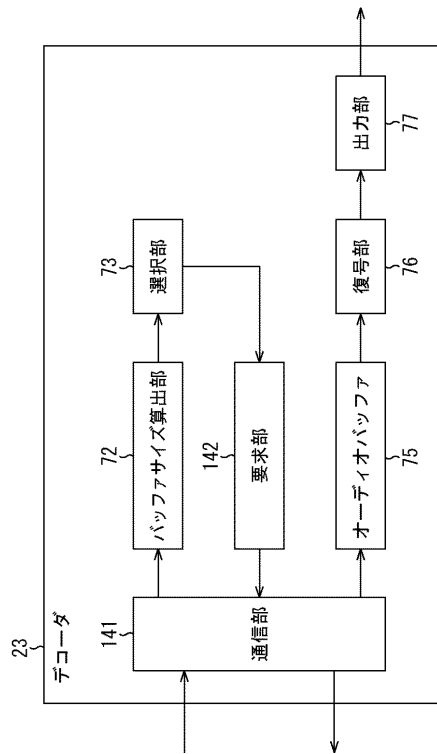
【図 12】

図12



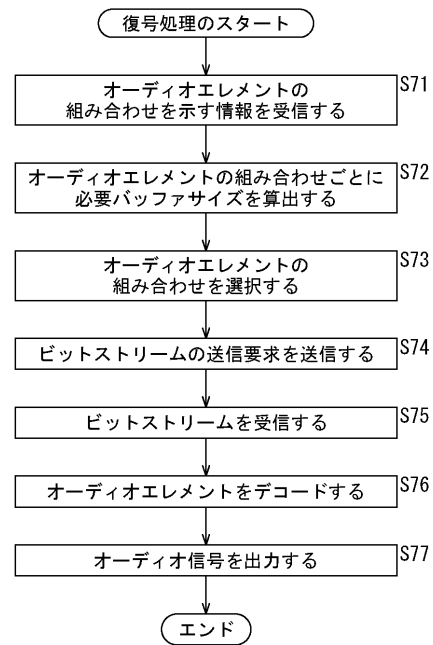
【図 13】

図13



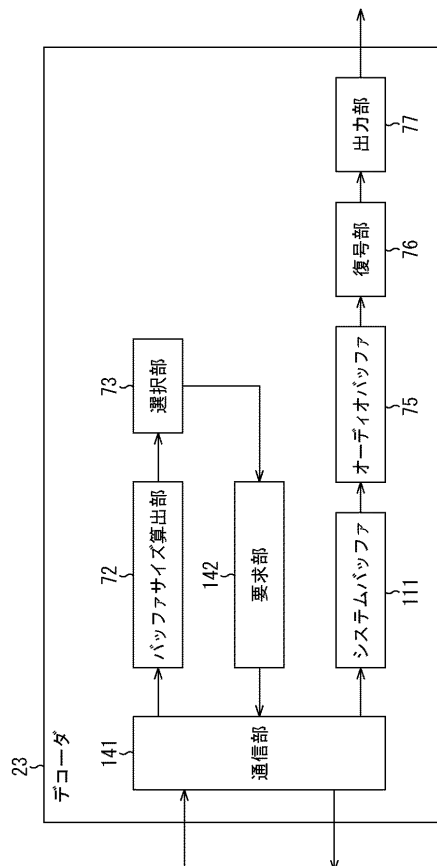
【図 14】

図14



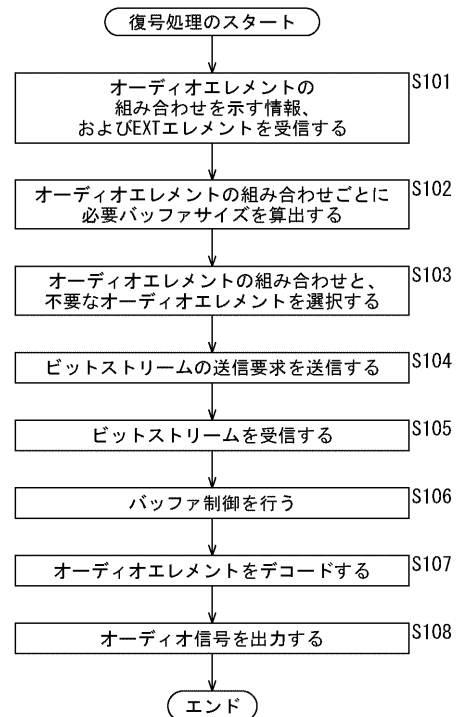
【図 15】

図15

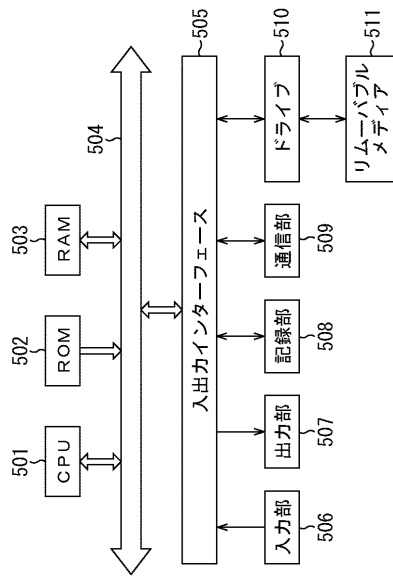


【図 16】

図16



【図17】
図17



フロントページの続き

- (72)発明者 史 潤宇
東京都港区港南1丁目7番1号 ソニー株式会社内
- (72)発明者 平林 光浩
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 堀 洋介

- (56)参考文献 特開2001-134294(JP, A)
米国特許出願公開第2012/0230497(US, A1)
特開2002-156998(JP, A)
特開2012-042972(JP, A)
特開2004-165776(JP, A)
特開2005-031289(JP, A)
国際公開第2014/007094(WO, A1)

- (58)調査した分野(Int.Cl., DB名)
- | | |
|------|--------|
| H04S | 7/00 |
| G10L | 19/008 |
| H04S | 3/00 |