

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2024-516133
(P2024-516133A)

(43)公表日 令和6年4月12日(2024.4.12)

(51)国際特許分類	F I	テーマコード (参考)
H 0 4 N 21/431 (2011.01)	H 0 4 N 21/431	5 B 0 5 0
H 0 4 N 21/84 (2011.01)	H 0 4 N 21/84	5 C 1 6 4
G 0 6 T 19/00 (2011.01)	G 0 6 T 19/00 6 0 0	

審査請求 未請求 予備審査請求 未請求 (全56頁)

(21)出願番号	特願2023-562735(P2023-562735)	(71)出願人	595020643 クゥアルコム・インコーポレイテッド QUALCOMM INCORPORATED
(86)(22)出願日	令和4年4月20日(2022.4.20)		
(85)翻訳文提出日	令和5年10月12日(2023.10.12)		
(86)国際出願番号	PCT/US2022/071809		
(87)国際公開番号	WO2022/226502		アメリカ合衆国、カリフォルニア州 9 2 1 2 1-1 7 1 4、サン・ディエゴ、 モアハウス・ドライブ 5 7 7 5
(87)国際公開日	令和4年10月27日(2022.10.27)		
(31)優先権主張番号	63/177,261	(74)代理人	110003708 弁理士法人鈴榮特許総合事務所
(32)優先日	令和3年4月20日(2021.4.20)		
(33)優先権主張国・地域又は機関	米国(US)	(72)発明者	ブアジジ、イメド アメリカ合衆国、カリフォルニア州 9 2 1 2 1-1 7 1 4、サン・ディエゴ、 モアハウス・ドライブ 5 7 7 5
(31)優先権主張番号	17/659,760		
(32)優先日	令和4年4月19日(2022.4.19)	(72)発明者	ストックハマー、トーマス アメリカ合衆国、カリフォルニア州 9
(33)優先権主張国・地域又は機関	米国(US)		最終頁に続く
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA 最終頁に続く		

(54)【発明の名称】 没入型メディアコンテンツをストリーミングするためのシーン記述のユーザ環境へのアンカリング

(57)【要約】

メディアデータを提示するための例示的なデバイスは、仮想シーン内の1つまたは複数の仮想オブジェクトを定義するメディアデータを記憶するように構成されたメモリと、回路内で実装され、仮想シーン内の1つまたは複数の仮想オブジェクトとシーンアンカーとを記述するデータを含むビットストリームのシーン記述を受信し、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表しており；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定し；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示するように構成された1つまたは複数のプロセッサと、を含む。

【選択図】 図 6

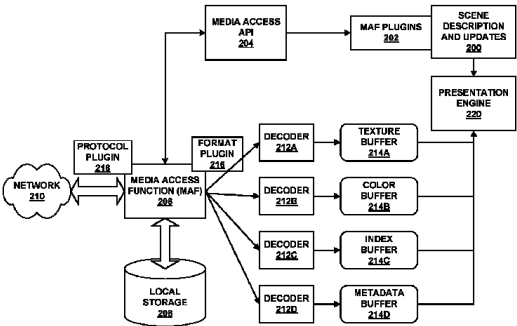


FIG. 6

【特許請求の範囲】

【請求項 1】

メディアデータを提示する方法であって、

ビットストリームのシーン記述を受信することと、前記シーン記述は、シーンアンカーと仮想シーン内の 1 つまたは複数の仮想オブジェクトとを記述するデータを含み、前記シーンアンカーは、前記仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、

前記シーンアンカーを使用して、前記仮想シーンと前記現実世界のプレゼンテーション環境との間の前記対応を決定することと、

前記決定された対応に従って、前記現実世界のプレゼンテーション環境内のロケーションにおいて前記 1 つまたは複数の仮想オブジェクトを提示することと、

を備える方法。

10

【請求項 2】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、ローカル、ビュー、またはステージのうちの 1 つまたは複数である、請求項 1 に記載の方法。

【請求項 3】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、アプリケーション定義空間であり、

前記方法は、前記アプリケーション定義空間を定義するデータを受信することをさらに備える、請求項 1 に記載の方法。

20

【請求項 4】

前記シーン記述は、シーン座標系から前記現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、請求項 1 に記載の方法。

【請求項 5】

前記変換は、回転、平行移動、または拡大縮小のうちの 1 つまたは複数を含む、請求項 4 に記載の方法。

【請求項 6】

前記 1 つまたは複数の仮想オブジェクトを提示することは、

前記仮想シーン内の前記 1 つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定することと、

前記現実世界のプレゼンテーション環境内の前記ロケーションを生成するために、前記変換に従って前記シグナリングされたロケーションを現実世界のロケーションに変換することと、

を備える、請求項 4 に記載の方法。

30

【請求項 7】

前記 1 つまたは複数の仮想オブジェクトを提示することは、前記シーンアンカーに従って、前記仮想シーンを前記現実世界のプレゼンテーション環境にアライメントさせることを備える、請求項 1 に記載の方法。

【請求項 8】

前記シーン記述は、ユーザから受信される入力を表すアクションを定義するデータをさらに含み、

前記方法は、前記アクションに応答して、前記 1 つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも 1 つの前記ロケーションに対する修正を決定することをさらに備え、

前記 1 つまたは複数の仮想オブジェクトを提示することは、前記 1 つまたは複数の仮想オブジェクトまたは前記仮想カメラのうちの前記少なくとも 1 つの前記ロケーションの前記修正に従って、前記 1 つまたは複数の仮想オブジェクトを提示することを備える、請求項 1 に記載の方法。

40

【請求項 9】

50

前記シーン記述は、前記１つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、請求項１に記載の方法。

【請求項１０】

メディアデータを提示するためのデバイスであって、
仮想シーン内の１つまたは複数の仮想オブジェクトを定義するメディアデータを記憶するように構成されたメモリと、

回路内で実装された１つまたは複数のプロセッサと、を備え、前記１つまたは複数のプロセッサは、

ビットストリームのシーン記述を受信することと、前記シーン記述は、シーンアンカーと仮想シーン内の１つまたは複数の仮想オブジェクトとを記述するデータを含み、前記シーンアンカーは、前記仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すものであり、

前記シーンアンカーを使用して、前記仮想シーンと前記現実世界のプレゼンテーション環境との間の前記対応を決定することと、

前記決定された対応に従って、前記現実世界のプレゼンテーション環境内のロケーションにおいて前記１つまたは複数の仮想オブジェクトを提示することと、

を行うように構成される、デバイス。

【請求項１１】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、ローカル、ビュー、またはステージのうちの１つまたは複数である、請求項１０に記載のデバイス。

【請求項１２】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、アプリケーション定義空間であり、

前記方法は、前記アプリケーション定義空間を定義するデータを受信することをさらに備える、請求項１０に記載のデバイス。

【請求項１３】

前記シーン記述は、シーン座標系から前記現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、請求項１０に記載のデバイス。

【請求項１４】

前記変換は、回転、平行移動、または拡大縮小のうちの１つまたは複数を含む、請求項１３に記載のデバイス。

【請求項１５】

前記１つまたは複数の仮想オブジェクトを提示するために、前記１つまたは複数のプロセッサは、

前記仮想シーン内の前記１つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定することと、

前記現実世界のプレゼンテーション環境内の前記ロケーションを生成するために、前記変換に従って前記シグナリングされたロケーションを現実世界のロケーションに変換することと、

を行うように構成される、請求項１３に記載のデバイス。

【請求項１６】

前記１つまたは複数の仮想オブジェクトを提示するために、前記１つまたは複数のプロセッサは、前記シーンアンカーに従って、前記仮想シーンを前記現実世界のプレゼンテーション環境にアライメントさせること、を行うように構成される、請求項１０に記載のデバイス。

【請求項１７】

前記シーン記述は、ユーザから受信される入力を表すアクションを定義するデータをさらに含み、

前記１つまたは複数のプロセッサは、前記アクションに応答して、前記１つまたは複数

10

20

30

40

50

の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つの前記ロケーションに対する修正を決定するようにさらに構成され、

前記1つまたは複数のプロセッサは、前記1つまたは複数の仮想オブジェクトまたは前記仮想カメラのうちの前記少なくとも1つの前記ロケーションの前記修正に従って、前記1つまたは複数の仮想オブジェクトを提示するように構成される、請求項10に記載のデバイス。

【請求項18】

前記シーン記述は、前記1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、請求項10に記載のデバイス。

【請求項19】

集積回路、
マイクロプロセッサ、および
ワイヤレス通信デバイス

のうちの少なくとも1つを備える、請求項10に記載のデバイス。

【請求項20】

実行されたとき、プロセッサに、

ビットストリームのシーン記述を受信することと、前記シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、前記シーンアンカーは、前記仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、

前記シーンアンカーを使用して、前記仮想シーンと前記現実世界のプレゼンテーション環境との間の前記対応を決定することと、

前記決定された対応に従って、前記現実世界のプレゼンテーション環境内のロケーションにおいて前記1つまたは複数の仮想オブジェクトを提示することと、

を行わせる命令を記憶したコンピュータ可読記憶媒体。

【請求項21】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、請求項20に記載のコンピュータ可読記憶媒体。

【請求項22】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、アプリケーション定義空間であり、

前記方法は、前記アプリケーション定義空間を定義するデータを受信することをさらに備える、請求項20に記載のコンピュータ可読記憶媒体。

【請求項23】

前記シーン記述は、シーン座標系から前記現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、請求項20に記載のコンピュータ可読記憶媒体。

【請求項24】

前記変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、請求項23に記載のコンピュータ可読記憶媒体。

【請求項25】

前記プロセッサに前記1つまたは複数の仮想オブジェクトを提示させる前記命令は、前記プロセッサに、

前記仮想シーン内の前記1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定することと、

前記現実世界のプレゼンテーション環境内の前記ロケーションを生成するために、前記変換に従って前記シグナリングされたロケーションを現実世界のロケーションに変換することと、

を行わせる命令を備える、請求項23に記載のコンピュータ可読記憶媒体。

【請求項26】

10

20

30

40

50

前記プロセッサに前記 1 つまたは複数の仮想オブジェクトを提示させる前記命令は、前記プロセッサに、前記シーンアンカーに従って、前記仮想シーンを前記現実世界のプレゼンテーション環境にアライメントさせることを行わせる命令を備える、請求項 20 に記載のコンピュータ可読記憶媒体。

【請求項 27】

前記シーン記述は、ユーザから受信される入力を表すアクションを定義するデータをさらに含み、

前記プロセッサに、前記アクションに応答して、前記 1 つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも 1 つの前記ロケーションに対する修正を決定させる命令をさらに備え、

前記プロセッサに前記 1 つまたは複数の仮想オブジェクトを提示させる前記命令は、前記 1 つまたは複数の仮想オブジェクトまたは前記仮想カメラのうちの前記少なくとも 1 つの前記ロケーションの前記修正に従って、前記プロセッサに、前記 1 つまたは複数の仮想オブジェクトを提示させる命令を備える、請求項 20 に記載のコンピュータ可読記憶媒体。

【請求項 28】

前記シーン記述は、前記 1 つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、請求項 20 に記載のコンピュータ可読記憶媒体。

【請求項 29】

メディアデータを提示するためのデバイスであって、

ビットストリームのシーン記述を受信するための手段と、前記シーン記述は、シーンアンカーと仮想シーン内の 1 つまたは複数の仮想オブジェクトとを記述するデータを含み、前記シーンアンカーは、前記仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、

前記シーンアンカーを使用して、前記仮想シーンと前記現実世界のプレゼンテーション環境との間の前記対応を決定するための手段と、

前記決定された対応に従って、前記現実世界のプレゼンテーション環境内のロケーションにおいて前記 1 つまたは複数の仮想オブジェクトを提示するための手段と、

を備えるデバイス。

【請求項 30】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、ローカル、ビュー、またはステージのうちの 1 つまたは複数である、請求項 29 に記載のデバイス。

【請求項 31】

前記シーンアンカーは、前記現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、前記タイプは、アプリケーション定義空間であり、

前記方法は、前記アプリケーション定義空間を定義するデータを受信することをさらに備える、請求項 29 に記載のデバイス。

【請求項 32】

前記シーン記述は、シーン座標系から前記現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、請求項 29 に記載のデバイス。

【請求項 33】

前記変換は、回転、平行移動、または拡大縮小のうちの 1 つまたは複数を含む、請求項 32 に記載のデバイス。

【請求項 34】

前記 1 つまたは複数の仮想オブジェクトを提示するための前記手段は、

前記仮想シーン内の前記 1 つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定するための手段と、

前記現実世界のプレゼンテーション環境内の前記ロケーションを生成するために、前記変換に従って前記シグナリングされたロケーションを現実世界のロケーションに変換する

10

20

30

40

50

ための手段と、

を備える、請求項 3 2 に記載のデバイス。

【請求項 3 5】

前記 1 つまたは複数の仮想オブジェクトを提示するための前記手段は、前記シーンアンカーに従って、前記仮想シーンを前記現実世界のプレゼンテーション環境にアライメントさせるための手段を備える、請求項 2 9 に記載のデバイス。

【請求項 3 6】

前記シーン記述は、ユーザから受信される入力を表すアクションを定義するデータをさらに含み、

前記アクションに応答して、前記 1 つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも 1 つの前記ロケーションに対する修正を決定するための手段をさらに備え、

前記 1 つまたは複数の仮想オブジェクトを提示するための前記手段は、前記 1 つまたは複数の仮想オブジェクトまたは前記仮想カメラのうちの前記少なくとも 1 つの前記ロケーションの前記修正に従って、前記 1 つまたは複数の仮想オブジェクトを提示するための手段を備える、請求項 2 9 に記載のデバイス。

【請求項 3 7】

前記シーン記述は、前記 1 つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、請求項 2 9 に記載のデバイス。

【発明の詳細な説明】

【優先権の主張】

【0 0 0 1】

【0001】本出願は、2022年4月19日に出願された米国特許出願第17／659，760号、および2021年4月20日に出願された米国仮出願第63／177，261号の優先権を主張し、その各々の内容全体が参照により本明細書に組み込まれる。2022年4月19日に出願された米国特許出願第17／659，760号は、2021年4月20日に出願された米国仮出願第63／177，261号の利益を主張する。

【技術分野】

【0 0 0 2】

【0002】本開示は、特に没入型（immersive）メディアコンテンツのための、メディアデータの記憶および輸送に関する。

【背景技術】

【0 0 0 3】

【0003】デジタルビデオ機能は、デジタルテレビジョン、デジタルダイレクトブロードキャストシステム、ワイヤレスブロードキャストシステム、携帯情報端末（PDA）、ラップトップまたはデスクトップコンピュータ、デジタルカメラ、デジタル記録デバイス、デジタルメディアプレーヤ、ビデオゲームデバイス、ビデオゲームコンソール、携帯電話または衛星無線電話、ビデオ遠隔会議デバイスなどを含む、広範囲にわたるデバイスに組み込まれ得る。デジタルビデオデバイスは、デジタルビデオ情報をより効率的に送信および受信するために、MPEG-2、MPEG-4、ITU-T H. 263またはITU-T H. 264／MPEG-4、Part 10、アドバンスドビデオコーディング（AVC：Advanced Video Coding）、ITU-T H. 265（高効率ビデオコーディング（HEVC：High Efficiency Video Coding）とも呼ばれる）によって定義された規格、およびそのような規格の拡張に記載されているビデオ圧縮技法などの、ビデオ圧縮技法を実装する。

【0 0 0 4】

【0004】ビデオ圧縮技法は、ビデオシーケンスに固有の冗長性を低減または除去するために、空間的予測および／または時間的予測を実行する。ブロックベースのビデオコーディングの場合、ビデオフレームまたはスライス、マクロブロックに区分され得る。各マクロブロックはさらに区分され得る。イントラコード化（I）フレームまたはスライス中

10

20

30

40

50

のマクロブロックは、隣接マクロブロックに関する空間的予測を使用して符号化される。インターコード化（PまたはB）フレームまたはスライス中のマクロブロックは、同じフレームもしくはスライス中の隣接マクロブロックに関する空間的予測、または他の参照フレームに関する時間的予測を使用し得る。

【0005】

【0005】ビデオデータが符号化された後、ビデオデータは送信または記憶のためにパケット化され得る。ビデオデータは、AVCなどの、国際標準化機構（ISO）ベースのメディアファイルフォーマットおよびその拡張などの、様々な規格のいずれかに準拠するビデオファイルへと組み立てられ得る。

【発明の概要】

【0006】

【0006】概して、本開示は、たとえば、拡張現実（AR）、仮想現実（VR）、または複合現実（MR：mixed reality）コンテンツのための、没入型メディアコンテンツをストリーミングおよび提示するための技法を説明する。そのようなコンテンツは、一般に、エクステンデッドリアリティ（XR）コンテンツと呼ばれ得る。概して、XRコンテンツは、1つまたは複数の仮想オブジェクトを含む仮想シーンを含み得る。仮想シーンは、3次元（3D）でナビゲート可能なメディアコンテンツとしてユーザに提示され得る。たとえば、ユーザは、コントローラおよび／または現実世界の動きを使用して、仮想シーンをナビゲートすることが可能であり得る。現実世界のプレゼンテーション環境内の適切な動きを可能とするために、および仮想オブジェクトが適切に提示されることを確実にするために、本開示は、仮想シーンを現実世界のプレゼンテーション環境にアンカーするための技法を説明する。たとえば、1つまたは複数の現実世界のアンカーポイントが、仮想シーンを現実世界のプレゼンテーション環境にアンカーするために識別および使用され得る。これらのアンカーポイントは、シーングラフを含み得るシーン記述において定義され得る。

【0007】

【0007】一例では、メディアデータを提示する方法は、ビットストリームのシーン記述を受信することと、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定することと、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示することと、を含む。

【0008】

【0008】別の例では、メディアデータを提示するためのデバイスは、仮想シーン内の1つまたは複数の仮想オブジェクトを定義するメディアデータを記憶するように構成されたメモリと、回路内で実装される1つまたは複数のプロセッサと、を含み、1つまたは複数のプロセッサは、ビットストリームのシーン記述を受信し、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すものであり；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定し；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示するように構成される。

【0009】

【0009】別の例では、コンピュータ可読記憶媒体は、実行されたとき、プロセッサに、ビットストリームのシーン記述を受信させ、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すものであり；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定させ；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにお

10

20

30

40

50

いて1つまたは複数の仮想オブジェクトを提示させる命令を記憶している。

【0010】

【0010】別の例では、メディアデータを提示するためのデバイスは、ビットストリームのシーン記述を受信するための手段と、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定するための手段と；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示するための手段と、を含む。

【0011】

【0011】1つまたは複数の例の詳細が添付の図面および以下の説明に記載される。他の特徴、目的、および利点は、説明および図面、ならびに特許請求の範囲から明らかになるう。

【図面の簡単な説明】

【0012】

【図1】【0012】ネットワークを介してメディアデータをストリーミングするための技法を実装する例示的なシステムを示すブロック図。

【図2】【0013】図1の取出し (retrieval) ユニットの構成要素の例示的なセットをより詳細に示すブロック図。

【図3】【0014】例示的なマルチメディアコンテンツの要素を示す概念図。

【図4】【0015】表現のセグメントに対応し得る、例示的なビデオファイルの要素を示すブロック図。

【図5】【0016】例示的なグラフィックスランゲージトランスミッションフォーマット (g l T F : G r a p h i c s L a n g u a g e T r a n s m i s s i o n F o r m a t) のシーングラフ構造を示すグラフ。

