

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6735415号
(P6735415)

(45) 発行日 令和2年8月5日(2020.8.5)

(24) 登録日 令和2年7月15日(2020.7.15)

(51) Int.Cl. F I
 HO4N 21/431 (2011.01) HO4N 21/431
 HO4N 21/4728 (2011.01) HO4N 21/4728

請求項の数 11 (全 56 頁)

(21) 出願番号	特願2019-513366 (P2019-513366)	(73) 特許権者	315002955
(86) (22) 出願日	平成29年9月6日(2017.9.6)		ノキア テクノロジーズ オーユー
(65) 公表番号	特表2019-526994 (P2019-526994A)		フィンランド共和国 02610 エスポー
(43) 公表日	令和1年9月19日(2019.9.19)		カラカーリ 7
(86) 国際出願番号	PCT/FI2017/050626	(74) 代理人	100127188
(87) 国際公開番号	W02018/046797		弁理士 川守田 光紀
(87) 国際公開日	平成30年3月15日(2018.3.15)	(72) 発明者	ハンヌクセラ ミスカ
審査請求日	平成31年4月15日(2019.4.15)		フィンランド共和国 33610 タンペ
(31) 優先権主張番号	15/261, 192		レルストホッリンリンネ 2
(32) 優先日	平成28年9月9日(2016.9.9)	(72) 発明者	アフラキベニ ペイマン
(33) 優先権主張国・地域又は機関	米国 (US)		フィンランド共和国 33720 タンペ
			レ アトミカツ 1 F 23
		審査官	松元 伸次
			最終頁に続く

(54) 【発明の名称】 オーディオビジュアルコンテンツの観察点および観察向きの制御された選択のための方法および装置

(57) 【特許請求の範囲】

【請求項 1】

再生デバイスの意図された動作を特定することと；
 前記再生デバイスの前記意図された動作を特定することに応じて制御信号を生成させることと；

を含み、

前記意図された動作は観察設定を含み、前記観察設定は観察点と観察向きのいずれかまたは両方を含み、前記観察点は、オーディオプレゼンテーション及び/又はビデオプレゼンテーションの取得又は再生のための、3次元空間内の点または体積を示し、前記観察向きは、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションがレンダリングされる向きを示し、

前記制御信号は、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションの前記再生デバイス上でのレンダリング動作に関連付けられており、前記制御信号は、前記再生デバイスの前記意図された動作を示し、

ビデオストリーム中の再生モード条件に関連付けられる前記観察設定は、前記再生デバイスによって前記意図された動作がいつ適用されるかを示す標示を含み、前記標示は再生モードを示し、前記再生モードは、連続再生と、連続再生開始のためのシーク又はランダムアクセスとの一つ又は両方を含む、

方法。

【請求項 2】

前記レンダリング動作は、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションの部分を選択することを含む、請求項1に記載の方法。

【請求項3】

前記観察設定は、前記再生デバイスの意図された動作に関連付けられた標示であって、最も取りうる視聴方向を示す標示を受信することを含む、請求項1又は2に記載の方法。

【請求項4】

前記観察設定に関連付けられた条件が満たされたかどうかを判断することは、前記再生デバイスがランダムアクセスモードであるかどうかを判断することを含む、請求項1から3のいずれかに記載の方法。

【請求項5】

制御信号を受信することと；

前記制御信号に従ってオーディオプレゼンテーション及び/又はビデオプレゼンテーションをレンダリングすることと；

を含む方法であって、

前記制御信号は、オーディオプレゼンテーション及び/又はビデオプレゼンテーションの再生デバイス上でのレンダリング動作に関連付けられており、

前記制御信号は、前記制御信号は、前記再生デバイスの意図された動作を示し、

前記意図された動作は観察設定を含み、前記観察設定は観察点と観察向きのいずれかまたは両方を含み、

前記観察点は、オーディオプレゼンテーション及び/又はビデオプレゼンテーションの取得又は再生のための、3次元空間内の点または体積を示し、前記観察向きは、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションがレンダリングされる向きを示し、

ビデオストリーム中の再生モード条件に関連付けられる前記観察設定は、前記再生デバイスによって前記意図された動作がいつ適用されるかを示す標示を含み、前記標示は再生モードを示し、前記再生モードは、連続再生と、連続再生開始のためのシーク又はランダムアクセスとの一つ又は両方を含み、

前記方法は更に、前記標示を適用するかどうかを決定することを含む、方法。

【請求項6】

前記制御信号に従って、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションの伝送単位を要求することを更に含む、請求項5に記載の方法。

【請求項7】

連続再生において無条件に適用される初期観察設定標示に基づいて後で要求される可能性がある伝送単位を判断することと；

前記後で要求される可能性がある前記伝送単位の情報を送信することと；

を更に含む、請求項5又は6に記載の方法。

【請求項8】

再生デバイスの意図された動作を特定する手段と；

前記再生デバイスの前記意図された動作を特定することに応じて制御信号を生成させる手段と；

を備え、

前記意図された動作は観察設定を含み、前記観察設定は観察点と観察向きのいずれかまたは両方を含み、前記観察点は、オーディオプレゼンテーション及び/又はビデオプレゼンテーションの取得又は再生のための、3次元空間内の点または体積を示し、前記観察向きは、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションがレンダリングされる向きを示し、

前記制御信号は、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションの前記再生デバイス上でのレンダリング動作に関連付けられており、前記制御信号は、前記再生デバイスの前記意図された動作を示し、