【図6】【0017】様々な構成要素間の重要なインターフェースを含む例示的なM P E G - Iアーキテクチャを示すブロック図。

【図7】【0018】例示的なサーキュラバッファを示す概念図。

【図8】【0019】例示的なM P E G - Iシーン記述アーキテクチャを示すブロック図。

【図9】【0020】シーン記述によって示される例示的なアンカーXR空間を示す概念図。

【図10】【0021】本開示の技法を実行することができる別の例示的なクライアントデバイス300を示すブロック図。

【図11】【0022】本開示の技法による、例示的な方法を示すフローチャート。

【発明を実施するための形態】

【0013】

【0023】概して、本開示は、たとえば拡張現実 (A R)、複合現実 (M R)、または仮想現実 (V R) コンテンツなどのエクステンデッドリアリティ (X R) コンテンツのための没入型メディアコンテンツをストリーミングするための技法を説明する。詳細には、本開示は、仮想環境と仮想環境が提示される現実世界のプレゼンテーション環境との間の対応をシグナリングするための技法を説明する。このようにして、クライアントデバイスは、現実世界のプレゼンテーション環境に適合するように、たとえば、仮想環境の境界が現実世界のプレゼンテーション環境の境界を超えないことを確実にするように、仮想環境を方向付けるおよび／またはスケーリングすることができる。したがって、クライアントデバイスは、仮想環境の境界に現実世界のプレゼンテーション環境の境界を超えさせ得るようなやり方で仮想環境を提示するのを回避することができ、これは、ユーザと壁、椅子、テーブルなどの現実世界のオブジェクトとの間の衝突を回避するのを助け得る。同様に、これらの技法は、より大きい現実世界の部屋が利用可能であるときに、より大きな現実世界のプレゼンテーション環境をユーザが利用できることを確実にし得る。

【0014】

【0024】これらの技法は、没入型メディアコンテンツを送るサーバデバイスが、没入型

10

20

30

40

50

メディアコンテンツの各受信者に複数の異なる現実世界のプレゼンテーション環境を表す詳細を備える必要があるという点で、ストリーミングされる没入型メディアコンテンツの文脈で有益であり得る。代わりに、これらの技法は、それぞれの各クライアントデバイスが、シーン記述において定義されるときにアンカーポイントに対応する適切な現実世界のアンカーポイントを決定することを可能にする。このようにして、没入型メディアコンテンツの個々の各提示が、没入型メディアコンテンツが提示される対応する現実世界のプレゼンテーション環境に適合することを依然として可能にしつつ、サーバによる過剰な処理が、回避され得る。

【0015】

[0025]本開示の技法は、ISOベースメディアファイルフォーマット、スケーラブルビデオコーディング（SVC）ファイルフォーマット、アドバンストビデオコーディング（AVC）ファイルフォーマット、第3世代パートナーシッププロジェクト（3GPP（登録商標））ファイルフォーマット、および／もしくはマルチビュービデオコーディング（MVC）ファイルフォーマット、または他の同様のビデオファイルフォーマットのうちのいずれかに従ってカプセル化されたビデオデータに準拠するビデオファイルに適用され得る。

【0016】

[0026]HTTPストリーミングでは、頻繁に使用される動作は、HEADと、GETと、部分GETとを含む。HEAD動作は、所与のユニフォームリソースロケータ（URL）またはユニフォームリソースネーム（URN）に関連付けられたペイロードを取り出すことなく、URLまたはURNに関連付けられたファイルのヘッダを取り出す。GET動作は、所与のURLまたはURNに関連付けられたファイル全体を取り出す。部分GET動作は、入力パラメータとしてバイト範囲を受信し、ファイルの連続するいくつかのバイトを取り出し、バイトの数は、受信されたバイト範囲に対応する。したがって、部分GET動作は1つまたは複数の個々のムービーフラグメントを得ることができるので、HTTPストリーミングのためのムービーフラグメントが与えられ得る。ムービーフラグメント中に、異なるトラックのいくつかのトラックフラグメントが存在し得る。HTTPストリーミングでは、メディアプレゼンテーションは、クライアントがアクセス可能であるデータの構造化された集合であり得る。クライアントは、ストリーミングサービスをユーザに提示するために、メディアデータ情報を要求し、ダウンロードし得る。

【0017】

[0027]HTTPストリーミングを使用して3GPPデータをストリーミングする例では、マルチメディアコンテンツのビデオおよび／またはオーディオデータに関する複数の表現が存在し得る。以下で説明されるように、異なる表現は、異なるコーディング特性（たとえば、ビデオコーディング規格の異なるプロファイルまたはレベル）、異なるコーディング規格もしくはコーディング規格の拡張（マルチビューおよび／またはスケーラブル拡張など）、または異なるビットレートに対応し得る。そのような表現のマニフェストは、メディアプレゼンテーション記述（MPD）データ構造において定義され得る。メディアプレゼンテーションは、HTTPストリーミングクライアントデバイスがアクセス可能であるデータの構造化された集合に対応し得る。HTTPストリーミングクライアントデバイスは、ストリーミングサービスをクライアントデバイスのユーザに提示するために、メディアデータ情報を要求し、ダウンロードし得る。メディアプレゼンテーションは、MPDの更新を含み得るMPDデータ構造中に記述され得る。

【0018】

[0028]メディアプレゼンテーションは、1つまたは複数の期間のシーケンスを含み得る。各期間は、次の期間の開始まで、または最後の期間の場合、メディアプレゼンテーションの終了まで継続し得る。各期間は、同じメディアコンテンツにおいて1つまたは複数の表現を含み得る。表現は、オーディオ、ビデオ、タイムドテキスト、または他のそのようなデータのいくつかの代替的な符号化バージョンのうちの1つであり得る。表現は、符号化タイプによって、たとえば、ビデオデータの場合、ビットレート、解像度、および／

またはコーデックによって、オーディオデータの場合、ビットレート、言語、および／またはコーデックによって異なり得る。表現という用語は、マルチメディアコンテンツの特定の期間に対応し、特定の方法で符号化された符号化オーディオデータまたはビデオデータのセクションを指すために使用され得る。

【0019】

【0029】特定の期間の表現は、表現が属する適応セットを示すMPD中の属性によって示されるグループに割り当てられ得る。同じ適応セット中の表現は、一般に、クライアントデバイスが、たとえば、帯域幅適応を実行するために、これらの表現の間で動的におよびシームレスに切り替えることができるという点で、互いに代替物と見なされる。たとえば、特定の期間のビデオデータの各表現は、対応する期間のマルチメディアコンテンツのビデオデータまたはオーディオデータなどのメディアデータを提示するために復号するのに表現のいずれかが選択され得るように、同じ適応セットに割り当てられ得る。1期間内のメディアコンテンツは、いくつかの例では、存在する場合、グループ0からの1つの表現、または各非ゼログループからの多くとも1つの表現の組合せのいずれかによって表され得る。期間の各表現についてのタイミングデータは、期間の開始時間に対して表現され得る。

10

【0020】

【0030】表現は、1つまたは複数のセグメントを含み得る。各表現は初期化セグメントを含むことがあり、または表現の各セグメントは自己初期化していることがある。存在するとき、初期化セグメントは、表現にアクセスするための初期化情報を含み得る。概して、初期化セグメントは、メディアデータを含まない。セグメントは、ユニフォームリソースロケータ（URL）、ユニフォームリソースネーム（URN）、またはユニフォームリソース識別子（URI）などの識別子によって一意に参照され得る。MPDは、各セグメントに識別子を与え得る。いくつかの例では、MPDはまた、URL、URN、またはURIによってアクセス可能なファイル内のセグメントのためのデータに対応し得るバイト範囲をrange属性の形態で与え得る。

20

【0021】

【0031】異なる表現は、異なるタイプのメディアデータの実質的に同時の取出しのために選択され得る。たとえば、クライアントデバイスは、セグメントを取り出す、オーディオ表現と、ビデオ表現と、タイムドテキスト表現とを選択し得る。いくつかの例では、クライアントデバイスは、帯域幅適応を実行するための特定の適応セットを選択し得る。すなわち、クライアントデバイスは、ビデオ表現を含む適応セット、オーディオ表現を含む適応セット、および／またはタイムドテキストを含む適応セットを選択し得る。代替的に、クライアントデバイスは、あるタイプのメディア（たとえば、ビデオ）の適応セットを選択し、他のタイプのメディア（たとえば、オーディオおよび／またはタイムドテキスト）の表現を直接選択し得る。

30

【0022】

【0032】図1は、ネットワークを介してメディアデータをストリーミングするための技法を実装する例示的なシステム10を示すブロック図である。この例では、システム10は、コンテンツ作成デバイス20と、サーバデバイス60と、クライアントデバイス40を含む。クライアントデバイス40およびサーバデバイス60は、インターネットを備え得るネットワーク74によって通信可能に結合される。いくつかの例では、コンテンツ作成デバイス20およびサーバデバイス60も、ネットワーク74または別のネットワークによって結合され得るか、あるいは直接通信可能に結合され得る。いくつかの例では、コンテンツ作成デバイス20およびサーバデバイス60は同じデバイスを備え得る。

40

【0023】

【0033】コンテンツ作成デバイス20は、図1の例では、オーディオソース22とビデオソース24とを備える。オーディオソース22は、たとえば、オーディオエンコーダ26によって符号化されるべき、キャプチャされたオーディオデータを表す電気信号を生成するマイクロフォンを備え得る。代替的に、オーディオソース22は、前に記録されたオ

50

オーディオデータを記憶する記憶媒体、コンピュータ化されたシンセサイザなどのオーディオデータ生成器、またはオーディオデータの任意の他のソースを備え得る。ビデオソース 24 は、ビデオエンコーダ 28 によって符号化されるべきビデオデータを生成するビデオカメラ、前に記録されたビデオデータで符号化された記憶媒体、コンピュータグラフィックスソースなどのビデオデータ生成ユニット、またはビデオデータの任意の他のソースを備え得る。コンテンツ作成デバイス 20 は、必ずしもすべての例においてサーバデバイス 60 に通信可能に結合されるとは限らないが、サーバデバイス 60 によって読み取られる別個のメディアにマルチメディアコンテンツを記憶し得る。

【0024】

[0034] 生のオーディオおよびビデオデータは、アナログまたはデジタルデータを備え得る。アナログデータは、オーディオエンコーダ 26 および／またはビデオエンコーダ 28 によって符号化される前にデジタル化され得る。オーディオソース 22 は、通話参加者が話している間、通話参加者からオーディオデータを取得し得、同時に、ビデオソース 24 は、通話参加者のビデオデータを取得し得る。他の例では、オーディオソース 22 は、記憶されたオーディオデータを備えるコンピュータ可読記憶媒体を備え得、ビデオソース 24 は、記憶されたビデオデータを備えるコンピュータ可読記憶媒体を備え得る。このようにして、本開示で説明される技法は、ライブ、ストリーミング、リアルタイムオーディオおよびビデオデータ、またはアーカイブされた、あらかじめ記録されたオーディオおよびビデオデータに適用され得る。

【0025】

[0035] ビデオフレームに対応するオーディオフレームは、概して、ビデオフレーム内に含まれるビデオソース 24 によってキャプチャされた（または生成された）ビデオデータと同時に、オーディオソース 22 によってキャプチャされた（または生成された）オーディオデータを含むオーディオフレームである。たとえば、通話参加者が一般に話すことによってオーディオデータを生成する間、オーディオソース 22 はオーディオデータをキャプチャし、同時に、すなわち、オーディオソース 22 がオーディオデータをキャプチャしている間、ビデオソース 24 は通話参加者のビデオデータをキャプチャする。したがって、オーディオフレームは、1 つまたは複数の特定のビデオフレームに時間的に対応し得る。したがって、ビデオフレームに対応するオーディオフレームは、一般に、オーディオデータとビデオデータとが同時にキャプチャされる状況、およびオーディオフレームとビデオフレームとが、それぞれ、同時にキャプチャされたオーディオデータとビデオデータとを備える状況に対応する。

【0026】

[0036] いくつかの例では、オーディオエンコーダ 26 は、符号化オーディオフレームのオーディオデータが記録された時間を表す、各符号化オーディオフレームにおけるタイムスタンプを符号化することができ、同様に、ビデオエンコーダ 28 は、符号化ビデオフレームのビデオデータが記録された時間を表す、各符号化ビデオフレームにおけるタイムスタンプを符号化することができる。そのような例では、ビデオフレームに対応するオーディオフレームは、あるタイムスタンプを備えるオーディオフレームと、同じタイムスタンプを備えるビデオフレームとを備え得る。コンテンツ作成デバイス 20 は、オーディオエンコーダ 26 および／またはビデオエンコーダ 28 がそこからタイムスタンプを生成し得るか、あるいはオーディオソース 22 およびビデオソース 24 がオーディオデータとビデオデータとをそれぞれタイムスタンプに関連付けるために使用し得る、内部クロックを含み得る。

【0027】

[0037] いくつかの例では、オーディオソース 22 は、オーディオデータが記録された時間に対応するデータをオーディオエンコーダ 26 に送ることができ、ビデオソース 24 は、ビデオデータが記録された時間に対応するデータをビデオエンコーダ 28 に送ることができる。いくつかの例では、オーディオエンコーダ 26 は、必ずしもオーディオデータが記録された絶対時間を示すことなしに、符号化オーディオデータの相対的時間順序を示

すために、符号化オーディオデータ中のシーケンス識別子を符号化し得、同様に、ビデオエンコーダ 28 も、符号化ビデオデータの相対的時間順序を示すためにシーケンス識別子を使用し得る。同様に、いくつかの例では、シーケンス識別子は、タイムスタンプにマッピングされるか、または場合によってはタイムスタンプと関連し得る。

【0028】

【0038】オーディオエンコーダ 26 は、一般に符号化オーディオデータのストリームを生成するが、ビデオエンコーダ 28 は、符号化ビデオデータのストリームを生成する。データの各個々のストリームは（オーディオかビデオかにかかわらず）エレメンタリストリームと呼ばれることがある。エレメンタリストリームは、表現のデジタル的にコード化された（場合によっては圧縮された）単一の構成要素である。たとえば、表現のコード化ビデオまたはオーディオ部分は、エレメンタリストリームであり得る。エレメンタリストリームは、ビデオファイル内にカプセル化される前に、パケット化エレメンタリストリーム（PES）に変換され得る。同じ表現内では、1つのエレメンタリストリームに属する PES パケットを他のものから区別するためにストリーム ID が使用され得る。エレメンタリストリームの基本データ単位は、パケット化エレメンタリストリーム（PES）パケットである。したがって、コーディングされたビデオデータは、一般にエレメンタリビデオストリームに対応する。同様に、オーディオデータは、1つまたは複数のそれぞれのエレメンタリストリームに対応する。

【0029】

【0039】ITU-T H. 264 / AVC および来るべき高効率ビデオコーディング（HEVC）規格などの、多くのビデオコーディング規格は、シンタックスと、セマンティクスと、エラーのないビットストリームのための復号プロセスとを定義し、これらのいずれも特定のプロファイルまたはレベルに準拠する。ビデオコーディング規格は、通常、エンコーダを指定しないが、エンコーダは、生成されたビットストリームがデコーダの規格に準拠することを保証することを課される。ビデオコーディング規格のコンテキストでは、「プロファイル」は、アルゴリズム、機能、またはツール、およびそれらに適用される制約のサブセットに対応する。たとえば、H. 264 規格によって定義される「プロファイル」は、H. 264 規格によって指定されたビットストリームシンタックス全体のサブセットである。「レベル」は、たとえば、ピクチャの解像度と、ビットレートと、ブロック処理レートとに関連するデコーダメモリおよび計算などの、デコーダリソース消費の制限に対応する。プロファイルは `profile_idc`（プロファイルインジケータ）値でシグナリングされ得るが、レベルは `level_idc`（レベルインジケータ）値でシグナリングされ得る。

【0030】

【0040】H. 264 規格は、たとえば、所与のプロファイルのシンタックスによって課される境界内で、復号されたピクチャの指定されたサイズなどの、ビットストリーム中のシンタックス要素によってとられる値に応じて、エンコーダおよびデコーダのパフォーマンスの大きい変動を必要とする可能性が依然としてあることを認める。H. 264 規格は、多くの適用例において、特定のプロファイル内でシンタックスのすべての仮定的使用を処理することが可能なデコーダを実装することが現実的でもなく、経済的でもないことをさらに認める。したがって、H. 264 規格は、ビットストリーム中のシンタックス要素の値に課された制約の規定されたセットとして「レベル」を定義する。これらの制約は、値に関する単純な制限であり得る。代替的に、これらの制約は、値の演算の組合せ（たとえば、ピクチャの幅×ピクチャの高さ×毎秒復号されるピクチャの数）に関する制約の形態をとり得る。H. 264 規格は、個々の実装形態が、サポートされるプロファイルごとに異なるレベルをサポートし得ることをさらに規定している。

【0031】

【0041】プロファイルに準拠するデコーダは、通常、プロファイル中で定義されたすべての機能をサポートする。たとえば、コーディング機能として、B ピクチャコーディングは、H. 264 / AVC のベースラインプロファイルではサポートされないが、H. 26

4 / A V C の他のプロファイルではサポートされる。レベルに準拠するデコーダは、レベルにおいて定義された制限を超えてリソースを必要としない任意のビットストリームを復号することが可能であるべきである。プロファイルおよびレベルの定義は、説明可能性のために役立ち得る。たとえば、ビデオ送信中に、プロファイル定義とレベル定義のペアが全送信セッションについてネゴシエートされ、同意され得る。より具体的には、H. 264 / A V C では、レベルは、処理される必要があるマクロブロックの数に関する制限と、復号ピクチャバッファ (D P B) サイズと、コード化ピクチャバッファ (C P B) サイズと、垂直動きベクトル範囲と、マクロブロック (M B) などの2つの連続するブロックごとの動きベクトルの最大数と、B ブロックが 8×8 ピクセル未満のサブブロック区分を有することができるかどうかとを定義し得る。このようにして、デコーダは、デコーダがビットストリームを適切に復号することが可能であるかどうかを決定し得る。

10

【0032】

【0042】図1の例では、コンテンツ作成デバイス20のカプセル化ユニット30は、ビデオエンコーダ28からのコード化ビデオデータを備えるエレメンタリストリームと、オーディオエンコーダ26からのコード化オーディオデータを備えるエレメンタリストリームとを受信する。いくつかの例では、ビデオエンコーダ28およびオーディオエンコーダ26は各々、符号化データからPESパケットを形成するためのパケットサイズを含み得る。他の例では、ビデオエンコーダ28およびオーディオエンコーダ26は各々、符号化データからPESパケットを形成するためのそれぞれのパケットサイズとインターフェースし得る。さらに他の例では、カプセル化ユニット30は、符号化オーディオデータと符号化ビデオデータとからPESパケットを形成するためのパケットサイズを含み得る。

20

【0033】

【0043】ビデオエンコーダ28は、様々なビットレートで、ピクセル解像度、フレームレート、様々なコーディング規格への準拠、様々なコーディング規格のための様々なプロファイルおよび／またはプロファイルのレベルへの準拠、(たとえば、2次元または3次元再生用の) 1つまたは複数のビューを有する表現、あるいは他のそのような特性などの様々な特性を用いてマルチメディアコンテンツの異なる表現を生成するために、様々な方法でマルチメディアコンテンツのビデオデータを符号化し得る。本開示で使用される表現は、オーディオデータ、ビデオデータ、テキストデータ(たとえば、クローズドキャプション用)、または他のそのようなデータのうちの1つを備え得る。表現は、オーディオエレメンタリストリームまたはビデオエレメンタリストリームなどのエレメンタリストリームを含み得る。各PESパケットは、PESパケットが属するエレメンタリストリームを識別するstream_idを含み得る。カプセル化ユニット30は、エレメンタリストリームを様々な表現のビデオファイル(たとえば、セグメント)にアSEMBLすることを担う。

30

【0034】

【0044】カプセル化ユニット30は、オーディオエンコーダ26およびビデオエンコーダ28から表現のエレメンタリストリームのPESパケットを受信し、PESパケットから対応するネットワークアブストラクションレイヤ(NAL)ユニットを形成する。コード化ビデオセグメントは、ビデオ電話、記憶、ブロードキャスト、またはストリーミングなどのアプリケーションに対処する「ネットワークフレンドリー」なビデオ表現を提供する、NALユニットに編成され得る。NALユニットは、ビデオコーディングレイヤ(VCL)NALユニットと非VCL NALユニットとに分類され得る。VCLユニットは、コア圧縮エンジンを含み得、ブロック、マクロブロック、および／またはスライスレベルデータを含み得る。他のNALユニットは、非VCL NALユニットであり得る。いくつかの例では、通常は1次コード化ピクチャとして提示される、1つの時間インスタンス中のコード化ピクチャは、1つまたは複数のNALユニットを含み得るアクセスユニット中に含まれ得る。

40

【0035】

【0045】非VCL NALユニットは、特に、パラメータセットNALユニットとSE

50

I N A Lユニットとを含み得る。パラメータセットは、（シーケンスパラメータセット（S P S）中の）シーケンスレベルヘッダ情報と、（ピクチャパラメータセット（P P S）中の）まれに変化するピクチャレベルヘッダ情報とを含み得る。パラメータセット（たとえば、P P SおよびS P S）がある場合、まれに変化する情報がシーケンスごとまたはピクチャごとに繰り返される必要はなく、したがって、コーディング効率が改善され得る。さらに、パラメータセットの使用は、重要なヘッダ情報の帯域外送信を可能にし、誤り耐性のための冗長送信の必要性を回避することができる。帯域外送信の例では、S E I N A Lユニットなどの、他のN A Lユニットとは異なるチャンネル上でパラメータセットN A Lユニットが送信され得る。

【0036】

10

【0046】補足エンハンスメント情報（S E I）は、V C L N A Lユニットからのコード化ピクチャサンプルを復号するためには必要でないが、復号、表示、誤り耐性、および他の目的に関するプロセスを支援し得る情報を含み得る。S E Iメッセージは、非V C L N A Lユニット中に含まれることがある。S E Iメッセージは、一部の規格仕様の規範的部分であり、したがって、規格に準拠するデコーダ実装のために常に必須であるとは限らない。S E Iメッセージは、シーケンスレベルS E IメッセージまたはピクチャレベルS E Iメッセージであり得る。S V Cの例におけるスケーラビリティ情報S E Iメッセージ、M V Cにおけるビュースケーラビリティ情報S E IメッセージなどのS E Iメッセージ内に、何らかのシーケンスレベル情報が含まれる場合がある。これらの例示的なS E Iメッセージは、たとえば、動作点の抽出およびそれらの動作点の特性に関する情報を搬送し得る。加えて、カプセル化ユニット30は、表現の特性を記述するメディアプレゼンテーション記述子（M P D）などの、マニフェストファイルを形成し得る。カプセル化ユニット30は、拡張可能マークアップ言語（X M L）に従ってM P Dをフォーマットし得る。

20

【0037】

【0047】カプセル化ユニット30は、マニフェストファイル（たとえば、M P D）とともに、マルチメディアコンテンツの1つまたは複数の表現についてのデータを出力インターフェース32に与え得る。出力インターフェース32は、ユニバーサルシリアルバス（U S B）インターフェース、C DまたはD V Dライターまたはバーナーなど、記憶媒体に書き込むためのネットワークインターフェースまたはインターフェース、磁気またはフラッシュ記憶媒体へのインターフェース、あるいはメディアデータを記憶または送信するための他のインターフェースを備え得る。カプセル化ユニット30は、マルチメディアコンテンツの表現の各々のデータを出力インターフェース32に与えることができ、出力インターフェース32は、ネットワーク送信または記憶媒体を介してそのデータをサーバデバイス60に送ることができる。図1の例では、サーバデバイス60は、様々なマルチメディアコンテンツ64を記憶する記憶媒体62を含み、各マルチメディアコンテンツ64は、それぞれのマニフェストファイル66と1つまたは複数の表現68A～68N（表現68）とを含む。いくつかの例では、出力インターフェース32はデータを直接ネットワーク74に送ることもできる。

30

【0038】

【0048】いくつかの例では、表現68は、適応セットに分離され得る。すなわち、表現68の様々なサブセットは、コーデック、プロファイルおよびレベル、解像度、ビューの数、セグメントのファイルフォーマット、表現を用いて表示されるべきテキストおよび／または復号され、たとえばスピーカーによって提示されるべきオーディオデータの言語または他の特性を識別し得るテキストタイプ情報、適応セット中の表現のためのシーンのカメラアングルまたは現実世界のカメラパースペクティブを記述し得るカメラアングル情報、特定の視聴者のコンテンツ適合性を記述するレーティング情報など、特性のそれぞれの共通のセットを含み得る。

40

【0039】

【0049】マニフェストファイル66は、特定の適応セットに対応する表現68のサブセットを示すデータと、適応セットについての共通の特性とを含み得る。マニフェストファ

50

イル 6 6 はまた、ビットレートなどの、適応セットの個々の表現についての個々の特性を表すデータを含み得る。このようにして、適応セットは、簡略化されたネットワーク帯域幅適応を可能にし得る。適応セット中の表現は、マニフェストファイル 6 6 の適応セット要素の子要素を使用して示され得る。

【0040】

[0050]サーバデバイス 6 0 は、要求処理ユニット 7 0 とネットワークインターフェース 7 2 とを含む。いくつかの例では、サーバデバイス 6 0 は、複数のネットワークインターフェースを含み得る。さらに、サーバデバイス 6 0 の機能のうちのいずれかまたはすべてが、ルータ、ブリッジ、プロキシデバイス、スイッチ、または他のデバイスなどの、コンテンツ配信ネットワークの他のデバイス上で実装され得る。いくつかの例では、コンテンツ配信ネットワークの中間デバイスは、マルチメディアコンテンツ 6 4 のデータをキャッシュし、サーバデバイス 6 0 の構成要素に実質的に準拠する構成要素を含み得る。概して、ネットワークインターフェース 7 2 は、ネットワーク 7 4 を介してデータを送信および受信するように構成される。

【0041】

[0051]要求処理ユニット 7 0 は、クライアントデバイス 4 0 などのクライアントデバイスから、記憶媒体 6 2 のデータについてのネットワーク要求を受信するように構成される。たとえば、要求処理ユニット 7 0 は、R. Fielding らによる、RFC 2616、「Hypertext Transfer Protocol-HTTP/1.1」、Network Working Group、IETF、1999 年 6 月に記載されている、ハイパーテキスト転送プロトコル (HTTP) バージョン 1.1 を実装し得る。すなわち、要求処理ユニット 7 0 は、HTTP GET 要求または部分 GET 要求を受信し、これらの要求に応答してマルチメディアコンテンツ 6 4 のデータを与えるように構成され得る。要求は、たとえば、セグメントの URL を使用して、表現 6 8 のうちの 1 つのセグメントを指定し得る。いくつかの例では、要求はまた、セグメントの 1 つまたは複数のバイト範囲を指定し、したがって部分 GET 要求を備え得る。要求処理ユニット 7 0 は、表現 6 8 のうちの 1 つのセグメントのヘッダデータを与えるための HTTP HEAD 要求をサービスするようにさらに構成され得る。いずれの場合も、要求処理ユニット 7 0 は、要求されたデータをクライアントデバイス 4 0 などの要求元デバイスに与えるための要求を処理するように構成され得る。

【0042】

[0052]追加または代替として、要求処理ユニット 7 0 は、eMBMS などのブロードキャストまたはマルチキャストプロトコルを介してメディアデータを配信するように構成され得る。コンテンツ作成デバイス 2 0 は、説明した方法と実質的に同じ方法で DASH セグメントおよび／またはサブセグメントを作成し得るが、サーバデバイス 6 0 は、eMBMS または別のブロードキャストもしくはマルチキャストネットワークトランスポートプロトコルを使用して、これらのセグメントまたはサブセグメントを配信し得る。たとえば、要求処理ユニット 7 0 は、クライアントデバイス 4 0 からマルチキャストグループ参加要求を受信するように構成され得る。すなわち、サーバデバイス 6 0 は、マルチキャストグループに関連付けられたインターネットプロトコル (IP) アドレスを、クライアントデバイス 4 0 を含む、特定のメディアコンテンツ (たとえば、ライブイベントのブロードキャスト) に関連付けられたクライアントデバイスに広告することができる。今度は、クライアントデバイス 4 0 が、マルチキャストグループに参加するための要求をサブミットすることができる。この要求は、ネットワーク 7 4 全体にわたって、たとえば、ネットワーク 7 4 を構成するルータに伝播され得、その結果、ルータは、マルチキャストグループに関連付けられた IP アドレスを宛先とするトラフィックを、クライアントデバイス 4 0 などの加入クライアントデバイスに向けさせられる。

【0043】

[0053]図 1 の例に示されるように、マルチメディアコンテンツ 6 4 は、メディアプレゼンテーション記述 (MPD) に対応し得るマニフェストファイル 6 6 を含む。マニフェ

10

20

30

40

50

ストファイル 66 は、異なる代替表現 68（たとえば、異なる品質を有するビデオサービス）の記述を含み得、その記述は、たとえば、表現 68 のコーデック情報と、プロファイル値と、レベル値と、ビットレートと、他の記述特性とを含み得る。クライアントデバイス 40 は、表現 68 のセグメントにアクセスする方法を決定するためにメディアプレゼンテーションの MPD を取り出し得る。

【0044】

【0054】詳細には、取出しユニット 52 は、ビデオデコーダ 48 の復号能力とビデオ出力 44 のレンダリング能力とを決定するために、クライアントデバイス 40 の構成データ（図示せず）を取り出し得る。構成データはまた、クライアントデバイス 40 のユーザによって選択された言語の選好、クライアントデバイス 40 のユーザによって設定された深度の選好に対応する 1 つまたは複数のカメラパースペクティブ、および／またはクライアントデバイス 40 のユーザによって選択されたレーティングの選好のうちのいずれかまたはすべてを含み得る。取出しユニット 52 は、たとえば、HTTP GET 要求と部分 GET 要求とをサブミットするように構成されたウェブブラウザまたはメディアクライアントを備え得る。取出しユニット 52 は、クライアントデバイス 40 の 1 つまたは複数のプロセッサまたは処理ユニット（図示せず）によって実行されるソフトウェア命令に対応し得る。いくつかの例では、取出しユニット 52 に関して説明される機能のすべてまたは部分が、ハードウェア、あるいはハードウェア、ソフトウェア、および／またはファームウェアの組合せで実装され得、ソフトウェアまたはファームウェアの命令を実行するために必須のハードウェアが設けられ得る。

【0045】

【0055】取出しユニット 52 は、クライアントデバイス 40 の復号能力およびレンダリング能力を、マニフェストファイル 66 の情報によって示される表現 68 の特性と比較し得る。取出しユニット 52 は、最初に、表現 68 の特性を決定するためにマニフェストファイル 66 の少なくとも一部分を取り出し得る。たとえば、取出しユニット 52 は、1 つまたは複数の適応セットの特性を記述するマニフェストファイル 66 の一部分を要求し得る。取出しユニット 52 は、クライアントデバイス 40 のコーディング能力およびレンダリング能力によって満たされ得る特性を有する表現 68 のサブセット（たとえば、適応セット）を選択し得る。次いで、取出しユニット 52 は、適応セット中の表現のビットレートを決定し、ネットワーク帯域幅の現在利用可能な量を決定し、ネットワーク帯域幅によって満たされ得るビットレートを有する表現のうちの 1 つからセグメントを取り出し得る。

【0046】

【0056】概して、より高いビットレート表現はより高品質のビデオ再生を生じ得るが、利用可能なネットワーク帯域幅が減少したとき、より低いビットレート表現は十分な品質のビデオ再生を与え得る。したがって、利用可能なネットワーク帯域幅が比較的高いとき、取出しユニット 52 は比較的高いビットレート表現からデータを取り出し得るが、利用可能なネットワーク帯域幅が低いとき、取出しユニット 52 は比較的低いビットレート表現からデータを取り出し得る。このようにして、クライアントデバイス 40 は、ネットワーク 74 上でマルチメディアデータをストリーミングしながら、ネットワーク 74 の変化するネットワーク帯域幅利用可能性にも適応し得る。

【0047】

【0057】追加または代替として、取出しユニット 52 は、eMBMS または IP マルチキャストなどのブロードキャストまたはマルチキャストネットワークプロトコルに従ってデータを受信するように構成され得る。そのような例では、取出しユニット 52 は、特定のメディアコンテンツに関連付けられたマルチキャストネットワークグループに参加するための要求をサブミットし得る。マルチキャストグループに参加した後、取出しユニット 52 は、サーバデバイス 60 またはコンテンツ作成デバイス 20 に発行されるさらなる要求なしに、マルチキャストグループのデータを受信し得る。取出しユニット 52 は、マルチキャストグループのデータがもはや必要とされないときに、たとえば、再生を停止する

ために、またはチャンネルを異なるマルチキャストグループに変更するために、マルチキャストグループを離れるための要求をサブミットし得る。

【0048】

【0058】ネットワークインターフェース54は、選択された表現のセグメントのデータを受信し、取出しユニット52に与えることができ、取出しユニット52は、今度は、そのセグメントをカプセル化解除ユニット50に与えることができる。カプセル化解除ユニット50は、ビデオファイルの要素を構成PESストリームにカプセル化解除し、符号化データを取り出すためにPESストリームをパケット化解除し、たとえば、ストリームのPESパケットヘッダによって示されるように、符号化データがオーディオストリームの一部であるのかビデオストリームの一部であるのかに応じて、符号化データをオーディオデコーダ46またはビデオデコーダ48のいずれかに送ることができる。オーディオデコーダ46は、符号化オーディオデータを復号し、復号オーディオデータをオーディオ出力42に送り、ビデオデコーダ48は、符号化ビデオデータを復号し、ストリームの複数のビューを含み得る復号ビデオデータをビデオ出力44に送る。

10

【0049】

【0059】ビデオエンコーダ28、ビデオデコーダ48、オーディオエンコーダ26、オーディオデコーダ46、カプセル化ユニット30、取出しユニット52、およびカプセル化解除ユニット50は各々、適用可能な場合、1つまたは複数のマイクロプロセッサ、デジタル信号プロセッサ(DSP)、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)、ディスクリート論理回路、ソフトウェア、ハードウェア、ファームウェア、またはそれらの任意の組合せなどの、様々な好適な処理回路のいずれかとして実装され得る。ビデオエンコーダ28およびビデオデコーダ48の各々は、1つまたは複数のエンコーダまたはデコーダ中に含まれ得、そのいずれも複合ビデオエンコーダ／デコーダ(コーデック)の一部として統合され得る。同様に、オーディオエンコーダ26およびオーディオデコーダ46の各々は、1つまたは複数のエンコーダまたはデコーダ中に含まれ得、そのいずれも複合コーデックの一部として統合され得る。ビデオエンコーダ28、ビデオデコーダ48、オーディオエンコーダ26、オーディオデコーダ46、カプセル化ユニット30、取出しユニット52、および／またはカプセル化解除ユニット50を含む装置は、集積回路、マイクロプロセッサ、および／またはセルラー電話などのワイヤレス通信デバイスを備え得る。

20

30

【0050】

【0060】クライアントデバイス40、サーバデバイス60、および／またはコンテンツ作成デバイス20は、本開示の技法に従って動作するように構成され得る。例として、本開示は、クライアントデバイス40およびサーバデバイス60に関して、これらの技法について説明する。しかしながら、コンテンツ作成デバイス20は、サーバデバイス60の代わりに(またはそれに加えて)これらの技法を実行するように構成され得ることを理解されたい。

【0051】

【0061】カプセル化ユニット30は、NALユニットが属するプログラムを識別するヘッダ、ならびにペイロード、たとえば、オーディオデータ、ビデオデータ、またはNALユニットが対応するトランスポートまたはプログラムストリームを記述するデータを備える、NALユニットを形成し得る。たとえば、H.264/AVCでは、NALユニットは1バイトのヘッダと変動するサイズのペイロードとを含む。そのペイロード中にビデオデータを含むNALユニットは、様々なグラニュラリティレベルのビデオデータを備え得る。たとえば、NALユニットは、ビデオデータのブロック、複数のブロック、ビデオデータのスライス、またはビデオデータのピクチャ全体を備え得る。カプセル化ユニット30は、エレメンタリストリームのPESパケットの形態でビデオエンコーダ28から符号化ビデオデータを受信し得る。カプセル化ユニット30は、各エレメンタリストリームを対応するプログラムに関連付け得る。

40

【0052】

50

[0062]カプセル化ユニット30はまた、複数のNALユニットからアクセスユニットをアセンブルし得る。概して、アクセスユニットは、ビデオデータのフレームを表すための1つまたは複数のNALユニットと、そのフレームに対応するオーディオデータが利用可能なとき、そのようなオーディオデータとを備え得る。アクセスユニットは、概して、1つの出力時間インスタンスにわたるすべてのNALユニット、たとえば1つの時間インスタンスにわたるすべてのオーディオデータおよびビデオデータを含む。たとえば、各ビューが20フレーム毎秒(fps)のフレームレートを有する場合、各時間インスタンスは0.05秒の時間間隔に対応し得る。この時間間隔中に、同じアクセスユニット(同じ時間インスタンス)のすべてのビューの固有のフレームは同時にレンダリングされ得る。一例では、アクセスユニットは、1次コード化ピクチャとして提示され得る、1つの時間インスタンス中のコード化ピクチャを備え得る。

10

【0053】

[0063]したがって、アクセスユニットは、共通の時間インスタンスのすべてのオーディオフレームおよびビデオフレーム、たとえば、時間Xに対応するすべてのビューを備え得る。本開示はまた、特定のビューの符号化ピクチャを「ビュー構成要素」と呼ぶ。すなわち、ビュー構成要素は、特定の時間における特定のビューの符号化ピクチャ(またはフレーム)を備え得る。したがって、アクセスユニットは、共通の時間インスタンスのすべてのビュー構成要素を備えるものと定義され得る。アクセスユニットの復号順序は、必ずしも出力順序または表示順序と同じである必要はない。

【0054】

20

[0064]メディアプレゼンテーションは、異なる代替表現(たとえば、異なる品質を有するビデオサービス)の記述を含み得るメディアプレゼンテーション記述(MPD)を含み得、その記述は、たとえば、コーデック情報と、プロファイル値と、レベル値とを含み得る。MPDは、マニフェストファイル66などのマニフェストファイルの一例である。様々なプレゼンテーションのムービーフラグメントにアクセスする方法を決定するために、クライアントデバイス40は、メディアプレゼンテーションのMPDを取り出し得る。ムービーフラグメントは、ビデオファイルのムービーフラグメントボックス(moofボックス)中に配置され得る。

【0055】

[0065]マニフェストファイル66(たとえば、MPDを備え得る)は、表現68のセグメントの利用可能性を広告し得る。すなわち、MPDは、表現68のうちの1つの第1のセグメントが利用可能になる壁時計時間を示す情報と、表現68内のセグメントの持続時間を示す情報とを含み得る。このようにして、クライアントデバイス40の取出しユニット52は、開始時間と、特定のセグメントに先行するセグメントの持続時間とに基づいて、各セグメントがいつ利用可能であるかを決定し得る。

30

【0056】

[0066]カプセル化ユニット30が、受信されたデータに基づいてNALユニットおよび/またはアクセスユニットをビデオファイルにアセンブルした後、カプセル化ユニット30はビデオファイルを出力のために出力インターフェース32に渡す。いくつかの例では、カプセル化ユニット30は、ビデオファイルをローカルに記憶するか、またはビデオファイルを直接クライアントデバイス40に送るのではなく、出力インターフェース32を介してビデオファイルをリモートサーバに送り得る。出力インターフェース32は、たとえば、送信機、トランシーバ、たとえば、オプティカルドライブ、磁気メディアドライブ(たとえば、フロッピー(登録商標)ドライブ)などの、コンピュータ可読媒体にデータを書き込むためのデバイス、ユニバーサルシリアルバス(USB)ポート、ネットワークインターフェース、または他の出力インターフェースを備え得る。出力インターフェース32は、ビデオファイルを、たとえば、送信信号、磁気メディア、光メディア、メモリ、フラッシュドライブ、または他のコンピュータ可読媒体などの、コンピュータ可読媒体に出力する。