ビデオストリーム中の再生モード条件に関連付けられる前記観察設定は、前記再生デバ

10

20

30

40

50

イスによって前記意図された動作がいつ適用されるかを示す標示を含み、前記標示は再生モードを示し、前記再生モードは、連続再生と、連続再生開始のためのシーク又はランダムアクセスとの一つ又は両方を含む、装置。

【請求項 9】

制御信号を受信する手段と；

前記制御信号に従ってオーディオプレゼンテーション及び/又はビデオプレゼンテーションをレンダリングする手段と；
を備え、

前記制御信号は、オーディオプレゼンテーション及び/又はビデオプレゼンテーションの再生デバイス上でのレンダリング動作に関連付けられており、

前記制御信号は、前記制御信号は、前記再生デバイスの意図された動作を示し、

前記意図された動作は観察設定を含み、前記観察設定は観察点と観察向きのいずれかまたは両方を含み、

前記観察点は、オーディオプレゼンテーション及び/又はビデオプレゼンテーションの取得又は再生のための、3次元空間内の点または体積を示し、前記観察向きは、前記オーディオプレゼンテーション及び/又は前記ビデオプレゼンテーションがレンダリングされる向きを示し、

ビデオストリーム中の再生モード条件に関連付けられる前記観察設定は、前記再生デバイスによって前記意図された動作がいつ適用されるかを示す標示を含み、前記標示は再生モードを示し、前記再生モードは、連続再生と、連続再生開始のためのシーク又はランダムアクセスとの一つ又は両方を含む、装置。

【請求項 10】

処理手段及び記憶手段を備える装置であって、前記記憶手段はプログラム命令を格納し、前記プログラム命令は、前記処理手段に実行されると、前記装置に、請求項 1 から 7 のいずれかに記載の方法を遂行させるように構成される、装置。

【請求項 11】

装置の処理手段に実行されると、前記装置に、請求項 1 から 7 のいずれかに記載の方法を遂行させるように構成されるプログラム命令を備える、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

一例示的实施形態は概してメディアコンテンツの符号化に関し、特に、ストリーミング仮想現実コンテンツおよびその他のオーディオビジュアルコンテンツの状況において関連する。

【背景】

【0002】

360°ビデオや他の同様のコンテンツなどの仮想現実オーディオビジュアルコンテンツは、そのようなコンテンツから得られる没入的視聴体験を好んだり楽しんだりする視聴者やコンテンツクリエイタの間でますます人気を集めつつある。このような仮想現実コンテンツの人気の高まりにより、高品質の視聴体験をもたらすストリーミング仮想現実コンテンツに対する視聴者の要求が高まっている。

【0003】

ストリーミング仮想現実コンテンツ環境において一貫した高品質の視聴体験を提供するニーズにより、様々な技術的課題がもたらされる。これは特に、コンテンツクリエイタが、コンテンツをどのように視聴者に提示すべきか、およびそのようなコンテンツを視聴者の視野内でどのように提示すべきかに関する創作的および/または監督的な決定をした場合に当てはまる。このような技術的課題は、視聴者にとって快適な視聴の向きの範囲が限られている場合や、視聴者の向きによって、視聴者が体験するコンテンツが、コンテンツ

10

20

30

40

50

クリエイタが意図したものから離れてしまいがちな場合に伴うことがある。

【摘要】

【0004】

したがって、ユーザ主導のオーディオビジュアルコンテンツの選択的レンダリングを行うために、一例示的实施形態による方法、装置、およびコンピュータプログラム製品を提供する。この点において、一例示的实施形態による方法、装置、およびコンピュータプログラム製品は、レンダリングされるオーディオビジュアルコンテンツの観察点および観察向きの選択の制御を行う。

【0005】

一例示的实施形態において方法が提供される。この方法は、オーディオビジュアルプレゼンテーションの伝送単位のセットにおける初期観察設定に関連付けられた標示を受信することを含む。本例示的实施形態の前記方法は、再生デバイスの意図された動作に関連付けられた標示を受信することを含む。本例示的实施形態の前記方法は、前記再生デバイスの前記意図された動作を特定することを含む。本例示的实施形態の前記方法は、前記再生デバイスの前記意図された動作を特定することに応じて制御信号を生成させることも含み、前記制御信号は、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。

10

【0006】

そのような方法のいくつかの例示的実装において、前記観察設定は観察点と観察向きとを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記再生デバイスの前記意図された動作に関連付けられた前記標示は、前記再生デバイスの連続再生モードにおける前記再生デバイスの意図された動作に関連付けられた標示と、前記再生デバイスのランダムアクセスモードにおける前記再生デバイスの意図された動作に関連付けられた標示とを含む。

20

【0007】

いくつかの例示的実装において、前記再生デバイスの前記意図された動作を特定することは、前記再生デバイスの前記意図された動作に関連付けられた条件が満たされたかどうかを判断することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記条件は、前記初期観察設定に関連付けられた少なくとも1つのリセット条件を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記

30

【0008】

いくつかの例示的実装において、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションの前記レンダリング動作は、前記オーディオビジュアルプレゼンテーションの部分を選択することを含む。