40

【0057】

50

[0067]ネットワークインターフェース54は、ネットワーク74を介してNALユニットまたはアクセスユニットを受信し、取出しユニット52を介してNALユニットまたはアクセスユニットをカプセル化解除ユニット50に与え得る。カプセル化解除ユニット50は、ビデオファイルの要素を構成PESストリームにカプセル化解除し、符号化データを取り出すためにPESストリームをパケット化解除し、たとえば、ストリームのPESパケットヘッダによって示されるように、符号化データがオーディオストリームの一部であるのかビデオストリームの一部であるのかに応じて、符号化データをオーディオデコーダ46またはビデオデコーダ48のいずれかに送ることができる。オーディオデコーダ46は、符号化オーディオデータを復号し、復号オーディオデータをオーディオ出力42に送り、ビデオデコーダ48は、符号化ビデオデータを復号し、ストリームの複数のビューを含み得る復号ビデオデータをビデオ出力44に送る。

10

[0058]

[0068]以下により詳細に説明されるように、クライアントデバイス40は、単独または任意の組合せで本開示の様々な技法を実行するように構成され得る。概して、取出し(retrieval)ユニット52は、上記で説明したようなメディアデータ(たとえば、シーンデータ)と、シーン記述を含むビットストリームとを取り出すように構成され得る。シーン記述は、メディアデータによって表される仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すアンカーポイントデータを含み得る。クライアントデバイス40は、アンカーポイントデータを使用して仮想シーンを現実世界のプレゼンテーション環境にアンカーするとともに、必要に応じて、たとえば、回転、平行移動(translation)、および/または拡大縮小(scaling)によって仮想シーンを変換するようにも構成され得る。

20

[0059]

[0069]図2は、図1の取出しユニット52の構成要素の例示的なセットをより詳細に示すブロック図である。この例では、取出しユニット52は、eMBMSミドルウェアユニット100と、DASHクライアント110と、メディアアプリケーション112とを含む。

[0060]

[0070]この例では、eMBMSミドルウェアユニット100は、eMBMS受信ユニット106と、キャッシュ104と、プロキシサーバユニット102とをさらに含む。この例では、eMBMS受信ユニット106は、たとえば、tools.ietf.org/html/rfc6726において利用可能なT. Pairaら、「FLUTE-File Delivery over Unidirectional Transport」、ネットワークワーキンググループ、RFC6726、2012年11月に記載されているFile Delivery over Unidirectional Transport (FLUTE)に従って、eMBMSを介してデータを受信するように構成される。すなわち、eMBMS受信ユニット106は、たとえば、ブロードキャスト/マルチキャストサービスセンター(BMSC)として働き得るサーバデバイス60から、ブロードキャストを介してファイルを受信し得る。

30

[0061]

[0071]eMBMSミドルウェアユニット100がファイルに関するデータを受信すると、eMBMSミドルウェアユニットは、受信されたデータをキャッシュ104中に記憶し得る。キャッシュ104は、フラッシュメモリ、ハードディスク、RAM、または任意の他の適切な記憶媒体などのコンピュータ可読記憶媒体を備え得る。

40

[0062]

[0072]プロキシサーバユニット102は、DASHクライアント110のためのサーバとして働き得る。たとえば、プロキシサーバユニット102は、DASHクライアント110にMPDファイルまたは他のマニフェストファイルを提供し得る。プロキシサーバユニット102は、MPDファイル中のセグメントに関する利用可能時間と、セグメントが取り出され得るハイパーリンクとを広告し得る。これらのハイパーリンクは、クライア

50

ントデバイス 40 に対応するローカルホストアドレスプレフィックス（たとえば、IPv4 の場合は 127.0.0.1）を含み得る。このようにして、DASH クライアント 110 は、HTTP GET または部分 GET 要求を使用して、プロキシサーバユニット 102 にセグメントを要求し得る。たとえば、リンク `http://127.0.0.1/rep1/seg3` から利用可能なセグメントの場合、DASH クライアント 110 は、`http://127.0.0.1/rep1/seg3` についての要求を含む HTTP GET 要求を構築し、その要求をプロキシサーバユニット 102 にサブミットし得る。プロキシサーバユニット 102 は、要求されたデータをキャッシュ 104 から取り出し、そのような要求に応答して、そのデータを DASH クライアント 110 に提供し得る。

【0063】

10

【0073】図 3 は、例示的なマルチメディアコンテンツ 120 の要素を示す概念図である。マルチメディアコンテンツ 120 は、マルチメディアコンテンツ 64（図 1）、または記憶媒体 62 に記憶された別のマルチメディアコンテンツに対応し得る。図 3 の例では、マルチメディアコンテンツ 120 は、メディアプレゼンテーション記述（MPD）122 と、複数の表現 124A～124N（表現 124）とを含む。表現 124A は、任意のヘッダデータ 126 と、セグメント 128A～128N（セグメント 128）とを含み、表現 124N は、任意のヘッダデータ 130 と、セグメント 132A～132N（セグメント 132）とを含む。文字 N は、便宜上、表現 124 の各々中の最後のムービーフラグメントを指定するために使用される。いくつかの例では、表現 124 間で異なる数のムービーフラグメントが存在し得る。

20

【0064】

【0074】MPD 122 は、表現 124 とは別のデータ構造を備え得る。MPD 122 は、図 1 のマニフェストファイル 66 に対応し得る。同様に、表現 124 は、図 1 の表現 68 に対応し得る。概して、MPD 122 は、コーディング特性およびレンダリング特性、適応セット、MPD 122 が対応するプロファイル、テキストタイプ情報、カメラアングル情報、レーティング情報、トリックモード情報（たとえば、時間サブシーケンスを含む表現を示す情報）、および／または（たとえば、再生中のメディアコンテンツ中へのターゲット広告挿入のための）リモート期間を取り出すための情報などの、表現 124 の特性を全体的に記述するデータを含み得る。

【0065】

30

【0075】ヘッダデータ 126 は、存在するとき、セグメント 128 の特性、たとえば、ランダムアクセスポイント（RAP、ストリームアクセスポイント（SAP）とも呼ばれる）の時間ロケーションを記述することができ、セグメント 128 のランダムアクセスポイントは、ランダムアクセスポイント、セグメント 128 内のランダムアクセスポイントへのバイトオフセット、セグメント 128 のユニフォームリソースロケータ（URL）、またはセグメント 128 の他の態様を含む。ヘッダデータ 130 は、存在するとき、セグメント 132 に関する同様の特性を記述し得る。追加または代替として、そのような特性は、MPD 122 内に完全に含まれ得る。

【0066】

40

【0076】セグメント 128、132 は、1 つまたは複数のコード化ビデオサンプルを含み、コード化ビデオサンプルの各々は、ビデオデータのフレームまたはスライスを含み得る。セグメント 128 のコード化ビデオサンプルの各々は、同様の特性、たとえば、高さ、幅、および帯域幅の要件を有し得る。そのような特性は MPD 122 のデータによって記述され得るが、そのようなデータは図 3 の例に示されていない。MPD 122 は、本開示で説明されるシグナリングされた情報のいずれかまたはすべてに加えて、3GPP 仕様によって説明される特性を含み得る。

【0067】

【0077】セグメント 128、132 の各々は、一意のユニフォームリソースロケータ（URL）に関連付けられ得る。したがって、セグメント 128、132 の各々は、DASH などのストリーミングネットワークプロトコルを使用して独立して取出し可能であり得

50

る。このようにして、クライアントデバイス40などの宛先デバイスは、セグメント128または132を取り出すためにHTTP GET要求を使用し得る。いくつかの例では、クライアントデバイス40は、セグメント128または132の特定のバイト範囲を取り出すためにHTTP部分GET要求を使用し得る。

【0068】

【0078】図4は、図3のセグメント128、132のうちの1つなどの表現のセグメントに対応し得る例示的なビデオファイル150の要素を示すブロック図である。セグメント128、132の各々は、図4の例に示されるデータの配列に実質的に一致するデータを含み得る。ビデオファイル150は、セグメントをカプセル化されると言われることがある。上記で説明したように、ISOベースメディアファイルフォーマットおよびその拡張によるビデオファイルは、データを「ボックス」と呼ばれる一連のオブジェクトに記憶する。図4の例では、ビデオファイル150は、ファイルタイプ（FTYP）ボックス152と、ムービー（MOOV）ボックス154と、セグメントインデックス（sidx）ボックス162と、ムービーフラグメント（MOOF）ボックス164と、ムービーフラグメントランダムアクセス（MFRA）ボックス166とを含む。図4はビデオファイルの一例を表しているが、他のメディアファイルは、ISOベースメディアファイルフォーマットおよびその拡張に従って、ビデオファイル150のデータと同様に構造化された他のタイプのメディアデータ（たとえば、オーディオデータ、タイムドテキストデータなど）を含み得ることを理解されたい。

【0069】

【0079】ファイルタイプ（FTYP）ボックス152は概して、ビデオファイル150のファイルタイプを記述する。ファイルタイプボックス152は、ビデオファイル150の最良の使用を記述する仕様を識別するデータを含み得る。ファイルタイプボックス152は、代替的に、MOOVボックス154、ムービーフラグメントボックス164、および／またはMFRAボックス166の前に配置され得る。

【0070】

【0080】いくつかの例では、ビデオファイル150などのセグメントは、FTYPボックス152の前にMPD更新ボックス（図示せず）を含み得る。MPD更新ボックスは、MPDを更新するための情報とともに、ビデオファイル150を含む表現に対応するMPDが更新されるべきであることを示す情報を含み得る。たとえば、MPD更新ボックスは、MPDを更新するために使用されるリソースのためのURIまたはURLを提供し得る。別の例として、MPD更新ボックスは、MPDを更新するためのデータを含み得る。いくつかの例では、MPD更新ボックスは、ビデオファイル150のセグメントタイプ（STYP）ボックス（図示せず）の直後にくることができ、ここで、STYPボックスは、ビデオファイル150のセグメントタイプを定義し得る。

【0071】

【0081】MOOVボックス154は、図4の例では、ムービーヘッダ（MVHD）ボックス156と、トラック（TRAK）ボックス158と、1つまたは複数のムービー拡張（MVEX）ボックス160とを含む。概して、MVHDボックス156は、ビデオファイル150の一般的特性を記述し得る。たとえば、MVHDボックス156は、ビデオファイル150が最初に生成されたとき、ビデオファイル150が最後に変更されたときを記述するデータ、ビデオファイル150の時間軸、ビデオファイル150の再生の持続時間、またはビデオファイル150を一般に記述する他のデータを含み得る。

【0072】

【0082】TRAKボックス158は、ビデオファイル150のトラックについてのデータを含み得る。TRAKボックス158は、TRAKボックス158に対応するトラックの特性を記述するトラックヘッダ（TKHD）ボックスを含み得る。いくつかの例では、TRAKボックス158はコード化ビデオピクチャを含み得るが、他の例では、トラックのコード化ビデオピクチャは、TRAKボックス158および／またはsidxボックス162のデータによって参照され得るムービーフラグメント164中に含まれ得る。

【0073】

【0083】いくつかの例では、ビデオファイル150は、2つ以上のトラックを含み得る。したがって、MOOVボックス154は、ビデオファイル150中のトラックの数に等しいいくつかのTRAKボックスを含み得る。TRAKボックス158は、ビデオファイル150の対応するトラックの特性を記述し得る。たとえば、TRAKボックス158は、対応するトラックについての時間および／または空間情報を記述し得る。カプセル化ユニット30（図1）が、ビデオファイル150などのビデオファイル中にパラメータセットトラックを含むとき、MOOVボックス154のTRAKボックス158と同様のTRAKボックスは、パラメータセットトラックの特性を記述し得る。カプセル化ユニット30は、パラメータセットトラックを記述するTRAKボックス内のパラメータセットトラック中にシーケンスレベルSEIメッセージが存在することをシグナリングし得る。

10

【0074】

【0084】MVEXボックス160は、たとえば、もしあれば、MOOVボックス154内に含まれるビデオデータに加えて、ビデオファイル150がムービーフラグメント164を含むことをシグナリングするように、対応するムービーフラグメント164の特性を記述し得る。ビデオデータをストリーミングするコンテキストでは、コード化ビデオピクチャは、MOOVボックス154ではなくムービーフラグメント164中に含まれ得る。したがって、すべてのコード化ビデオサンプルは、MOOVボックス154ではなくムービーフラグメント164中に含まれ得る。

【0075】

20

【0085】MOOVボックス154は、ビデオファイル150中のムービーフラグメント164の数に等しいいくつかのMVEXボックス160を含み得る。MVEXボックス160の各々は、ムービーフラグメント164のうちの対応する1つの特性を記述し得る。たとえば、各MVEXボックスは、ムービーフラグメント164のうちの対応する1つの持続時間を記述するムービー拡張ヘッダボックス（MEHD）ボックスを含み得る。

【0076】

【0086】上述のように、カプセル化ユニット30は、シーケンスデータセットを、実際のコード化ビデオデータを含まないビデオサンプル中に記憶し得る。ビデオサンプルは、概して、特定の時間インスタンスにおけるコード化ピクチャの表現である、アクセスユニットに対応し得る。AVCのコンテキストでは、コード化ピクチャは、アクセスユニットのすべてのピクセルを構成するための情報を含んでいる1つまたは複数のVCLNALユニットと、SEIメッセージなどの他の関連する非VCLNALユニットとを含む。したがって、カプセル化ユニット30は、ムービーフラグメント164のうちの1つの中に、シーケンスレベルSEIメッセージを含み得る、シーケンスデータセットを含み得る。カプセル化ユニット30はさらに、シーケンスデータセットおよび／またはシーケンスレベルSEIメッセージの存在を、ムービーフラグメント164のうちの1つに対応するMVEXボックス160のうちの1つの内のムービーフラグメント164のうちの1つの中に存在するものとしてシグナリングし得る。

30

【0077】

【0087】SIDXボックス162は、ビデオファイル150の任意の要素である。すなわち、3GPPファイルフォーマット、または他のそのようなファイルフォーマットに準拠するビデオファイルは、必ずしもSIDXボックス162を含むとは限らない。3GPPファイルフォーマットの例によれば、SIDXボックスは、セグメント（たとえば、ビデオファイル150内に含まれるセグメント）のサブセグメントを識別するために使用され得る。3GPPファイルフォーマットは、サブセグメントを、「対応するメディアデータボックスと、ムービーフラグメントボックスによって参照されるデータを含むメディアデータボックスとをもつ、1つまたは複数の連続するムービーフラグメントボックスの自己完結型セットは、そのムービーフラグメントボックスに続き、同じトラックに関する情報を含む次のムービーフラグメントボックスに先行しなければならない」と定義する。3GPPファイルフォーマットはまた、SIDXボックスが、「ボックスによって文書化さ

40

50

れる（サブ）セグメントのサブセグメントへの参照のシーケンスを含む。参照されたサブセグメントは、プレゼンテーション時間において連続する。同様に、セグメントインデックスボックスによって参照されるバイトは、常にセグメント内で連続する。参照されたサイズは、参照された材料中のバイト数のカウントを与える」ことを示している。

【0078】

【0088】S I D Xボックス162は、概して、ビデオファイル150中に含まれるセグメントの1つまたは複数のサブセグメントを表す情報を提供する。たとえば、そのような情報は、サブセグメントが開始および／または終了する再生時間、サブセグメントのバイトオフセット、サブセグメントがストリームアクセスポイント（S A P）を含む（たとえば、それで開始する）かどうか、S A Pのタイプ（たとえば、S A Pが瞬時デコードリフレッシュ（I D R）ピクチャであるか、クリーンランダムアクセス（C R A）ピクチャであるか、切断リンクアクセス（B L A）ピクチャであるかどうかなど）、サブセグメント中の（再生時間および／またはバイトオフセットに関する）S A Pの位置などを含み得る。

10

【0079】

【0089】ムービーフラグメント164は、1つまたは複数のコード化ビデオピクチャを含み得る。いくつかの例では、ムービーフラグメント164は、1つまたは複数のピクチャグループ（G O P）を含み得、G O Pの各々は、いくつかのコード化ビデオピクチャ、たとえばフレームまたはピクチャを含み得る。さらに、上記で説明したように、ムービーフラグメント164は、いくつかの例では、シーケンスデータセットを含み得る。ムービーフラグメント164の各々は、ムービーフラグメントヘッダボックス（M F H D、図4に図示せず）を含み得る。M F H Dボックスは、ムービーフラグメントのシーケンス番号などの、対応するムービーフラグメントの特性を記述し得る。ムービーフラグメント164は、ビデオファイル150中のシーケンス番号の順に含まれ得る。

20

【0080】

【0090】M F R Aボックス166は、ビデオファイル150のムービーフラグメント164内のランダムアクセスポイントを記述し得る。これは、ビデオファイル150によってカプセル化されたセグメント内の特定の時間ロケーション（すなわち、再生時間）へのシークを実施することなどの、トリックモードを実施するのを支援し得る。M F R Aボックス166は、概して随意であり、いくつかの例では、ビデオファイル中に含まれる必要がない。同様に、クライアントデバイス40などのクライアントデバイスは、ビデオファイル150のビデオデータを正しく復号し表示するために、必ずしもM F R Aボックス166を参照する必要がない。M F R Aボックス166は、ビデオファイル150のトラックの数に等しいか、またはいくつかの例ではビデオファイル150のメディアトラック（たとえば、非ヒントトラック）の数に等しい、いくつかのトラックフラグメントランダムアクセス（T F R A）ボックス（図示せず）を含み得る。

30

【0081】

【0091】いくつかの例では、ムービーフラグメント164は、I D Rピクチャなどの、1つまたは複数のストリームアクセスポイント（S A P）を含み得る。同様に、M F R Aボックス166は、S A Pのビデオファイル150内のロケーションの指示を提供し得る。したがって、ビデオファイル150のS A Pからビデオファイル150の時間サブシーケンスが形成され得る。時間サブシーケンスはまた、S A Pに従属するPフレームおよび／またはBフレームなどの、他のピクチャを含み得る。時間サブシーケンスのフレームおよび／またはスライス、サブシーケンスの他のフレーム／スライスに依存する時間サブシーケンスのフレーム／スライスが適切に復号され得るように、セグメント内に構成され得る。たとえば、データの階層構成において、他のデータの予測のために使用されるデータも時間サブシーケンス中に含まれ得る。

40

【0082】

【0092】図5は、例示的なグラフィックスランゲージトランスミッションフォーマット（g l T F : Graphics Language Transmission Format）のシーングラフ構造を

50

示すグラフである。没入型 3 次元 (3D) ユーザ体験を可能にするための技術の 1 つは、シーン記述である。シーン記述は、3D シーンの構成を記述するために使用され得、ここで、シーン記述は、シーン内の様々な 2 次元 (2D) および 3D アセットを参照するおよび位置付けるデータを含む。シーン記述において提供される情報は、次いで、現実的なシーンを生成する物理ベースレンダリング (PBR) のような技法を使用して、3D シーンを適切にレンダリングするようにプレゼンテーションエンジンによって使用され得る。

【0083】

【0093】シーン記述は、通常、シーンのジオメトリのオブジェクトベース階層を表す、有向非巡回グラフ、典型的には、プレインツリー構造であるシーングラフを含む。グラフの葉ノードは、画像、テクスチャ、またはメディアデータバッファなどの幾何プリミティブを表す。グラフにおける各ノードは、その子へのポインタを保持する。子ノードは、とりわけ、他のノード、ジオメトリ要素、変換行列、メディアデータバッファへのアクセス、レンダリングのためのカメラ情報などの群であり得る。

10

【0084】

【0094】空間変換は、グラフのノードとして表され、変換行列によって表される。変換ノードの典型的な使用法は、その子ノードにおけるオブジェクトの回転、平行移動、または拡大縮小を記述することである。シーングラフは、アニメーションプロパティを経時的に変化させることを可能にするアニメーションノードもサポートし、したがって、動的コンテンツを記述する。

【0085】

20

【0095】シーングラフのアーキテクチャは、たとえば、レンダリングのためにグラフを横切る (traversing) 間、処理の複雑さの減少という利点を有する。グラフ表現によって簡略化されている例示的な動作は、カリング (culling) 動作であり、ここで、親ノードの空間が現在のビュー錐台 (view frustum) のレンダリングに対して可視でないかまたは関連しない (詳細度カリング) と見なされる場合、グラフの分岐 (branch) は、処理から落とされる (are dropped)。

【0086】

【0096】(典型的には、ゲームエンジン、VFX 設計ツール、または AR / VR オーサリングツールの心臓に) シーン記述のための多くの独自のソリューションがあるが、いくつかのソリューションは、標準化もされている。詳細には、2001 年に、XML 構文を使用する仮想現実モデリング言語 (VRML) が、WEB 使用法のために標準化される最初のシーン記述ソリューションであった。後に、2005 年にリリースされた OpenGL を使用するオープンソースプロジェクトである Open Scene Graph が、いくつかのコンピュータゲーム、および Delta 3D または Flight Gear などのレンダリングプラットフォームにおける構成要素として使用されている。2010 年に、ISO / IEC によって標準化された X3D は、VRML の後継になった。X3D は、シーングラフ記述のためにバイナリフォーマットと JSON フォーマットとを導入し、マルチテクスチャレンダリング、シェーディング、リアルタイム環境ライトニング、およびカリングなどの新しい能力を特徴付けた。2015 年に、OpenGL も管理する Khronos グループが、グラフィックスライブラリトランスミッションフォーマット (glTF: Graphics Library Transmission Format) をリリースした。glTF は、3D コンテンツツールおよびサービスのための共通記述フォーマットとして効率的で相互運用可能であるように意図された JSON フォーマットに基づくシーン記述フォーマットである。

30

40

【0087】

【0097】図 5 に示されたものなどの典型的な glTF のシーングラフは、コンテンツプロパティと、コンテンツデータへのアクセスと、可能な動的アニメーションと、レンダリングのためのパラメータとを記述するノードのツリーである。図 5 の例では、シーングラフ構造は、子ノード 172 を有するシーンノード 170 を含む。ノード 172 は、子のカメラノード 174 と、メッシュノード 176 と、スキーンノード 178 とを有する。メッシュ

50

ュノードは、子のマテリアルノード 180 と、アクセサノード 182 とを有する。アニメーションノード 184 は、ノード 172 およびアクセサノード 182 について述べられるやはり親であり、スキンノード 178 は、アクセサノード 182 へのエッジを有する。アクセサノード 182 は、子のバッファビューノード 188 を有し、これ自体は、子のバッファノード 194 を有する。マテリアルノード 180 は、子のテクスチャノード 186 を有する。テクスチャノード 186 は、子のサンブラノード 190 と、画像ノード 192 とを有する。

【0088】

【0098】典型的には、カメラノード 174 は、レンダリングが行われるはずであるビューを記述するデータを含む。3Dオブジェクトは、メッシュノード 176 などのメッシュサブノードに記述されてもよく、その子ノードは、バッファを通じたメディアデータへのアクセスの仕方に関する情報と、テクスチャ情報（たとえば、アクセサノード 182、バッファビューノード 188、バッファノード 194、マテリアルノード 180、テクスチャノード 186、サンブラノード 190、および画像ノード 192）とを提供する。アニメーションノード 184 などのアニメーションノードは、経時的に 3D オブジェクトの動的变化を記述するデータを含み得る。

【0089】

【0099】g l T F 2.0 は、交換可能および相互運用可能なシーン記述のために堅実で効率的なベースラインを提供する。しかしながら、g l T F 2.0 は、静的なシーンおよびアセットに伝統的に焦点を当ててきており、これにより、没入型環境における動的でリッチな 3D シーンの要求および必要に対処するのを不適にさせている。

【0090】

【0100】g l T F 2.0 における特定のギャップは、以下を含む。

【0091】

- ・ ビデオおよび移動メッシュおよびポイントクラウドのような時限 (timed) メディアについてのサポートはない。

【0092】

- ・ オーディオに対するサポートはない。

【0093】

- ・ シーンおよびシーン内のアセットとのインタラクションのための限定されたサポート。

【0094】

- ・ たとえば AR 体験にとって極めて重要であるローカルおよびリアルタイムメディアに対するサポートはない。

【0095】

【0101】M P E G は、家庭で成長した古いシーングラフソリューションとゼロから新しいものを定義する可能性とを含めて、異なる代替を注意深く評価した後に没入型マルチメディアのためのそのソリューションを開発するために、g l T F 2.0 が提供する拡張機構を活用することが決定されている。

【0096】

【0102】図 6 は、様々な構成要素間の重要なインターフェースを含む例示的な M P E G - I アーキテクチャを示すブロック図である。この例では、M P E G - I アーキテクチャは、シーン記述および更新 200 と、メディアアクセス機能 (M A F) 206 と、フォーマットプラグイン 216 と、プロトコルプラグイン 218 と、メディアアクセスアプリケーションプログラミングインターフェース (A P I) 204 と、M A F プラグイン 202 と、プレゼンテーションエンジン 220 と、ネットワーク 210 と、ローカルストレージ 208 と、デコーダ 212 A ~ 212 D (デコーダ 212) と、テクスチャバッファ 214 A と、カラーバッファ 214 B と、インデックスバッファ 214 C と、メタデータバッファ 214 D (バッファ 214) とを含む。シーン記述および更新 200 は、図 5 のシーングラフに対応し得る。

10

20

30

40

50

【0097】

【0103】MPEGは、没入型メディアプレゼンテーションの消費のためのエントリポイントとして働くシーン記述に関する作業をガイドするために、MPEG-Iのためのアーキテクチャを発展させた。設計は、メディアアクセスおよびレンダリングパイプライン全体を通じてのデータ交換のための手段としてバッファ（たとえば、テクスチャバッファ214A、カラーバッファ214B、インデックスバッファ214C、およびメタデータバッファ214D）に主に焦点を当てている。それは、バッファ214を通してアクセス可能にされるシーン記述および更新200によって参照されるメディアを要求するためにメディアアクセスAPI204も定義する。この設計は、gltf2.0の原理と上手く整合し、ビデオ復号インターフェースとよく統合する。

10

【0098】

【0104】シーン内の時限メディアおよびメタデータへのアクセスを与えるために、時限アクセサ（MPEG__アクセサ__時限拡張）を定義するようにgltf拡張が指定される。gltf内のアクセサは、バッファビューを通して見られるバッファ内に記憶されたようなデータのタイプおよびレイアウトを定義する。時限データがバッファから読み出されるとき、バッファ内のデータは、時間とともに動的に変化することが予想される。バッファ要素は、時限データとともに使用されるサーキュラバッファのサポートを追加するために拡張される。MPEG拡張によれば、時限メディアおよび／またはメタデータを含むシーンは、データにアクセスするために時限アクセサ拡張を使用するものとする。時限アクセサは、基礎をなすデータバッファが動的であることを示すための通常のアクセサの拡張である。

20

【0099】

【0105】図7は、例示的なサーキュラバッファを示す概念図である。時限データアクセスをサポートするために、バッファ要素は、サーキュラバッファ機能を提供するようにMPEGによって拡張されている。拡張は、「MPEG__サーキュラ__バッファ」と名付けられ、「バッファ」構造の一部として含まれ得る。時限データへのアクセスを提供するバッファは、「MPEG__サーキュラ__バッファ」拡張を含むことができる。

【0100】

【0106】図7の例では、サーキュラバッファ240は、空データ234A、234B、（これらは使用可能なメディアデータを含まないサーキュラバッファ240の一部に対応し得る）と、それぞれのヘッダ230A～230Dとペイロード232A～232Dとを含む復号されたメディアフレームとを含む。この例では4つのフレームが示されているが、概して、サーキュラバッファ240は、サーキュラバッファ240に割り振られたメモリの量およびフレームのサイズに従って、任意の個数のフレームを記憶することができる。

30

【0101】

【0107】図1のクライアントデバイス40は、読み出しポインタ236と書き込みポインタ238とを維持することができる。概して、読み出しポインタ236は、読み出される次のメディアフレーム（この場合、ヘッダ230Aとペイロード232Aとを含むフレーム）の開始点を識別し、一方、書き込みポインタ238は、続くフレームが書き込まれることが可能であるサーキュラバッファ240の一部を表す。上記で説明したように、クライアントデバイス40は、バッファオーバーフローとアンダーフローとを回避するために、読み出しポインタ236が書き込みポインタ238を通り過ぎないように、およびまた書き込みポインタ238が読み出しポインタ236を追い越すのを防ぐように、サーキュラバッファ240を管理し得る。

40

【0102】

【0108】この例では、クライアントデバイス40は、記憶されるいくつかのフレーム、各フレームの寸法、およびフレームのテクセルごとのサンプルフォーマットに基づいて、サーキュラバッファ240にメモリを割り振ることができる。複数のフレームは、バッファをスワップし、バッファへのスムーズなアクセスを確実にするために使用され得る。ク

50

クライアントデバイス 40 のアクセサは、示されるように、読み出しポインタ 236 と書き込みポインタ 238 の両方を維持することができる。フレームは、レンダリングのために読み出しポインタ 236 で読み出される。メディアデコーダからの新たな入力フレームは、書き込みポインタ 238 において挿入される。そのフレーム内の先のデータは上書きされてもよく、フレームバッファはそれに応じてサイズ変更されてもよい。クライアントデバイス 40 は、バッファにオフセットとして読み出しポインタ 236 および書き込みポインタ 238 についての値を記憶することができる。

【0103】

【0109】クライアントデバイス 40 のレンダラは、タイムスタンプ（書き込み__ポインタ）>タイムスタンプ（読み出し__ポインタ）であることを確実にすることができる。バッファ中のフレームを新しい復号されたビデオフレームで上書きするとき、レンダラは、読み出し__ポインタが最も古いタイムスタンプを有するフレームに移動されることを確実にし得る。これにより、フレームドロップをもたらし得るが、バッファ内の同じフレームへの同時アクセスが実行されないことも確実にし得る。

【0104】

【0110】MPEG__メディアによって識別される MPEG メディア拡張は、シーンに使用される MPEG メディアアイテムの配列を提供する。

【0105】

【0111】「MPEG__ビデオ__テクスチャ」によって識別される MPEG ビデオテクスチャ拡張は、MPEG__メディアオブジェクトによってリストされるメディアおよびそのそれぞれのトラックに g l T F テクスチャオブジェクトをリンクする可能性を提供する。MPEG ビデオテクスチャ拡張は、時限アクセサオブジェクトを使用して、復号される時限テクスチャが利用可能にされる時限アクセサへの参照も提供する。

【0106】

【0112】図 8 は、例示的な MPEG-I シーン記述アーキテクチャを示すブロック図である。MPEG-I シーン記述ソリューションは、異なる入力モダリティからのメディアを混合するシーンを記述するように設計されている。これらの異なるモダリティとのインタラクションは、シーン記述を消費するプレゼンテーションエンジンへの専用インターフェースを介して保証される。見る人のローカルな現実世界とのインタラクションは、i-1 および i-i インターフェースを介して実行される。これらのインターフェースは、OpenXR などの既存の API を活用することができる。

【0107】

【0113】この例では、MPEG-I シーン記述アーキテクチャは、ローカルカメラ 250 と、ローカルマイクロフォン 252 と、コントローラデバイス 254 と、センサ 256 と、シーングラフ 262 と、シーングラフ更新 264 と、2D/3D メディアデータ 258 と、プレゼンテーションエンジン 260 とを含む。ローカルカメラ 250 は、仮想シーンのためのアンカーポイントとして使用され得るローカルの現実世界のオブジェクトの画像をキャプチャするために使用され得る。ローカルマイクロフォン 252 は、ユーザ音声をキャプチャするために使用され得る。コントローラデバイス 254 は、ユーザ入力を受信するための 1 つまたは複数のボタン、トラックパッドなどを含むゲームコントローラデバイスであり得る。センサ 256 は、ジャイロスコープ、動きセンサ、光センサ、またはユーザの姿勢、仮想カメラの向き、ユーザのロケーションなどを検出するための他のタイプのセンサを含み得る。2D/3D メディアデータ 258 は、仮想シーン、たとえば、壁、床、天井、椅子、テーブル、または任意の他の仮想オブジェクトを含む 1 つまたは複数の仮想オブジェクトと、仮想シーン内のオブジェクトのロケーションとをレンダリングするためのデータを含み得る。シーングラフ 262 およびシーングラフ更新 264 は、図 6 のシーン記述および更新 200 ならびに図 5 のシーングラフに対応し得る。

【0108】

【0114】OpenXR は、広範囲の XR デバイスに対処する XR アプリケーションを開発するためのアプリケーションプログラミングインターフェース（API）である。XR

は、人間によるインタラクションを介してコンピュータによって生成される現実世界環境と仮想世界環境との混合を指す。XRは、仮想現実（VR）、拡張現実（AR）、および複合現実（MR）などの技術を含む。OpenXRは、アプリケーションとXRランタイムとの間のインターフェースである。ランタイムは、フレーム構成、ユーザにより引き起こされるアクション、およびトラッキング情報などの機能进行处理する。

【0109】

【0115】OpenXRは、レイヤードAPI（layered API）であるように設計されており、これは、ユーザまたはアプリケーションがアプリケーションとランタイム実装との間にAPIレイヤを挿入することができることを意味する。これらのAPIレイヤは、上のレイヤからOpenXR関数をインターセプトし、次いで、その層がない場合にその層なしに実行されるものとは異なる動作を実行することによって、追加の機能を提供する。最も単純な場合には、レイヤは、同じ引数を有する次のレイヤを単に呼び出す。より複雑なレイヤは、レイヤまたはその下のランタイムに存在しないAPI機能を実装することができる。この機構は、本質的に、OpenXRに設計されているアーキテクチャ化された「ファンクションシミング」または「インターセプト」機能であり、APIコールを「フッキング」するより非公式な方法を置き換えることを意味する。

【0110】

【0116】アプリケーションは、利用可能なAPIレイヤのリストを取得するためにxrEnumerateApiLayerProperties関数を呼び出すことによってアプリケーションにとって利用可能であるAPIレイヤを決定することができる。次いで、アプリケーションは、インスタンスを生成するときに、このリストから所望のAPIレイヤを選択し、それらをxrCreateInstance関数に提供することができる。

【0111】

【0117】APIレイヤは、基礎をなす（underlying）ランタイムによってサポートされてもされなくてもよいOpenXR関数を実装することができる。これらの新しい機能をエクスポートするために、APIレイヤは、この機能をOpenXR拡張の形でエクスポートしなければならない。それは、関連付けられた拡張なしで新しいOpenXR関数をエクスポートしてはならない。

【0112】

【0118】OpenXRインスタンスは、OpenXRアプリケーションがOpenXRランタイムと通信することを可能にするオブジェクトである。アプリケーションは、xrCreateInstanceを呼び出し、結果として生じるXrInstanceオブジェクトへのハンドルを受信することによってこの通信を達成する。

【0113】

【0119】XrInstanceオブジェクトは、OpenXR関連アプリケーションの状態を、任意のそのような状態をアプリケーションのグローバルアドレス空間に記憶することなく、記憶し、追跡する。これは、このオブジェクトがアプリケーションに不透明であるので、アプリケーションが複数のインスタンスを生成するとともに、アプリケーションのOpenXRの状態を安全にカプセル化することを可能にする。OpenXRランタイムは、作成および使用され得る同時XrInstanceオブジェクトの数を制限することができるが、それらは、プロセスごとに少なくとも1つのXrInstanceオブジェクトの生成および使用法をサポートしなければならない。

【0114】

【0120】スペースは、アプリケーションが作成し、次いで、APIコールにおいて使用するXrSpaceハンドルによって表される。アプリケーションが座標を返す関数を呼び出すときはいつでも、それらの座標が表現される基準系を指定するためにXrSpaceを提供する。同様に、関数に座標を提供するとき、アプリケーションは、ランタイムがこれらの座標を解釈するためにどのXrSpaceを使用すべきかを指定する。

【0115】

10

20

30

40

50

【0121】Open XRは、アプリケーションがそれらの空間推論を自動実行するために使用するよく知られている基準空間のセットを定義する。これらの基準空間は、VIEW、LOCAL、およびSTAGEである。各基準空間は、その原点が位置する場所とその軸が向けられる方法とを確立するよく定義された意味を有する。

【0116】

【0122】ランタイムの追跡システムが経時的に世界のその理解を改善するランタイムは、独立して空間を追跡することができる。たとえば、LOCAL空間およびSTAGE空間がそれらの原点を世界における静的な位置にそれぞれマッピングするとしても、インサイドアウト追跡システムを有するランタイムは、各原点を所定の位置に維持するために、継続的に各空間の原点に対する微調整を導入することができる。

10

【0117】

【0123】よく知られている基準空間を超えて、ランタイムは、経時的にモーションコントローラのポーズを追跡するポーズアクション空間などの他の独立して追跡される空間をエクスポーズする。

【0118】

【0124】上記で説明したように、シーングラフ262とシーングラフ更新264とを含むシーン記述は、没入型メディアデータを消費するためのデータセットにおけるエントリポイントとして働く。シーン記述は、シーンのすべての仮想オブジェクトと、それらの空間的関係および時間的関係とを記述する。現在、シーン記述（そして他の要素はない）は、シーンの仮想オブジェクトと現実世界のオブジェクトとの間の関係を記述する。しかしながら、ARアプリケーションとMRアプリケーションとをサポートするために、本開示は、仮想オブジェクトと現実世界のオブジェクトとの間のそのような関係を説明する必要性を認識する。

20

【0119】

【0125】より詳細には、本開示は、現実世界空間にシーンをアンカーするための技法を説明する。一例では、使用されるアンカーは、Open XR空間にマッピングされ得る。

【0120】

【0126】図9は、シーン記述によって示される例示的なアンカーXR空間を示す概念図である。本開示の技法によれば、シーン記述ノードは、シーンがXR空間にアンカーされることを示すその空間への参照を含み得る。アンカーXR空間は、基準空間、たとえば、ローカル、ビュー、またはステージであり得る。代替として、アンカーXR空間は、アプリケーションまたはユーザ入力に基づいて実行時にセットアップされるアプリケーション定義空間であってもよい。シーン記述は、アンカーXR空間のタイプを示す。アンカー空間がアプリケーション定義空間である場合、シーン記述は、アプリケーションが守らせるべきXR空間に対するいくつかの制約を提供することができる。

30

【0121】

【0127】シーン記述は、シーン座標系のアンカーXR空間への変換を示し得る。たとえば、シーン記述座標空間の原点の回転、平行移動、および拡大縮小は、XR空間にアライメントする（align）ように実行され得る。何も変換がない場合、シーン記述の座標系は、アンカー空間のものとアライメントされる。

40

【0122】

【0128】Open XRのようなXRランタイムシステムは、XR空間に対する境界空間のクエリを可能にする。シーン記述は、プレゼンテーションエンジンがシーンの広がり、すなわちシーンのバウンディングボックスを、アンカーXR空間のバウンディングボックスにアライメントすることを要求することができる。シーン記述は、このアライメント（alignment）が2つの空間を一致させるための拡大縮小を含むか否かを示すことができる。拡大縮小は、両方の次元であってもよく、または拡大縮小は、シーンの空間的な広がりとの関係を保存することができる。拡大縮小が適用されない場合、プレゼンテーションエンジンは、シーンバウンディングボックスの長いエッジをXR空間の長いエッジにアライメントし、次いで、XR空間バウンディングボックスの中心と一緒に置かれるようにシー