【0009】

別の例示的实施形態において、装置が提供される。前記装置は、少なくとも1つのプロセッサと、コンピュータプログラムコードを含む少なくとも1つのメモリと、を備え、前記少なくとも1つのメモリおよび前記コンピュータプログラムコードは、前記少なくとも1つのプロセッサによって、少なくとも、オーディオビジュアルプレゼンテーションの伝送単位のセットにおける初期観察設定に関連付けられた標示を受信することと、再生デバイスの意図された動作に関連付けられた標示を受信することと、前記再生デバイスの前記意図された動作を特定することと、前記再生デバイスの前記意図された動作を特定することに応じて制御信号を生成させることと、を前記装置にさせるように構成され、前記制御信号は、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。

40

【0010】

いくつかの例示的実装において、前記観察設定は観察点と観察向きとを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記再生デバイスの前記意図された動作に関連付けられた前記標示は、前記再生デバイスの連続再生モードにお

50

る前記再生デバイスの意図された動作に関連付けられた標示と、前記再生デバイスのランダムアクセスモードにおける前記再生デバイスの意図された動作に関連付けられた標示とを含む。

【0011】

いくつかの例示的実装において、前記再生デバイスの前記意図された動作を特定することは、前記再生デバイスの前記意図された動作に関連付けられた条件が満たされたかどうかを判断することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記条件は、前記初期観察設定に関連付けられた少なくとも1つのリセット条件を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記条件は、前記初期観察設定に関連付けられた少なくとも1つの維持条件を含む。

10

【0012】

いくつかの例示的実装において、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションの前記レンダリング動作は、前記オーディオビジュアルプレゼンテーションの部分を選択することを含む。

【0013】

さらなる例示的実施形態において、コンピュータプログラム製品が提供される。このコンピュータプログラム製品は、コンピュータによって実行可能なプログラムコード命令が記憶された、少なくとも1つの非一時的コンピュータ可読記憶媒体を含み、前記コンピュータによって実行可能なプログラムコード命令は、オーディオビジュアルプレゼンテーションの伝送単位のセットにおける初期観察設定に関連付けられた標示を受信し、再生デバイスの意図された動作に関連付けられた標示を受信し、前記再生デバイスの前記意図された動作を特定し、前記再生デバイスの前記意図された動作を特定することに応じて制御信号を生成させる、ように構成されたプログラムコード命令を含み、前記制御信号は、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられ、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションの前記レンダリング動作は、前記オーディオビジュアルプレゼンテーションの部分を選択することを含む。

20

【0014】

いくつかの例示的実装において、前記観察設定は観察点と観察向きとを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記再生デバイスの前記意図された動作に関連付けられた前記標示は、前記再生デバイスの連続再生モードにおける前記再生デバイスの意図された動作に関連付けられた標示と、前記再生デバイスのランダムアクセスモードにおける前記再生デバイスの意図された動作に関連付けられた標示とを含む。

30

【0015】

いくつかの例示的実装において、前記再生デバイスの前記意図された動作を特定することは、前記再生デバイスの前記意図された動作に関連付けられた条件が満たされたかどうかを判断することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記条件は、前記初期観察設定に関連付けられた少なくとも1つのリセット条件を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記条件は、前記初期観察設定に関連付けられた少なくとも1つの維持条件を含む。

40

【0016】

さらに別の例示的実施形態において、装置が提供される。前記装置は、オーディオビジュアルプレゼンテーションの伝送単位のセットにおける初期観察設定に関連付けられた標示を受信し、再生デバイスの意図された動作に関連付けられた標示を受信し、前記再生デバイスの前記意図された動作を特定し、前記再生デバイスの前記意図された動作を特定することに応じて制御信号を生成させる、ための手段を備え、前記制御信号は、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。

【0017】

50

いくつかの例示的実装において、前記観察設定は観察点と観察向きとを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記再生デバイスの前記意図された動作に関連付けられた前記標示は、前記再生デバイスの連続再生モードにおける前記再生デバイスの意図された動作に関連付けられた標示と、前記再生デバイスのランダムアクセスモードにおける前記再生デバイスの意図された動作に関連付けられた標示とを含む。

【0018】

いくつかの例示的実装において、前記再生デバイスの前記意図された動作を特定することは、前記再生デバイスの前記意図された動作に関連付けられた条件が満たされたかどうかを判断することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記条件は、前記初期観察設定に関連付けられた少なくとも1つのリセット条件を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記条件は、前記初期観察設定に関連付けられた少なくとも1つの維持条件を含む。

10

【0019】

いくつかの例示的実装において、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションの前記レンダリング動作は、前記オーディオビジュアルプレゼンテーションの部分を選択することを含む。

【0020】

さらなる例示的実施形態において、方法が提供される。前記方法は、オーディオビジュアルプレゼンテーションの伝送単位のセットに関連付けられた観察設定を検出することと、前記観察設定に関連付けられた条件が満たされたかどうかを判断することと、前記観察設定に関連付けられた条件が満たされたかどうかの判断に応じて、オーディオビジュアルプレゼンテーションの前記伝送単位のサブセットを選択することと、制御信号を生成させることと、を含み、前記制御信号は、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションの前記伝送単位の前記選択されたサブセットのレンダリング動作に関連付けられている。

20

【0021】

そのような方法のいくつかの例示的実装において、前記観察設定は観察点と観察向きのいずれかまたは両方を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定は、前記オーディオビジュアルプレゼンテーションの前記伝送単位のセットに関連付けられた最も取りうる視聴方向の標示を含む。