50

ンバウンディングボックスをセンタリングする。

【0123】

【0129】図9の例では、アンカーXR空間は、見る人のリビングルームの床に対応するタイプ「ステージ」である。

【0124】

【0130】本開示の技法によれば、MPEG__シーン__アンカー拡張が、シーンノードにg l T F 2 . 0拡張として追加され得る。ARシーンは、現実世界のXR空間へのシーンのアンカリングを記述するためのMPEG__シーン__アンカー拡張を含むことができる。この拡張は、extensionsUsed要素において宣言されてもよく、g l T F 2 . 0シーンのextensionsRequired要素に含まれ得る。

10

【0125】

【0131】たとえば、図1のクライアントデバイス40によって実行されるプレゼンテーションエンジンがシーンアンカー拡張をサポートしないとき、プレゼンテーションエンジンは、完全な仮想シーンをレンダリングすることに戻り得る。

【0126】

【0132】以下の表1は、MPEG__シーン__アンカー拡張の要素の例示的なセットと、それらのプロパティとを表す。

【0127】

【表1-1】

表1

20

名前	タイプ	デフォルト	説明
xrReferenceSpace Type	列挙	STAGE	基準タイプは、VIEW=1、LOCAL=2、STAGE=3、またはAPPLICATION=4のうちの1つであり得る。
aligned (アライメントされた)	列挙	NOT_ ALIGNED	アライメントされたフラグは、以下の値、すなわち、NOT__ALIGNED=0、ALIGNED__NOTSCALED=1、ALIGNED__SCALED=2のうちの1つをとることができる。ALIGNED__SCALEDが設定される場合、シーンバウンディングボックスは、XR空間のバウンディングボックスにアライメントされる。
(aligned==NOT__ ALIGNED)の場合、{			
transformation (変換)	行列4×4	N/A	シーン空間をXR空間にアンカーするためのシーン空間の変換。
position(位置)	配列(番号)		XR空間におけるシーン座標系の原点の位置。

30

40

【0128】

50

【表 1 - 2】

orientation (向き)	配列(番号)		四元数はアンカー空間におけるシーンの 回転を記述する。 centerPosition(中心位置)および orientation(向き)は、 変換の代替として使用される。
}			
(xrReference Type== APPLICATION)の 場合、{			
xrActionSpace	ストリング		アプリケーションは、たとえば コントローラアクションに対応する アクション空間を定義した。 アンカーとして使用されるXRActionは、 パス、たとえば、「入力/サムスティック/ 右」によって識別される。
xrObjectId	番号		アプリケーションとコンテンツサーバとの間で 共有される識別子。クライアントは、 サーバにアプリケーションのリストを 送ることができ、またはクライアントは、 オブジェクト検出のためにサーバに メディアストリームを送ることができ、 次いで、識別されたオブジェクトおよび それらのロケーションのリストを得る。 共有されたオブジェクトは、 アンカーポイントとして使用され得る。 XRSpaceは、アンカーオブジェクトが 空間内でそれらを追跡するために 作成され得る。
}			
poseFeedback	番号		ポーズ情報が送られる アップリンクストリームを記述するアクセサ。
sceneUnderstandi ngStream	番号		ローカルでキャプチャされたビューが ストリーミングされる アップリンクストリームを記述するアクセサ

10

20

30

40

【0129】

【0133】以下のデータセットは、上述されたシーンアンカー拡張の例示的なインスタンスを表す。

【0130】

【数1】

```

{
  "scene": 0,
  "scenes": [
    {
      "extensions": {
        "MPEG_scene_anchor": {
          "xrReferenceSpaceType": 4, # Application
          "aligned": 0, # NOT_ALIGNED
          "xrActionSpace": "/user/hand/right/input/grip",
          "sceneUnderstandingStream": 15
        }
      },
      "name": "Scene",
      "nodes": [
        759,
        760,
        761,
        762,
        763,
        764
      ]
    }
  ]
}

```

10

20

30

40

50

【0131】

【0134】図10は、本開示の技法を実行し得る別の例示的なクライアントデバイス300を示すブロック図である。クライアントデバイス300は、図1のクライアントデバイス40に対応し得る。この例では、クライアントデバイス300は、ネットワークインターフェース302と、取出しユニット304と、ユーザインターフェースデバイス306と、カメラ308と、センサ310と、マイクロフォン312と、ディスプレイ314と、デコーダ316と、シーンデータ320と、プレゼンテーションエンジン330とを含む。様々なユニットおよびデバイスのいずれかまたは全部は、ソフトウェア、ハードウェア、またはファームウェアに実装され得る。ソフトウェアまたはファームウェアに実装されるとき、命令を記憶するためのメモリおよび命令を実行するための1つまたは複数のプロセッサなどの必須のハードウェアも用意され得ることを理解されたい。ハードウェアは、1つまたは複数のデジタル信号プロセッサ(DSP)、汎用マイクロプロセッサ、特定用途向け集積回路(ASIC)、フィールドプログラマブルロジックアレイ(FPGA)、または他の同等の集積回路もしくはディスクリート論理回路などの1つまたは複数の処

理回路を含み得る。

【0132】

【0135】ネットワークインターフェース302は、一般に、ネットワークを介してデータを送受信するように構成される。たとえば、ネットワークインターフェース302は、イーサネット（登録商標）データを送受信するように構成されたネットワークインターフェースカード（NIC）を含んでもよい。代替として、ネットワークインターフェース302は、様々なIEEE802.11ネットワークプロトコルのいずれかなどの様々なプロトコルのいずれかに従ってネットワークデータを送受信するように構成されたワイヤレスネットワークインターフェースであってもよい。ネットワークインターフェース302は、図1のネットワークインターフェース54に対応し得る。

10

【0133】

【0136】取出し（retrieval）ユニット304は、図1および図2の取出しユニット52に対応し得る。取出しユニット304は、様々な取出しプロトコルに従ってネットワークインターフェース302を介してメディアデータを取り出すように構成され得る。たとえば、取出しユニット304は、DASHに従ってユニキャストを介してメディアデータを取り出すように構成され得る。別の例として、取出しユニット304は、たとえば、eMBMSに従って、ブロードキャストまたはマルチキャストを介してメディアデータを受信するように構成され得る。さらに別の例として、取出しユニット304は、リアルタイムトランスポートプロトコル（RTP）に従ってメディアデータを取り出すように構成され得る。メディアデータを取り出すための他のプロトコルが使用されることも可能である。

20

【0134】

【0137】デコーダ316は、図1のオーディオデコーダ46およびビデオデコーダ48に対応し得る。追加または代替として、デコーダ316は、図6のデコーダ212に対応し得る。概して、デコーダ316は、2次元ビデオ、3次元ビデオ、オーディオ、メタデータ等などの様々なタイプのメディアデータを復号する。デコーダ316は、復号されたメディアデータをシーンデータ320に出力することができる。

【0135】

【0138】シーンデータ320は、メディアデータ322を記憶するための1つまたは複数のメモリ（ストレージデバイス）を表す。加えて、本開示の技法によれば、シーンデータ320は、シーングラフ324も記憶する。シーングラフ324は、図5のシーングラフ、図6のシーン記述および更新200、ならびに／または図8のシーングラフ262およびシーングラフ更新264に対応し得る。シーングラフ324は、（カメラ308および／またはセンサ310によって検出されるような）ユーザがユーザの頭部を動かしたとき、（カメラ308、センサ310、および／またはユーザインターフェースデバイス306によって検出される動きに応答して衝突を通して検出されるような）ユーザが仮想オブジェクトとインタラクトする（たとえば、接触する）とき、またはシーングラフ324において定義されるアニメーションを通してなど、様々な入力に従って更新され得る。

30

【0136】

【0139】シーングラフ324はまた、本開示の技法による少なくとも1つのシーンアンカーも含む。シーンアンカーは、現実世界のプレゼンテーション環境における基準空間、たとえば、ローカル、ビュー、またはステージであり得る。OpenXRの仕様によれば、ビュー空間は、上／下、左／右、および前／後の動きを有する主たる見る人（ユーザ）のためのビュー変換を生成するために使用されるビュー原点を追跡する。局所空間は、上／下、左／右、および前／後の動きを有するピッチおよびロールを除外するように重力アライメントされたワールドロック原点を確立する。ステージ空間は、空であるランタイム定義された平らな矩形空間であり、ステージの軸に対してアライメントされた上／下、左／右、および前／後の動きを伴って床上で矩形の中心に原点を有しつつ上を歩き回られてもよい。

40

【0137】

50

[0140]したがって、ビュー空間は、ユーザビューポートの方向および向きにかかわらず、ディスプレイ上に永続的に提示される仮想オブジェクト、たとえばレチクル、ヘッドアップディスプレイ（HUD）などのために使用され得る。ローカル空間は、現実世界のプレゼンテーション環境の物理的な床に対して配置されていない仮想オブジェクトのために使用され得る。ステージ空間は、ユーザが、インタラクトすることができ、現実世界のプレゼンテーション環境の物理的な床（ステージ）に対して配置することができる他の仮想オブジェクトのために使用され得る。

【0138】

[0141]ユーザインターフェースデバイス306は、ユーザがクライアントデバイス300に入力を与えるためのコントローラまたは他のデバイスを含み得る。そのようなコントローラは、1つまたは複数のボタン、トラックパッド、アナログ制御スティック、アナログトリガなどを含み得る。ユーザ入力デバイス306のこれらの種々の要素を使用して、ユーザは、仮想世界の移動、仮想オブジェクトとのインタラクション（たとえば、ドアまたは収納箱を開くこと）、メニューを開くことなどを表す入力を与えることができる。

【0139】

[0142]カメラ308は、現実世界のプレゼンテーション環境の画像またはビデオデータをキャプチャするために使用されるカメラを表す。いくつかの例では、カメラ308は、現実世界のプレゼンテーション環境内のユーザのロケーションおよび向きを決定するために使用され得る、赤外線ライトボックスなどのクライアントデバイス300に結合されたデバイスから発せられる光を検出することができる。追加または代替として、カメラ308は、床、天井、壁、椅子、テーブル等などの現実世界のオブジェクトを検出することができ、クライアントデバイス300は、現実世界のオブジェクトとの潜在的な衝突をユーザに警告することができる。同様に、センサ310は、現実世界のプレゼンテーション環境におけるユーザの姿勢、ロケーション、および／または向きの情報を検出するためのジャイロスコープおよび／または他のセンサであってもよい。マイクロフォン312は、たとえば、仮想プレゼンテーションまたはゲームのために、たとえば、複数の人が仮想環境とインタラクトしているときに他のユーザに送られるユーザの音声をキャプチャするためのマイクロフォンであってもよい。

【0140】

[0143]ディスプレイ314は、ユーザに視覚出力を提供するためのディスプレイを表す。ディスプレイ314は、ユーザの眼に別個の出力を提供するための2つのスクリーン、たとえば、左眼スクリーンと右眼スクリーンとを含むことができる。代替として、ディスプレイ314は、左眼提示と右眼提示との間で迅速に交替するように構成されている単一ディスプレイを含んでもよく、ユーザの眼は、ディスプレイと同期されたやり方で交互にシャッタされてもよい。

【0141】

[0144]この例におけるプレゼンテーションユニット330は、アンカーポイント検出ユニット332と変換ユニット334とを含む。プレゼンテーションユニット330は、アンカーポイントタイプ、たとえば、ローカル、ビュー、またはステージを決定するためにシーングラフ324からデータを受信することができる。プレゼンテーションユニット330は、カメラ308からの画像入力と、センサ310からのセンサデータとを受信することもできる。アンカーポイント検出ユニット332は、シーングラフ324のデータから使用されるアンカーポイントのタイプを決定し、カメラ308からの画像データを使用して対応する現実世界のアンカーポイントを識別することができる。プレゼンテーションユニット330は、仮想シーンを現実世界のプレゼンテーション環境とアライメントするために、現実世界のプレゼンテーション環境において識別されたアンカーポイントを使用することができる。たとえば、現実世界のアンカーポイントは、床、表面（たとえば、テーブル）などの上の点であり得る。いくつかの例では、現実世界のアンカーポイントを表すために、クイックレスポンス（QR）コードなど、現実世界のオブジェクト上の視覚マーカが使用され得る。

【0142】

[0145]いくつかの例では、プレゼンテーションユニット330は、クライアントデバイス300によって実行されるアプリケーション（図示せず）から、現実世界のプレゼンテーション環境に対応するアプリケーション定義空間を表すデータを受信することができる。たとえば、ユーザは、アプリケーション定義空間を定義するためにユーザインターフェースデバイス306を使用することができる。プレゼンテーションユニット330は、アプリケーション定義空間を表すデータをアプリケーションから受信し得る。アプリケーション定義空間は、現実世界のアンカーポイントを含み得る。

【0143】

[0146]変換ユニット334は、現実世界のプレゼンテーション環境に一致するように仮想シーンを変換することができる。たとえば、変換ユニット334は、仮想シーンの向きと現実世界のプレゼンテーション環境の向きとを決定し、次いで、仮想シーンおよび現実世界のプレゼンテーション環境の向きがアライメントされるように仮想シーンを変換することができる。すなわち、シーングラフ324は、シーン座標系から現実世界のプレゼンテーション環境の現実世界の座標系への変換を表すデータを含むことができる。そのような変換は、回転および／または平行移動を含み得る。追加または代替として、変換ユニット334は、拡大縮小変換に従って現実世界のプレゼンテーション環境内に適合するように仮想シーンを拡大または縮小してもよい。

【0144】

[0147]仮想シーンアンカーポイントを現実世界のアンカーポイントにアライメントさせ、任意の必要な変換を行った後、プレゼンテーションユニット330は、ディスプレイ314を介してメディアデータ322を提示することができる。たとえば、メディアデータ322は、仮想オブジェクトと、テクスチャと、色と、仮想オブジェクトについての位置とを定義するデータを含んでもよい。プレゼンテーションユニット330は、（カメラ308およびセンサ310から受信されたデータにより決定される）ユーザの仮想カメラの向きに対して、対応するロケーションに仮想オブジェクトを提示することができる。プレゼンテーションユニット330は、ユーザインターフェースデバイス306、カメラ308、および／またはセンサ310から検出されたユーザの動きに従って、ならびに／あるいはシーングラフ324へ更新されたものに基づいて、プレゼンテーションを更新することができる。

【0145】

[0148]たとえば、ユーザは、たとえば、仮想オブジェクトと衝突することによって、または仮想オブジェクトと相互作用するためにボタンを押すことによって、仮想オブジェクトとインタラクトすることができる。プレゼンテーションユニット330は、そのようなアクション、たとえば、仮想オブジェクトの移動、回転、変形などに応答して、仮想オブジェクト上にレンダリングするためのアニメーションを決定することができる。そのようなアニメーションは、たとえば、図5の例示的なシーングラフのアニメーションノード184において定義され得る。したがって、続いて出力されるフレームにおいて、プレゼンテーションユニット330は、カメラ308およびセンサ310から受信されたデータによって決定されるユーザの眺めに対しての仮想オブジェクトのロケーションと向きとを更新することができる。そのような更新は、光源のロケーション、ならびに仮想オブジェクトについてのテクスチャおよび色の情報に基づく仮想オブジェクトの提示に対する修正をさらに含み得る。

【0146】

[0149]いくつかの例では、シーングラフ324は、表1に関して上記で説明したMPEG__シーン__アンカーのためのデータを含み得る。表1に示されるように、MPEG__シーン__アンカーは、poseFeedback要素と、sceneUnderstandingStream要素とを含むことができる。カメラ308およびセンサ310によって検出されるようにユーザのポーズが変化するとき、クライアントデバイス300は、ネットワークインターフェース302を使用するネットワークインターフェースを使用し

て `poseFeedback` 要素によって示されるアップリンクストリームを介してポーズ情報更新を送ることができる。308によってキャプチャされたローカルキャプチャされたビューがアップロードされることになる場合、クライアントデバイス300は、ネットワークインターフェース302を使用して、`sceneUnderstandingStream`要素によって示されたアップリンクストリームを介してローカルキャプチャされたビューを送ることができる。

【0147】

【0150】このようにして、クライアントデバイス300は、仮想シーン内の1つまたは複数の仮想オブジェクトを定義するメディアデータを記憶するように構成されたメモリと、回路内で実装された1つまたは複数のプロセッサと、を含む、メディアを提示するためのデバイスの一例を表し、1つまたは複数のプロセッサは、仮想シーン内の1つまたは複数の仮想オブジェクトとシーンアンカーとを記述するデータを含むビットストリームのシーン記述を受信し、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表しており；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定し；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションに1つまたは複数の仮想オブジェクトを提示するように構成される。

【0148】

【0151】図11は、本開示の技法による例示的な方法を示すフローチャートである。図11の方法は、たとえば、図1のクライアントデバイス40、図6または図8のMPEG-Iアーキテクチャに適合するデバイスもしくはデバイスのセット、あるいは図10のクライアントデバイス300によって実行され得る。例示および説明のために、図11の方法は、図10のクライアントデバイス300に関して説明される。しかしながら、他のデバイスまたはデバイスのセットがこの方法または同様の方法を実行するように構成され得ることを理解されたい。

【0149】

【0152】最初に、クライアントデバイス300は、シーン記述を含むビットストリームを受信することができる(350)。たとえば、クライアントデバイス300は、DASH、RTP、eMBMS、または他のユニキャスト、ブロードキャスト、もしくはマルチキャストのネットワークプロトコルに従ってビットストリームを取り出すことができる。すなわち、シーン記述を受信することは、CD-ROM、DVD、ブルーレイ(登録商標)ディスクなどの伝達媒体または記憶媒体などのコンピュータ可読媒体、ハードドライブまたはフラッシュドライブなどのストレージデバイスなどからネットワークを介したワイヤレスまたはワイヤード送信および受信などのシーン記述を取得するための任意の方法を含み得る。ビットストリームは、壁、天井、床、ドア、テーブル、椅子等などの1つまたは複数の仮想オブジェクトについてのメディアデータをさらに含み得る。シーン記述は、シーングラフ更新によって定期的に更新されるシーングラフを含むことができる。たとえば、シーン記述は、図5の例示的なシーングラフ、図6のシーン記述および更新200、図8のシーングラフ262およびシーングラフ更新264、または図10のシーングラフ324に対応し得る。

【0150】

【0153】クライアントデバイス300のプレゼンテーションユニット330は、シーン記述からアンカーポイントを決定することができる(352)。本開示の技法によれば、シーン記述は、表1に関して上記で説明したようにMPEG__シーン__アンカーなどのアンカーポイントについてのデータを含み得る。アンカーポイントについてのデータは、たとえば、基準空間タイプ(たとえば、現実世界のプレゼンテーション環境)が基準空間のビュー、ローカル、ステージ、またはアプリケーションタイプとなるべきかを示すデータを含み得る。アンカーポイントについてのデータは、シーンデータが現実世界のプレゼンテーション環境に一致するように変換(たとえば、回転、平行移動、および/または拡大縮小)されるべきかをさらに示し得る。アンカーポイントは、コントローラを介して受信

された動き、またはユーザによって装着されたデバイスの現実世界の再配置など、ユーザが実行し得る様々なアクションをさらに含み得る。アンカーポイントは、プレゼンテーションアプリケーションとサーバとの間で共有されるオブジェクト識別子を含むこともできる。そのようなオブジェクトは、アンカーポイントとして使用され得る。概して、シーン記述は、シーンアンカーポイントを、床上の特定のロケーション（たとえば、床の中間点）などの現実世界のアンカーポイントに関連付けるデータを含み得る。

【0151】

【0154】プレゼンテーションユニット330は、現実世界のプレゼンテーション環境を表すデータを受信することもできる（354）。いくつかの例では、このデータは、ユーザから入力データを受信するアプリケーションによって提供され得る。たとえば、ユーザは、たとえば、ユーザインターフェースデバイス302を用いて現実世界のプレゼンテーション環境を物理的にまたは仮想的に追跡することによって、現実世界のプレゼンテーション環境を定義するために、ユーザインターフェースデバイス306を使用することができる。いくつかの例では、プレゼンテーションユニット330は、カメラ308を使用して現実世界のプレゼンテーション環境を自動的に検出することができる。いくつかの例では、クライアントデバイス300は、カメラ308によってキャプチャされた画像および／またはビデオデータを受信し、シーン記述によって示されるようなsceneUnderstandingStreamを介してこのデータをアップロードし、次いで、ネットワークインターフェース302を介して、現実世界のプレゼンテーション環境における現実世界のアンカーポイントの位置を示すデータを受信することができる。