30

【0022】

いくつかの例示的実装において、前記制御信号は、再生デバイスの意図された動作の標示を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定に関連付けられた条件が満たされたかどうかを判断することは、前記再生デバイスに関連付けられた向きを特定することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定に関連付けられた条件が満たされたかどうかを判断することは、前記再生デバイスが連続再生モードであるかどうかを判断することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定に関連付けられた条件が満たされたかどうかを判断することは、前記再生デバイスがランダムアクセスモードであるかどうかを判断することを含む。

40

【0023】

さらに別の例示的実施形態において、装置が提供される。前記装置は、オーディオビジュアルプレゼンテーションの伝送単位のセットに関連付けられた観察設定を検出し、前記観察設定に関連付けられた条件が満たされたかどうかを判断し、前記観察設定に関連付けられた条件が満たされたかどうかの判断に応じて、オーディオビジュアルプレゼンテーションの前記伝送単位のサブセットを選択し、制御信号を生成させる、ための手段を備え、前記制御信号は、前記再生デバイス上での前記オーディオビジュアルプレゼンテーションの前記伝送単位の前記選択されたサブセットのレンダリング動作に関連付けられている。

【0024】

50

そのような装置のいくつかの例示的実装において、前記観察設定は観察点と観察向きのいずれかまたは両方を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定は、前記オーディオビジュアルプレゼンテーションの前記伝送単位のセットに関連付けられた最も取りうる視聴方向の標示を含む。

【0025】

いくつかの例示的実装において、前記制御信号は、再生デバイスの意図された動作の標示を含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定に関連付けられた条件が満たされたかどうかを判断することは、前記再生デバイスに関連付けられた向きを特定することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定に関連付けられた条件が満たされたかどうかを判断することは、前記再生デバイスが連続再生モードであるかどうかを判断することを含む。そのようないくつかの例示的実装、およびその他の例示的実装において、前記観察設定に関連付けられた条件が満たされたかどうかを判断することは、前記再生デバイスがランダムアクセスモードであるかどうかを判断することを含む。

10

【図面の簡単な説明】

【0026】

本開示の特定の例示的実施形態の概略を説明したが、ここで添付の図面を参照する。これらの図面は必ずしも縮尺に従ったものではない。

【0027】

【図1】図1は、本発明の一例示的実施形態による実装を実行しうる例示的システム環境を示す図である。

20

【0028】

【図2】図2は、本発明の一例示的実施形態により具体的に構成されうる装置のブロック図である。

【0029】

【図3】図3は、本発明の一例示的実施形態による仮想現実ビデオプレゼンテーションの状況において実行される一例示的実装のブロック図である。

【0030】

【図4】図4は、本発明の一例示的実施形態による多視点ビデオプレゼンテーションの状況において実行される一例示的実装のブロック図である。

30

【0031】

【図5】図5は、本発明の一例示的実施形態による多視点ビデオプレゼンテーションの状況において実行される別の例示的実装のブロック図である。

【0032】

【図6A】図6Aは、本発明の一例示的実施形態により、図2の装置などによって実行される一連の動作を示すフローチャートである。

【0033】

【図6B】図6Bは、本発明の一例示的実施形態により、図2の装置などによって実行される一連の動作を示すフローチャートである。

【0034】

40

【図7】図7は、本明細書に含まれる例示的実施形態のいくつかの説明に関連して言及する用語を視覚的に表現した図である。

【0035】

【図8】図8は、本明細書に含まれる例示的実施形態のいくつかの説明に関連して言及する用語を視覚的に表現した別の図である。

【0036】

【図9】図9は、本明細書に含まれる例示的実施形態のいくつかの説明に関連して言及する用語を視覚的に表現した別の図である。

【0037】

【図10】図10は、本発明の一例示的実施形態による包括的なオーディオビジュアル仮

50

想現実アプリケーションの状況において実行される一例示的実装のブロック図である。

【0038】

【図11】図11は、本発明の一例示的実施形態による、画像またはビデオの符号化のために仮想現実画像またはビデオコンテンツを処理する状況において実行される、例示的な画像の貼合せ、投影、およびマッピングプロセスのブロック図である。

【詳細説明】

【0039】

以降では、添付の図面を参照して、いくつかの実施形態をより詳細に説明する。ここで示すのは本発明のいくつかの実施形態であって、すべての実施形態ではない。実際には、本発明の様々な実施形態は多様な形式で具体化されてもよく、本明細書に定める実施形態に制限されると解釈するべきではない。むしろ、これらの実施形態は、適用される法的要件を本開示が満たすように提供される。全体を通して、同じ参照符号は同じ要素を示す。本明細書で用いる「データ」、「コンテンツ」、「情報」、および同様の用語は、本発明の実施形態により伝送、受信、および/または記憶可能なデータを指して同じ意味で用いられる場合がある。したがって、そのような用語のいずれかを用いることによって、本発明の実施形態の趣旨および範囲が制限されるとみなすべきではない。