【0152】

【0155】次いで、プレゼンテーションユニット330は、仮想シーンについてのデータを受信することができる（356）。たとえば、ビットストリームは、オブジェクト自体を定義するデータと、仮想シーン内のオブジェクトのロケーションと、オブジェクトについてのテクスチャと、オブジェクトについての色とを含む1つまたは複数の仮想オブジェクトについてのデータを含み得る。プレゼンテーションユニット330は、決定された現実世界のアンカーポイントにおいて現実世界のプレゼンテーション環境に仮想シーンをアンカーすることができる（358）。プレゼンテーションユニット330は、たとえば、回転、平行移動、および／または拡大縮小を使用して、仮想シーンを現実世界のプレゼンテーション環境とアライメントするように変換することもできる（360）。次いで、プレゼンテーションユニット330は、仮想シーン内の適切な位置において仮想オブジェクトを提示することを含めて、仮想シーンを提示することができる（362）。

【0153】

【0156】ユーザは、ユーザインターフェースデバイス306を使用してまたはカメラ308および／またはセンサ310によって検出され得る物理的移動を使用して、仮想シーンを通して（ユーザの眼を表す）仮想カメラを動かすことができる。そのような動きにตอบสนองして、プレゼンテーションユニット330は、新しい仮想カメラロケーションを考慮するためにプレゼンテーションを更新することができる。たとえば、「ビュー」アンカーされたオブジェクトについては、位置は変更されないままであり得るが、「ローカル」または「ステージ」アンカーされたオブジェクトについては、位置は、ユーザ移動または仮想オブジェクトアニメーションに基づいて変化し得る。

【0154】

【0157】このようにして、図11の方法は、ビットストリームのシーン記述を受信することと、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定することと；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示することを含む方法の一例を表す。

【0155】

【0158】本開示の技法の様々な例が、以下の条項において要約される。

【0156】

【0159】条項1：メディアデータを提示する方法であって、仮想シーン内の1つまたは複数の仮想オブジェクトとシーンアンカーとを記述するデータを含むビットストリームのシーン記述を受信することと、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定することと、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示することとを備える方法。

【0157】

【0160】条項2：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項1に記載の方法。

【0158】

【0161】条項3：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項1に記載の方法。

【0159】

【0162】条項4：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項1から3のいずれかに記載の方法。

【0160】

【0163】条項5：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、請求項4に記載の方法。

【0161】

【0164】条項6：1つまたは複数の仮想オブジェクトを提示することは、仮想シーンを現実世界のプレゼンテーション環境にアライメントさせることを備える、条項1から5のいずれかに記載の方法。

【0162】

【0165】条項7：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項1に記載の方法。

【0163】

【0166】条項8；シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項1に記載の方法。

【0164】

【0167】条項9：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項1に記載の方法。

【0165】

【0168】条項10：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、請求項9に記載の方法。

【0166】

【0169】条項11：1つまたは複数の仮想オブジェクトを提示することは、仮想シーンを現実世界のプレゼンテーション環境にアライメントさせることを備える、条項1に記載の方法。

【0167】

【0170】条項12：メディアデータを提示するためのデバイスであって、項1から11のいずれかに記載の方法を実行するためのものまたは手段を備えるデバイス。

10

20

30

40

50

【0168】

【0171】条項13：1つまたは複数の手段は、回路内で実装された1つまたは複数のプロセッサを備える、条項12に記載デバイス。

【0169】

【0172】条項14：ビットストリームのデータを記憶するように構成されたメモリをさらに備える、条項12に記載デバイス。

【0170】

【0173】条項15：デバイスは、集積回路、マイクロプロセッサ、およびワイヤレス通信デバイスのうちの少なくとも1つを備える、条項12に記載デバイス。

【0171】

【0174】条項16：メディアデータを取り出すためのデバイスであって、仮想シーン内の1つまたは複数の仮想オブジェクトとシーンアンカーとを記述するデータを含むビットストリームのシーン記述を受信するための手段と、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定するための手段と、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示するための手段とを備えるデバイス。

【0172】

【0175】条項17：メディアデータを提示する方法であって、ビットストリームのシーン記述を受信することと、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定することと、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示することと、を備える方法。

【0173】

【0176】条項18：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項17に記載の方法。

【0174】

【0177】条項19：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項17に記載の方法。

【0175】

【0178】条項20：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項17に記載の方法。

【0176】

【0179】条項21：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項20に記載の方法。

【0177】

【0180】条項22：1つまたは複数の仮想オブジェクトを提示することは、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定することと、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換することとを備える、条項20に記載の方法。

【0178】

【0181】条項23：1つまたは複数の仮想オブジェクトを提示することは、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせることを備える、条項17に記載の方法。

【0179】

【0182】条項24：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、方法は、アクションに応答して1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定することをさらに備え、1つまたは複数の仮想オブジェクトを提示することは、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従って1つまたは複数の仮想オブジェクトを提示することを備える、条項17に記載の方法。

【0180】

【0183】条項25：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項17に記載の方法。

10

【0181】

【0184】条項26：メディアデータを提示するためのデバイスであって、仮想シーン内の1つまたは複数の仮想オブジェクトを定義するメディアデータを記憶するように構成されたメモリと、回路内で実装された1つまたは複数のプロセッサとを備え、1つまたは複数のプロセッサは、

ビットストリームのシーン記述を受信することと、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すものであり；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定することと；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示することと、を行うように構成された、デバイス。

20

【0182】

【0185】条項27：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項26に記載のデバイス。

【0183】

【0186】条項28：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項26に記載のデバイス。

30

【0184】

【0187】条項29：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項26に記載のデバイス。

【0185】

【0188】条項30：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項29に記載のデバイス。

【0186】

【0189】条項31：1つまたは複数の仮想オブジェクトを提示するために、1つまたは複数のプロセッサは、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定し、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換するように構成される、条項29に記載のデバイス。

40

【0187】

【0190】条項32：1つまたは複数の仮想オブジェクトを提示するために、1つまたは複数のプロセッサは、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせることを行うように構成される、条項26に記載のデバイス。

【0188】

【0191】条項33：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、1つまたは複数のプロセッサは、アクションに応答して1つまた

50

は複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定するようにさらに構成され、1つまたは複数のプロセッサは、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従って1つまたは複数の仮想オブジェクトを提示するように構成される、条項26に記載のデバイス。

【0189】

[0192]条項34：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項26に記載のデバイス。

【0190】

[0193]条項35：デバイスは、集積回路、マイクロプロセッサ、およびワイヤレス通信デバイスのうちの少なくとも1つを備える、条項26に記載のデバイス。

【0191】

[0194]条項36：実行されたとき、プロセッサに、ビットストリームのシーン記述を受信することと、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すものであり；シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定することと；決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示することと、を行わせる命令を記憶したコンピュータ可読記憶媒体。

【0192】

[0195]条項37：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項36に記載のコンピュータ可読記憶媒体。

【0193】

[0196]条項38：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項36に記載のコンピュータ可読記憶媒体。

【0194】

[0197]条項39；シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項36に記載のコンピュータ可読記憶媒体。

【0195】

[0198]条項40：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項39に記載のコンピュータ可読記憶媒体。

【0196】

[0199]条項41：プロセッサに1つまたは複数の仮想オブジェクトを提示させる命令は、プロセッサに、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定させ、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換させる命令を備える、条項39に記載のコンピュータ可読記憶媒体。

【0197】

[0200]条項42；プロセッサに1つまたは複数の仮想オブジェクトを提示させる命令は、プロセッサに、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせること、を行わせる命令を備える、条項36に記載のコンピュータ可読記憶媒体。

【0198】

[0201]条項43：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、プロセッサに、アクションに応答して1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定さ

せる命令をさらに備え、プロセッサに1つまたは複数の仮想オブジェクトを提示させる命令は、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従ってプロセッサに1つまたは複数の仮想オブジェクトを提示させる命令を備える、条項36に記載のコンピュータ可読記憶媒体。

【0199】

[0202]条項44：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項36に記載のコンピュータ可読記憶媒体。

【0200】

[0203]条項45：メディアデータを提示するためのデバイスであって、ビットストリームのシーン記述を受信するための手段と、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定するための手段と、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示するための手段とを備えるデバイス。

【0201】

[0204]条項46：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項45に記載のデバイス。

【0202】

[0205]条項47：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項45に記載のデバイス。

【0203】

[0206]条項48：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項45に記載のデバイス。

【0204】

[0207]条項49：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項48に記載のデバイス。

【0205】

[0208]条項50：1つまたは複数の仮想オブジェクトを提示するための手段は、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定するための手段と、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換するための手段とを備える、条項48に記載のデバイス。

【0206】

[0209]条項51：1つまたは複数の仮想オブジェクトを提示するための手段は、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせるための手段を備える、条項45に記載のデバイス。

【0207】

[0210]条項52：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、アクションに応答して1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定するための手段をさらに備え、1つまたは複数の仮想オブジェクトを提示するための手段は、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従って1つまたは複数の仮想オブジェクトを提示するための手段を備える、条項45に記載のデバイス。

【0208】

10

20

30

40

50

【0211】条項 5 3：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項 4 5 に記載のデバイス。

【0209】

【0212】条項 5 4：メディアデータを提示する方法であって、ビットストリームのシーン記述を受信することと、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定することと、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示することとを備える方法。

10

【0210】

【0213】条項 5 5：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項 5 4 に記載の方法。

【0211】

【0214】条項 5 6：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することとをさらに備える、条項 5 4 または 5 5 に記載の方法。

【0212】

20

【0215】条項 5 7：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項 5 4 から 5 6 のいずれかに記載の方法。

【0213】

【0216】条項 5 8：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項 5 7 に記載の方法。

【0214】

【0217】条項 5 9：1つまたは複数の仮想オブジェクトを提示することは、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定することと、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換することとを備える、条項 5 7 または 5 8 に記載の方法。

30

【0215】

【0218】条項 6 0：1つまたは複数の仮想オブジェクトを提示することは、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせることを備える、条項 5 4 から 5 9 のいずれかに記載の方法。

【0216】

【0219】条項 6 1：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、方法は、アクションに応答して1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定することとをさらに備え、1つまたは複数の仮想オブジェクトを提示することは、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従って1つまたは複数の仮想オブジェクトを提示することとを備える、条項 5 4 から 6 0 のいずれかに記載の方法。

40

【0217】

【0220】条項 6 2：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項 5 4 から 6 1 のいずれかに記載の方法。

【0218】

【0221】条項 6 3：メディアデータを提示するためのデバイスであって、仮想シーン内の1つまたは複数の仮想オブジェクトを定義するメディアデータを記憶するように構成さ

50

れたメモリと、回路内で実装され、ビットストリームのシーン記述を受信し、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すものであり、シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定し、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示するように構成された1つまたは複数のプロセッサとを備えるデバイス。

【0219】

【0222】条項64：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項63に記載のデバイス。

10

【0220】

【0223】条項65：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項63または64に記載のデバイス。

【0221】

【0224】条項66：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項63から65のいずれかに記載のデバイス。

20

【0222】

【0225】条項67：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項66に記載のデバイス。

【0223】

【0226】条項68：1つまたは複数の仮想オブジェクトを提示するために、1つまたは複数のプロセッサは、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定し、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換するように構成される、条項66または67に記載のデバイス。

【0224】

30

【0227】条項69：1つまたは複数の仮想オブジェクトを提示するために、1つまたは複数のプロセッサは、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせること、を行わせるように構成される、条項63から68のいずれかに記載のデバイス。

【0225】

【0228】条項70：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、1つまたは複数のプロセッサは、アクションに応答して1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定するようにさらに構成され、1つまたは複数のプロセッサは、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従って1つまたは複数の仮想オブジェクトを提示するように構成される、条項63から69のいずれかに記載のデバイス。

40

【0226】

【0229】条項71：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項63から70のいずれかに記載のデバイス。

【0227】

【0230】条項72：デバイスは、集積回路、マイクロプロセッサ、およびワイヤレス通信デバイスのうちの少なくとも1つを備える、条項63から71のいずれかに記載のデバイス。

50

【0228】

【0231】条項73：実行されたとき、プロセッサに、ビットストリームのシーン記述を受信させ、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表すものであり、シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定させ、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示させる命令を記憶したコンピュータ可読記憶媒体。

【0229】

【0232】条項74：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項73に記載のコンピュータ可読記憶媒体。

10

【0230】

【0233】条項75：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項73または74に記載のコンピュータ可読記憶媒体。

【0231】

【0234】条項76：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項73から75のいずれかに記載のコンピュータ可読記憶媒体。

20

【0232】

【0235】条項77：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項76に記載のコンピュータ可読記憶媒体。

【0233】

【0236】条項78：プロセッサに1つまたは複数の仮想オブジェクトを提示させる命令は、プロセッサに、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定させ、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換させる命令を備える、条項76または77に記載のコンピュータ可読記憶媒体。

30

【0234】

【0237】条項79：プロセッサに1つまたは複数の仮想オブジェクトを提示させる命令は、プロセッサに、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせることを行わせる命令を備える、条項73から78のいずれかに記載のコンピュータ可読記憶媒体。

【0235】

【0238】条項80：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、プロセッサに、アクションに応答して1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定させる命令をさらに備え、プロセッサに1つまたは複数の仮想オブジェクトを提示させる命令は、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従ってプロセッサに1つまたは複数の仮想オブジェクトを提示させる命令を備える、条項73から79のいずれかに記載のコンピュータ可読記憶媒体。

40

【0236】

【0239】条項81：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項73から80のいずれかに記載のコンピュータ可読記憶媒体。

【0237】

【0240】条項82：メディアデータを提示するためのデバイスであって、ビットストリ

50

ームのシーン記述を受信するための手段と、シーン記述は、シーンアンカーと仮想シーン内の1つまたは複数の仮想オブジェクトとを記述するデータを含み、シーンアンカーは、仮想シーンと現実世界のプレゼンテーション環境との間の対応を表し、シーンアンカーを使用して仮想シーンと現実世界のプレゼンテーション環境との間の対応を決定するための手段と、決定された対応に従って現実世界のプレゼンテーション環境内のロケーションにおいて1つまたは複数の仮想オブジェクトを提示するための手段とを備えるデバイス。

【0238】

【0241】条項83：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、ローカル、ビュー、またはステージのうちの1つまたは複数である、条項82に記載のデバイス。

10

【0239】

【0242】条項84：シーンアンカーは、現実世界のプレゼンテーション環境についてのタイプを定義するデータを含み、タイプは、アプリケーション定義空間であり、方法は、アプリケーション定義空間を定義するデータを受信することをさらに備える、条項82または83に記載のデバイス。

【0240】

【0243】条項85：シーン記述は、シーン座標系から現実世界のプレゼンテーション環境の座標系への変換を表すデータを含む、条項82から84のいずれかに記載のデバイス。

【0241】

20

【0244】条項86：変換は、回転、平行移動、または拡大縮小のうちの1つまたは複数を含む、条項85に記載のデバイス。

【0242】

【0245】条項87：1つまたは複数の仮想オブジェクトを提示するための手段は、仮想シーン内の1つまたは複数の仮想オブジェクトについてのシグナリングされたロケーションを決定するための手段と、現実世界のプレゼンテーション環境内のロケーションを生成するために変換に従ってシグナリングされたロケーションを現実世界のロケーションに変換するための手段とを備える、請求項85または86に記載のデバイス。

【0243】

【0246】条項88：1つまたは複数の仮想オブジェクトを提示するための手段は、シーンアンカーに従って仮想シーンを現実世界のプレゼンテーション環境にアライメントさせるための手段を備える、条項82から87のいずれかに記載のデバイス。

30

【0244】

【0247】条項89：シーン記述は、ユーザから受信した入力を表すアクションを定義するデータをさらに含み、アクションに応答して1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションに対する修正を決定するための手段をさらに備え、1つまたは複数の仮想オブジェクトを提示するための手段は、1つまたは複数の仮想オブジェクトまたは仮想カメラのうちの少なくとも1つのロケーションの修正に従って1つまたは複数の仮想オブジェクトを提示するための手段を備える、条項82から88のいずれかに記載のデバイス。

40

【0245】

【0248】条項90：シーン記述は、1つまたは複数の仮想オブジェクトについてのオブジェクト識別子を定義するデータをさらに含む、条項82から89のいずれかに記載のデバイス。

【0246】

【0249】1つまたは複数の例では、説明された機能は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組合せにおいて実装され得る。ソフトウェアで実装される場合、機能は、1つまたは複数の命令またはコードとして、コンピュータ可読媒体上に記憶されるか、あるいはコンピュータ可読媒体を介して送信され、ハードウェアベースの処理ユニットによって実行され得る。コンピュータ可読媒体は、データ記憶媒体な

50

どの有形媒体に対応する、コンピュータ可読記憶媒体を含み得るか、または、たとえば、通信プロトコルに従って、ある場所から別の場所へのコンピュータプログラムの転送を促進する任意の媒体を含む通信媒体を含み得る。このようにして、コンピュータ可読媒体は、概して、(1) 非一時的である有形コンピュータ可読記憶媒体、または(2) 信号もしくは搬送波などの通信媒体に、相当し得る。データ記憶媒体は、本開示で説明される技法の実装のための命令、コードおよび／またはデータ構造を取り出すために、1つまたは複数のコンピュータあるいは1つまたは複数のプロセッサによってアクセスされ得る、任意の利用可能な媒体であり得る。コンピュータプログラム製品は、コンピュータ可読媒体を含んでもよい。

【0247】

10

【0250】限定ではなく例として、そのようなコンピュータ可読記憶媒体は、RAM、ROM、EEPROM（登録商標）、CD-ROMもしくは他の光ディスクストレージ、磁気ディスクストレージ、もしくは他の磁気ストレージデバイス、フラッシュメモリ、または、命令もしくはデータ構造の形態の所望のプログラムコードを記憶するために使用され得、コンピュータによってアクセスされ得る任意の他の媒体を備えることができる。さらに、いかなる接続もコンピュータ可読媒体と適切に呼ばれる。たとえば、命令が、同軸ケーブル、光ファイバケーブル、ツイストペア、デジタル加入者回線(DSL)、または赤外線、無線、およびマイクロ波などのワイヤレス技術を使用して、ウェブサイト、サーバ、または他のリモートソースから送信される場合、同軸ケーブル、光ファイバケーブル、ツイストペア、DSL、または赤外線、無線、およびマイクロ波などのワイヤレス技術は媒体の定義に含まれる。しかしながら、コンピュータ可読記憶媒体およびデータ記憶媒体が、接続、搬送波、信号、または他の一時的媒体を含むのではなく、代わりに非一時的な有形の記憶媒体を対象とすることを理解されたい。本明細書で使用されるディスク(disk)およびディスク(disc)は、コンパクトディスク(disc)(CD)、レーザーディスク（登録商標）(disc)、光ディスク(disc)、デジタル多用途ディスク(disc)(DVD)、フロッピーディスク(disk)、およびBlu-ray（登録商標）ディスク(disc)を含み、ディスク(disk)は、通常、データを磁氣的に再生し、ディスク(disc)は、データをレーザーで光学的に再生する。上記の組合せも、コンピュータ可読媒体の範囲内に含まれるべきである。