【0040】

また、本明細書で用いる「回路構成」という用語は、(a)ハードウェアのみの回路実装(例えば、アナログ回路構成および/またはデジタル回路構成での実装など)、(b)本明細書に記載の1つ以上の機能を装置に実行させるべく協働する、1つ以上のコンピュータ可読メモリに記憶されたソフトウェアおよび/またはファームウェア命令を含むコンピュータプログラム製品と、回路との組合せ、および(c)ソフトウェアまたはファームウェアが物理的に存在しなくても動作のためにソフトウェアまたはファームウェアを必要とする、例えばマイクロプロセッサ(単数または複数)またはその一部である、回路、を指す。「回路構成」のこの定義は、任意の請求項を含め本明細書におけるこの用語の使用すべてに当てはまる。さらなる例として、本明細書で用いる「回路構成」という用語は、1つ以上のプロセッサおよび/またはその部分(単数または複数)、ならびに付随するソフトウェアおよび/またはファームウェアを備える、実装も含む。別の例として、本明細書で用いる「回路構成」という用語は、例えば、携帯電話用のベースバンド集積回路もしくはアプリケーションプロセッサ集積回路、またはサーバ、セルラネットワークデバイス、他のネットワークデバイス、および/または他のコンピューティングデバイスにおける類似の集積回路も含む。

【0041】

本明細書で定める、非一時的物理記憶媒体(例えば、揮発性または不揮発性メモリデバイス)を指す「コンピュータ可読記憶媒体」は、電磁信号を指す「コンピュータ可読伝送媒体」とは区別される。

【0042】

本明細書で用いる「マッピング」という用語は、状況に応じて、投影による平面上の画像データが、2次元平面にマッピングされるプロセス、または、当該プロセスから生成される画像フレームにマッピングされるプロセスのいずれかを指す。

【0043】

本明細書で用いる「観察向き」という用語は、レンダリングされる向きを指す。多くの状況において、これは通常は、コンテンツのレンダリングに用いるヘッドマウントディスプレイの向きと相対的な向きである。

【0044】

本明細書で用いる「観察点」という用語は、仮想現実オーディオ/ビデオの取得または再生のための3次元空間内の点または体積を指す。観察点は、通常は仮想現実オーディオ/ビデオの取得に用いるデバイスまたはリグの中心点、ならびに、オーディオおよびビデオトラックが配置されている3次元空間内の観察者の頭部の位置と同じである。場合によっては、観察点は、撮影デバイスまたはリグの1つの中心点ではなく、例えば、円形など

10

20

30

40

50

の軌道、領域、または体積に対応してもよい。場合によっては、観察者の頭部の位置を追跡し、頭部の回転に加えて頭部の動きに対してレンダリングを調整する場合、当該観察者の頭部の初期位置または参照位置として観察点を認識してもよい。

【 0 0 4 5 】

本明細書で用いる「観察設定」という用語は、観察点および観察向きを指す。使用可能な観察点が1つのみであるプレゼンテーションの状況では、観察設定用に当該観察点を明確に標示または確定する必要はない。

【 0 0 4 6 】

本明細書で用いる「投影」または「VR投影」という用語は、ある球状画像が、平面や立方体などの形状に投影されるプロセス、または、当該プロセスから生成される画像フレームに投影されるプロセスのいずれかを指す。VR投影の例には、正距円筒パノラマやキューブマップ投影などがある。状況によっては、投影という用語は、キューブマップなどの3次元形状を1つ以上の2次元平面にマッピングすることをさらに含むと理解されてもよい。このようなマッピングは、多数の2次元平面を同じフレーム（例えば1つの2次元平面など）にパッキングすることを含んでもよい。

【 0 0 4 7 】

本明細書で用いる「ビューポート」または「VRビューポート」という用語は、全方位視野のサブセットを指す。「ビューポート」という用語は、ユーザに対して現在表示されている全方位ビジュアルコンテンツのサブセット、および/または残りのビジュアルコンテンツから（例えば品質的な区別や分離可能な部分として、または動き制限タイルセットなどにより）区別して符号化される全方位ビジュアルコンテンツのサブセットを指す場合がある。これら2つの定義は修飾語によって区別してもよい。すなわち、前者をレンダリングされたビューポート、後者を符号化されたビューポートと呼んでもよい。場合によっては、ビューポートは向きと視野で表現されても、特定の投影フォーマット用の2次元座標系内で矩形などの領域によって表現されてもよい。後者の一例は、正距円筒パノラマ画像内の矩形である。ビューポートは複数のビューポートで構成されてもよい。これら複数のビューポートは結合して1つのビューポートを形成し、それぞれ異なるプロパティ（画質など）を有してもよい。

【 0 0 4 8 】

本明細書で用いる「向き」（例えばビューポートの向きなど）という用語は、座標系の角度座標で表現される場合がある。角度座標はヨー、ピッチ、ロールと呼ばれる場合があり、これらは、例えばy、x、zといった特定の座標軸周りの回転角度を標示する。ヨー、ピッチ、およびロールは、例えば、ビューポートの向きを標示するために用いてもよい。状況によっては、ビューポートの向きを制限してもよく、例えばロールを0に制限してもよい。そのようないくつかの例や他の例では、ヨーとピッチによって当該ビューポートの中心点のオイラー角を角度で標示する。ほとんどの状況において、ヨーはピッチより先に適用され、ヨーがY軸周りの回転で、ピッチがX軸周りの回転となる。同様に、ほとんどの状況において、角度は原点から離れる方向に見て時計方向に増加する。図7を参照すると、軸700はY軸702およびX軸704を含む。図7に示すように、ヨー706はY軸702周りの回転として描かれ、ピッチ708はX軸704周りの回転として描かれている。図8を参照すると、軸800は、Y軸804、X軸806、およびZ軸808を介して3次元空間802をマッピングするために用いられる。図8に示すように、ピッチ810およびヨー812を用いて、ベクトル816に沿ったビューポート814の中心点のオイラー角を標示することができる。

【 0 0 4 9 】

いくつかの例示的実装では、ビューポートの視野(Field Of View: FOV)を水平FOV(HorFov)および垂直FOV(VerFov)によって表現してもよい。状況によっては、例えば、HorFovがビューポートの水平視野を角度で標示し、VerFovがビューポートの垂直視野を角度で標示するように、HorFovおよびVerFovを定義してもよい。HorFovおよびVerFovを使用してビューポートのFOV

10

20

30

40

50

を表現する例を、図9に示している。図9では、図8と同じ3次元空間802を軸800（Y軸804、X軸806、およびZ軸808を含む）と共にマッピングしている。ビューポート814も同様に空間802内に配置されている。図9では、ピッチおよび/またはヨーを用いてビューポート814の中心点のオイラー角を表すのではなく、ビューポート814の視野をHorFov902およびVerFov904として表現できる例を示している。

【0050】

本明細書で用いる「グローバル座標系」という用語は、説明したような、観察点に原点を有する3次元座標系を指す場合がある。

【0051】

本明細書で用いる「ランダムアクセス」という用語は、ストリームの開始点以外の任意のポイントから復号を開始して、正確なまたは近似の再構築されたメディア信号（復号されたピクチャの表現など）をリカバリするデコーダの能力を指す場合がある。ランダムアクセスポイントおよびリカバリポイントは、ランダムアクセス動作を特徴付けるために用いてもよい。ランダムアクセスポイントは、復号を開始できる、メディアストリーム内の位置として定義されてもよい。この位置は、例えばビデオビットストリーム内のアクセス単位や符号化ピクチャなどである。リカバリポイントは、メディアストリーム内または再構築された信号内の最初の場所として定義されてもよく、復号が対応するランダムアクセスポイントから開始された場合、リカバリポイントに位置する、または出力順でリカバリポイントの次に位置する復号ピクチャなどのすべてのメディアのコンテンツが正確または略正確であることを特徴としてもよい。ランダムアクセスポイントがリカバリポイントと同じである場合、ランダムアクセス動作は瞬間的であり、同じでない場合は漸進的であってもよい。

【0052】

ランダムアクセスポイントにより、例えば、ローカルに格納されたメディアストリーム、およびメディアストリーミングに対してシーク、早送り、早戻しなどの動作を行うことができる。オンデマンドストリーミングに関連する状況では、サーバはシーク要求に対し、要求されたシーク動作の宛先に最も近い（かつ、多くの場合はその前の）ランダムアクセスポイントからデータの伝送を開始することによって応答することができ、および/または、デコーダは、要求されたシーク動作の宛先に最も近い（かつ、多くの場合はその前の）ランダムアクセスポイントから復号を開始することができる。ビットレートが異なる複数の符号化ストリーム間の切替えは、ユニキャストストリーミングにおいて伝送ビットレートと予想されるネットワークスループットとを一致させ、ネットワークの輻輳を回避するためによく用いられる方法である。別のストリームへの切替えをランダムアクセスポイントで行うことができる。また、ランダムアクセスポイントにより、ブロードキャストやマルチキャストを選択することができる。さらに、ランダムアクセスポイントは、ソースシーケンスにおけるシーンカットへの応答として、またはイントラピクチャ更新要求への応答として、符号化することができる。

【0053】

いくつかの例示的実装では、メディアファイルフォーマット標準の利用を想定している。これらの標準には、ISOベースメディアファイルフォーマット（ISO/IEC 14496-12、ISO BMFFと略されることもある）、MPEG-4ファイルフォーマット（ISO/IEC 14496-14、MP4フォーマットとしても知られる）、ネットワーク抽象レイヤ（Network Abstraction Layer：NAL）ユニット構造のビデオ用ファイルフォーマット（ISO/IEC 14496-15）、および3GPPファイルフォーマット（3GPP TS 26.244、3GPフォーマットとしても知られる）があるが、これらに限定されない。ISOベースメディアファイルフォーマットに基づいて、前述のすべてのファイルフォーマット（ISOベースメディアファイルフォーマット自体を除く）が派生している。

【0054】

ISOBMFFの概念、構造、および仕様の一部を、実施形態を実装しうる基礎となるコンテナファイルフォーマットの例として以下に示している。しかしながら、本発明の様子はISOBMFFに限定されず、それに基づいて本発明を部分的または完全に具現化しうる、可能な基礎の1つとして説明するものである。

【0055】

ISOBMFFの1つの構成単位はボックスと呼ばれる。各ボックスはヘッダとペイロードを含むことができる。ボックスヘッダはボックスのタイプおよびサイズ（通常はバイト単位）を標示する。ボックスは他のボックスを含むことができる。ISOファイルフォーマットでは、特定のタイプのボックス内に指定することができるボックスタイプを規定している。また、一部のボックスは各ファイル内に存在することが必須であるが、他のボックスの存在は任意である。さらに、一部のボックスタイプのボックスは、1つのファイル内に2つ以上存在することができる。したがって、ISOBMFFはボックスの階層構造を規定するものとみなすことができる。ISOベースメディアファイルの各ボックスは、4文字のコード（Four-Character Code：4CC、fourCC）で識別できる。4文字のコードは、32ビットの符号なし整数を用いて（特定の変換によって文字を8ビット値、特定のビットエンディアン、および特定のバイトエンディアンにすることを想定して）区別なく表現できる。ヘッダは、ボックスのタイプおよびサイズに関する情報を提供できる。