20

【0248】

30

【0251】命令は、1つまたは複数のデジタル信号プロセッサ(DSP)などの1つまたは複数のプロセッサ、汎用マイクロプロセッサ、特定用途向け集積回路(ASIC)、フィールドプログラマブル論理アレイ(FPGA)、あるいは他の等価な集積回路またはディスクリット論理回路によって実行され得る。したがって、本明細書で使用される「プロセッサ」という用語は、上記の構造、または、本明細書で説明された技法の実装に好適な任意の他の構造のいずれかを指すことがある。さらに、いくつかの態様では、本明細書で説明された機能は、符号化および復号のために構成された専用ハードウェアおよび／またはソフトウェアモジュール内に提供されるか、あるいは複合コーデックに組み込まれ得る。また、本技法は、1つまたは複数の回路または論理要素で十分に実装され得る。

【0249】

40

【0252】本開示の技法は、ワイヤレスハンドセット、集積回路(IC)またはICのセット(たとえば、チップセット)を含む、多種多様なデバイスまたは装置で実装され得る。本開示では、開示される技法を実行するように構成されたデバイスの機能的態様を強調するために様々な構成要素、モジュール、またはユニットについて説明したが、それらの構成要素、モジュール、またはユニットを、必ずしも異なるハードウェアユニットによって実現する必要があるとは限らない。むしろ、上記で説明したように、様々なユニットが、好適なソフトウェアおよび／またはファームウェアとともに、上記で説明した1つまたは複数のプロセッサを含めて、コーデックハードウェアユニットにおいて組み合わせられ得るか、または相互動作可能なハードウェアユニットの集合によって与えられ得る。

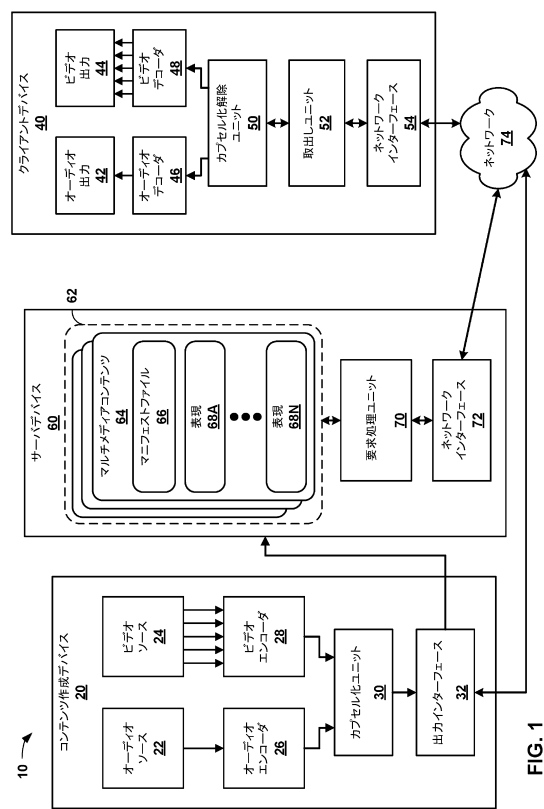
【0250】

50

[0253]様々な例について説明してきた。これらおよび他の例は、以下の特許請求の範囲内に入る。

【図面】

【図 1】



【図 2】

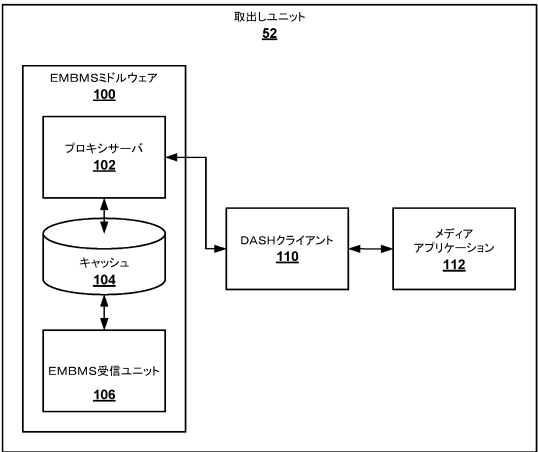


FIG. 2

【 図 3 】

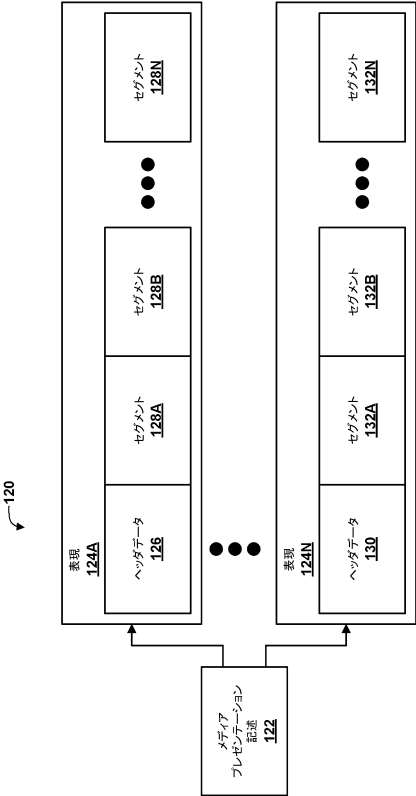


FIG. 3

【 図 4 】

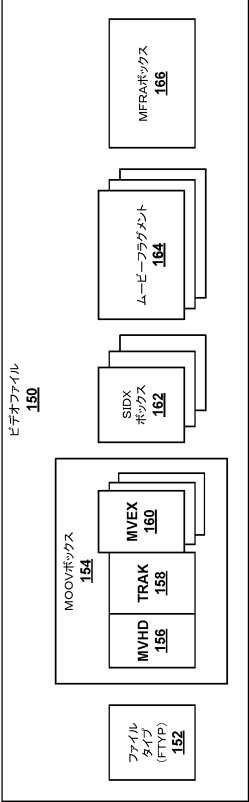


FIG. 4

【 図 5 】

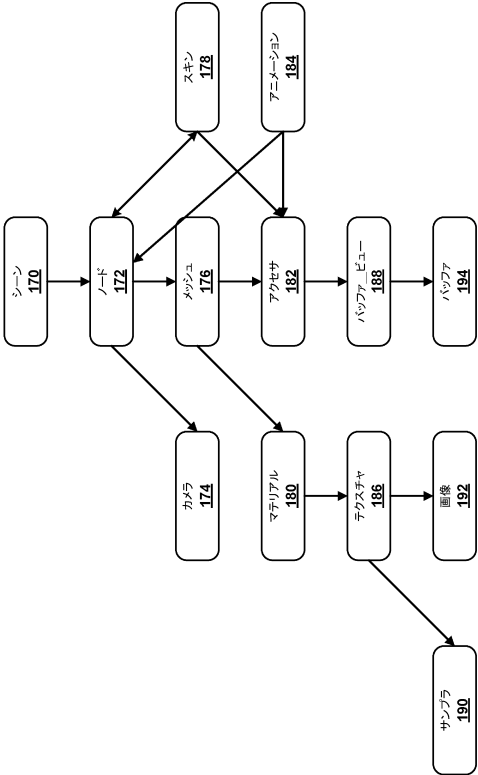


FIG. 5

【 図 6 】

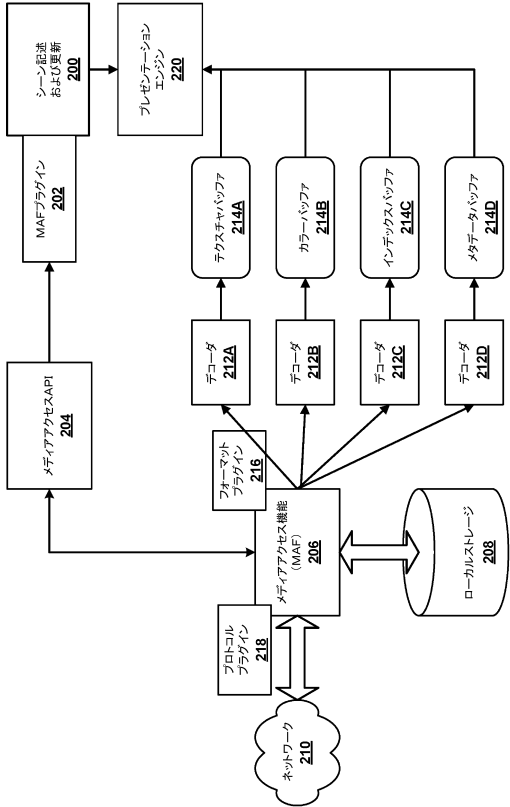


FIG. 6

【図 7】

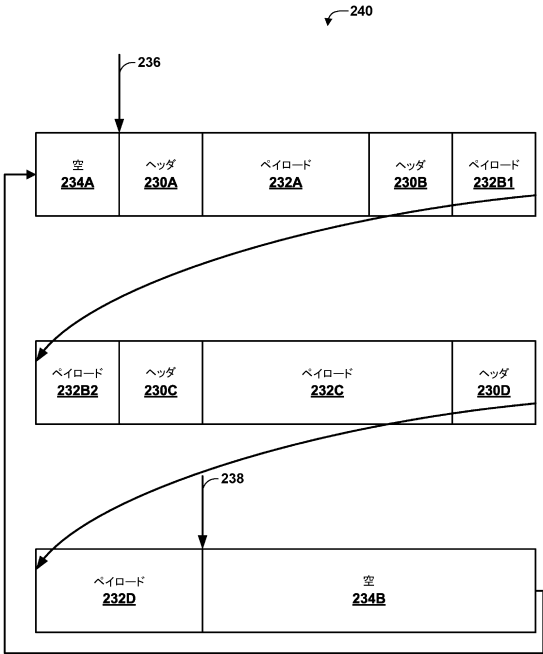


FIG. 7

【図 9】

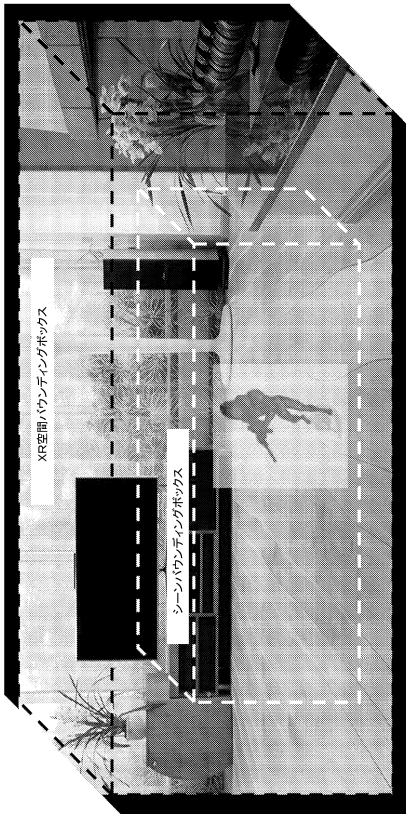


FIG. 9

【図 8】

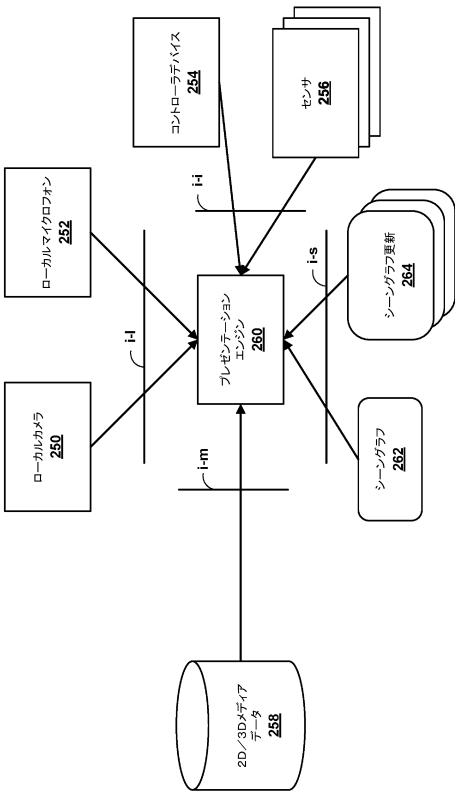


FIG. 8

【図 10】

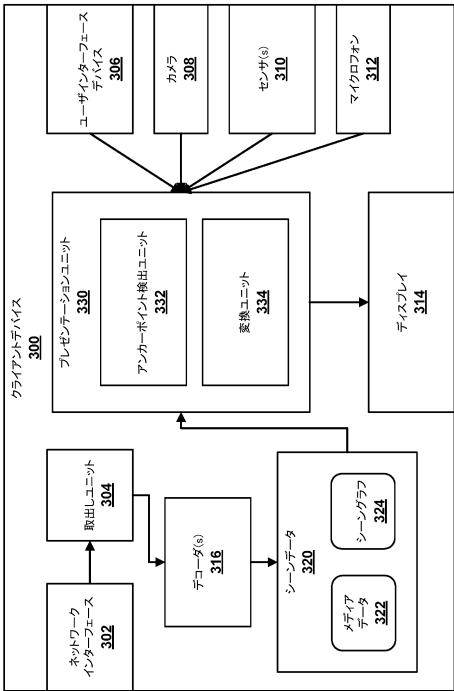


FIG. 10

【図 11】

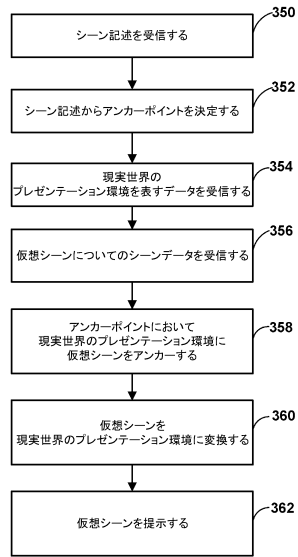


FIG. 11

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/071809

A. CLASSIFICATION OF SUBJECT MATTER INV. H04N21/44 H04N21/81 H04N21/854 H04N21/643 H04N21/845 H04N21/84 ADD. According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) H04N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	YU YOU (NOKIA) ET AL: "Updates to MPEG & 5G (XR use cases and requirements)", 134. MPEG MEETING; 20210426 - 20210430; ONLINE; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. m56667 18 April 2021 (2021-04-18), XP030295190, Retrieved from the Internet: URL:https://dms.mpeg.expert/doc_end_user/documents/134_OnLine/wg11/m56667-v1-Updates toMPEG&5G(XR)usecasesandrequirements.docx. zip Updates to MPEG & 5G (XR) use cases and requirements.docx [retrieved on 2021-04-18] page 1, paragraph 2 - page 2, paragraph 2 page 2, paragraph 3.1.1 - page 4, paragraph 3.1.2 page 5, paragraph 4.2 -/--	1-37
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 27 June 2022		Date of mailing of the international search report 06/07/2022
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Fantini, Federico

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2022/071809

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	figure 2 ----- US 2019/287306 A1 (WIESER ANTHONY ARNOLD [GB] ET AL) 19 September 2019 (2019-09-19) abstract paragraph [0033] - paragraph [0040] paragraph [0042] - paragraph [0043] paragraph [0050] - paragraph [0051] figures 5A, 5B -----	1-37
A	EP 3 588 249 A1 (KONINKLIJKE PHILIPS NV [NL]) 1 January 2020 (2020-01-01) abstract paragraph [0005] - paragraph [0006] paragraph [0057] - paragraph [0058] paragraph [0082] - paragraph [0085] paragraph [0088] - paragraph [0089] paragraph [0111] - paragraph [0112] -----	1-37
A	QUALCOMM INCORPORATED: "Proposed Updates to Baseline Technologies", 3GPP DRAFT; S4-200042, 3RD GENERATION PARTNERSHIP PROJECT (3GPP), MOBILE COMPETENCE CENTRE ; 650, ROUTE DES LUCIOLES ; F-06921 SOPHIA-ANTIPOLIS CEDEX ; FRANCE / vol. SA WG4, no. Wroclaw, Poland; 20200120 - 20200124 14 January 2020 (2020-01-14), XP051843521, Retrieved from the Internet: URL:https://ftp.3gpp.org/tsg_sa/WG4_CODECS/TSGS4_107_Wroclaw/Docs/S4-200042.zip S4-200042 - Updates to TR 26.928 - based on Telco Agreements.docx [retrieved on 2020-01-14] page 6, paragraph 4.1.1 page 8, paragraph 4.1.2 - page 11, paragraph 4.1.5 page 14, paragraph 4.3.1 page 21, paragraph 4.4.1 page 32, paragraph 4.6.6 page 37, paragraph 4.9.2.3 figures 4.1-5, 4.4.1-2 -----	1-37
A	EP 3 734 970 A1 (SAMSUNG ELECTRONICS CO LTD [KR]) 4 November 2020 (2020-11-04) abstract paragraph [0079] - paragraph [0091] paragraph [0165] - paragraph [0168] table 2 figure 5 ----- -/--	1-37

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2022/071809

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2018/276890 A1 (WANG YEKUI [US]) 27 September 2018 (2018-09-27) abstract paragraph [0035] - paragraph [0042] figure 11 -----	1-37
A	"Augmented Reality Framework (ARF) AR framework architecture", ETSI DRAFT SPECIFICATION; ARF 003, EUROPEAN TELECOMMUNICATIONS STANDARDS INSTITUTE (ETSI), 650, ROUTE DES LUCIOLES ; F-06921 SOPHIA-ANTIPOLIS ; FRANCE , vol. ISG ARF Augmented Reality Framework, no. V0.0.31 21 January 2020 (2020-01-21), pages 1-49, XP014362810, Retrieved from the Internet: URL:docbox.etsi.org/ISG/ARF/70-Draft/003/A RF-003v0031.docx [retrieved on 2020-01-21] page 9, paragraph 3.1 - page 12, paragraph 4.2 page 15, paragraph 5.2.5 page 16, paragraph 5.3.2 page 19, paragraph 5.6.4 page 20, paragraph 5.8 - paragraph 5.8.2 page 28, paragraph 6.14 page 29, paragraph 6.17 figure 2 -----	1-37
A	US 2021/099773 A1 (BOUAZIZI IMED [US] ET AL) 1 April 2021 (2021-04-01) abstract paragraph [0019] - paragraph [0022] paragraph [0037] - paragraph [0040] paragraph [0143] - paragraph [0151] paragraph [0156] - paragraph [0157] paragraph [0167] - paragraph [0168] figure 8 -----	1-37

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2022/071809

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 2019287306 A1		19-09-2019	CN 112243583 A	19-01-2021
			EP 3769509 A1	27-01-2021
			US 2019287306 A1	19-09-2019
			WO 2019182772 A1	26-09-2019

EP 3588249 A1		01-01-2020	BR 112021000289 A2	06-04-2021
			CA 3105400 A1	02-01-2020
			CN 112602042 A	02-04-2021
			EP 3588249 A1	01-01-2020
			EP 3811185 A1	28-04-2021
			JP 2022501685 A	06-01-2022
			KR 20210024071 A	04-03-2021
			TW 202016692 A	01-05-2020
			US 2021264658 A1	26-08-2021
			WO 2020002115 A1	02-01-2020

EP 3734970 A1		04-11-2020	CN 111869201 A	30-10-2020
			EP 3734970 A1	04-11-2020
			US 2021235058 A1	29-07-2021
			WO 2020145668 A1	16-07-2020

US 2018276890 A1		27-09-2018	AU 2018237595 A1	29-08-2019
			BR 112019019287 A2	14-04-2020
			CN 110431522 A	08-11-2019
			EP 3602261 A1	05-02-2020
			KR 20190131062 A	25-11-2019
			SG 11201907476X A	30-10-2019
			TW 201840201 A	01-11-2018
			US 2018276890 A1	27-09-2018
			WO 2018175903 A1	27-09-2018

US 2021099773 A1		01-04-2021	CN 114503599 A	13-05-2022
			TW 202127899 A	16-07-2021
			US 2021099773 A1	01-04-2021
			WO 2021067593 A1	08-04-2021

フロントページの続き

,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,D
K,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),O
A(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,B
B,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD
GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,IT,JM,JO,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,
LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO
RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,Z
M,ZW

2121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

Fターム (参考) 5B050 AA10 BA08 BA09 BA11 BA20 CA07 CA08 DA05 EA05 EA10
EA12 EA13 EA26 FA02 FA05 FA10
5C164 MA03S MB13P MB43S UB81P