【0056】

ISOBMFFによると、ファイルはメディアデータおよびメタデータを含むことができ、それらのデータは個別のボックスに含めることができる。一例示的实施形態において、メディアデータはメディアデータ（mdat）ボックス内に提供でき、メタデータを格納するためにムービー（moov）ボックスを用いることができる。場合によっては、ファイルを操作可能にするために、mdatボックスとmoovボックスの両方が存在しなければならない。ムービー（moov）ボックスは1つ以上のトラックを含むことができ、各トラックは1つの対応するトラック（trak）ボックスに配置できる。各トラックは、トラックタイプを指定する4文字のコードで識別されるハンドラに関連付けられる。ビデオ、オーディオ、および画像の各シーケンスのトラックはメディアトラックと総称され、これらのメディアトラックは1つの基本メディアストリームを含む。他のトラックタイプには、ヒントトラックおよび時限メタデータトラックがある。トラックはオーディオフレームやビデオフレームなどのサンプルを含む。メディアトラックは、メディア圧縮フォーマット（およびISOBMFFでのカプセル化）に従ってフォーマットされたサンプル（メディアサンプルと呼ばれることもある）を指す。ヒントトラックは、標示された通信プロトコルによる伝送用にパケットを構築するための定型の命令を含むヒントサンプルを指す。この定型の命令は、パケットヘッダの構築、およびパケットペイロードの構築の手引きを含むことができる。パケットペイロードの構築では、他のトラックまたはアイテムに存在するデータを参照することができる。この場合、例えば、他のトラックまたはアイテムに存在するデータは、パケット構築プロセス中に、特定のトラックまたはアイテム内のどのデータがパケットにコピーされるかの指示に関する参照によって標示できる。時限メタデータトラックは、参照されたメディアサンプルおよび/またはヒントサンプルを記述するサンプルを指してもよい。1つのメディアタイプを提示するために、1つのメディアトラックを選択できる。

【0057】

「trak」ボックスはサンプルテーブル（Sample Table）ボックスを含む。サンプルテーブルボックスは、例えば、トラック内のメディアサンプルのすべての時間とデータのインデックス付けを含むことができる。通常、サンプルテーブルボックスはサンプル記述（Sample Description）ボックスを含む必要がある。サンプル記述ボックスは、通常、当該ボックスに含まれるサンプルエントリ数を指定する、エントリ数フィールドも含む。ほとんどの実装において、サンプル記述ボックスは少なくとも1つのサンプルエントリを含む必要がある。サンプルエントリのフォーマットは、そのトラックのハンドラタイプに応じて

10

20

30

40

50

異なる。サンプルエントリは、使用された符号化タイプに関する詳細情報と、その符号化に必要な任意の初期化情報を提供する。

【 0 0 5 8 】

ムービーフラグメント機能により、そうしなければムービーボックス内に存在するであろうメタデータを複数のデータ部分に分割することができる。各データ部分は、トラックの特定の期間に対応しうる。換言すると、ムービーフラグメント機能により、ファイルメタデータとメディアデータのインターリーブを可能にすることができる。これにより、ムービーボックスのサイズを制限し、前述の使用例を実現することができる。

【 0 0 5 9 】

いくつかの例において、ムービーフラグメント用のメディアサンプルはmdatボックス内に存在してもよい。一方、ムービーフラグメントのメタデータ用にmoofボックスを提供してもよい。moofボックスは、以前はmoovボックスに含まれていた、任意の再生持続時間に関する情報を含むことができる。moovボックスは依然としてそれ自体で有効なムービーを表すが、加えて、同じファイル内でその後ムービーフラグメントがあることを標示するmvexボックスを含むことができる。ムービーフラグメントは、moovボックスに関連付けられているプレゼンテーションを時間的に拡張することができる。

【 0 0 6 0 】

ムービーフラグメント内には、トラックごとにゼロから複数の任意の数のトラックフラグメントのセットがあってもよい。トラックフラグメントは、ゼロから複数の任意の数のトラックラン (track run) を含むことができる。各トラックランは、そのトラックに対するサンプルの連続実行を記述する (したがってチャンクに類似している) 。これらの構造内の多くのフィールドは任意であり、デフォルト値を用いることができる。moofボックスに含まれるメタデータは、moovボックスに含まれるメタデータのサブセットに限定でき、場合によっては異なるように符号化できる。moofボックスに含めることができるボックスに関する詳細は、I S O B M F F 仕様で確認できる。自己完結型のムービーフラグメントを定義することもできる。このムービーフラグメントは、ファイルの順序において連続したmoofボックスとmdatボックスとで構成され、mdatボックスは (moofボックスがメタデータを提供する) 前述のムービーフラグメントのサンプルを含み、他のムービーフラグメント (すなわち、他の任意のmoofボックス) のサンプルは含まない。

【 0 0 6 1 】

I S O B M F F およびその派生 (例えばN A L ユニット構造化ビデオ用のファイルフォーマット (I S O / I E C 1 4 4 9 6 - 1 5) など) におけるサンプルグルーピングは、グルーピング基準に基づいて、トラック内の各サンプルを1つのサンプルグループのメンバーとして割り当てることとして定義することができる。サンプルグルーピングにおけるサンプルグループは、連続したサンプルに限定されず、不連続のサンプルも含むことができる。1つのトラック内のサンプルに対し2つ以上のサンプルグルーピングがありうるため、各サンプルグルーピングは、グルーピングのタイプを標示するタイプフィールドを有してもよい。サンプルグルーピングは、(1) サンプルグループへのサンプルの割当てを表すSampleToGroupボックス (sbgpボックス) と、(2) そのグループのプロパティを記述した、各サンプルグループのサンプルグループエントリを含むSampleGroupDescriptionボックス (sgpdボックス) の、2つのリンクされたデータ構造によって表現することができる。異なるグルーピング基準に基づくSampleToGroupボックスおよびSampleGroupDescriptionボックスのインスタンスが複数存在してもよい。これらは、グルーピングのタイプを標示するために用いるタイプフィールドによって区別してもよい。「sbgp」ボックスと「sgpd」ボックスはgrouping_typeの値を用いて、またこれらのボックスのいくつかの変形ではgrouping_type_parameterの値も用いて、リンクできる。「sbgp」ボックスは、特定のサンプルが属するサンプルグループ記述エントリのインデックスを標示する。

【 0 0 6 2 】

M a t r o s k a ファイルフォーマットでは、ビデオトラック、オーディオトラック、ピクチャトラック、または字幕トラックのいずれも1つのファイルに格納することができ

10

20

30

40

50

る（が、これに限定されない）。Matroskaファイルの拡張子には、ビデオ用の .mkv（字幕およびオーディオ付き）、立体ビデオ用の .mk3d、オーディオ専用ファイル用の .mka、および字幕専用の .mksがある。Matroskaは、WebMなどの派生ファイルフォーマット用の基本フォーマットとして用いることができる。

【0063】

Matroskaは、EBML（Extensible Binary Meta Language）を基礎として用いている。EBMLは、XMLの原理に基づく、バイナリとオクテット（バイト）が整列されたフォーマットを規定している。EBML自体は、バイナリマークアップの技法による一般的な記述である。Matroskaファイルは、EBML「文書」を構成する要素（Element）で構成される。これらの要素は、要素ID、その要素のサイズの記述子、およびバイナリデータ自体を含んでいる。要素は入れ子にすることができる。

10

【0064】

MatroskaのSegment要素は、他のトップレベル（レベル1）要素のコンテナである。Matroskaファイルは1つのセグメント（ただしそのみに制限されない）を含むことができる。Matroskaファイル内のマルチメディアデータは、複数のクラスタ（すなわちCluster要素）として編成される。各クラスタは一般的に数秒のマルチメディアデータを含む。クラスタはBlockGroup要素を含み、BlockGroup要素はBlock要素を含む。Cues要素は、ランダムアクセスまたはシークを支援しうるメタデータを含み、シークポイントに対するファイルポイントまたは対応するタイムスタンプを含みうる。

【0065】

URI（Uniform Resource Identifier）は、リソースの名前の識別に用いる文字列として定義できる。このような識別により、特定のプロトコルを用いて、ネットワークを介したリソースの表現との相互作用が可能になる。URIは、URIの具体的な構文および関連付けられたプロトコルを指定するスキームによって定義される。URL（Uniform Resource Locator）およびURN（Uniform Resource Name）はURIの形式である。URLは、Webリソースを識別し、そのリソースの表現を操作または取得する手段を指定して、主要なアクセス方式とネットワーク上の場所を指定するURIとして定義できる。URNは、特定の名称空間で名称によってリソースを識別するURIとして定義できる。URNは、リソースの場所またはそのアクセス方法を示すことなくリソースを識別するために用いることができる。

20

30

【0066】

HTTP（Hypertext Transfer Protocol）は、ビデオストリーミングアプリケーションにおけるような、インターネットを介したリアルタイムマルチメディアコンテンツの配信に広く用いられている。Microsoft（登録商標）Smooth Streaming、Apple（登録商標）Adaptive HTTP Live Streaming、Adobe（登録商標）Dynamic Streamingなどの、HTTPを介した適応型ストリーミングの商用ソリューションが複数公開され、標準化プロジェクトが進められている。適応型HTTPストリーミング（Adaptive HTTP Streaming：AHS）は、第3世代パートナーシッププロジェクト（Third Generation Partnership Project：3GPP）リリース9の packets 交換ストリーミング（Packet-Switched Streaming：PSS）サービス（3GPP TS 26.234 Release 9: "Transparent end-to-end packet-switched streaming service (PSS); protocols and codecs"（透過的なエンドツーエンドの packets 交換ストリーミングサービス（PSS）；プロトコルおよびコーデック）において最初に標準化された。MPEGはMPEG DASH標準（ISO/IEC 23009-1: "Dynamic adaptive streaming over HTTP (DASH)-Part 1: Media presentation description and segment formats"（HTTPを介した動的適応型ストリーミング（DASH）、パート1：メディアプレゼンテーション記述およびセグメントフォーマット）のベースとして3GPP AHSリリース9を採用した。MPEG DASHと3GPP-DASHは互いに似た技術であるため、総称してDASHと呼ばれることがある。DASHの概念、フォーマット、および動作の一部を、実施形態を実装しうるビデオストリーミングシステムの例

40

50

として以下に示している。本発明の態様はDASHに限定されず、それに基づいて本発明を部分的または完全に具現化しうる、可能な基礎の1つとして説明するものである。

【0067】

DASHでは、マルチメディアコンテンツをHTTPサーバに格納し、HTTPを用いて配信できる。コンテンツは、メディアプレゼンテーション記述(Media Presentation Description: MPD)とセグメント(Segment)という2つの部分としてサーバに格納できる。MPDは、利用可能なコンテンツ、その様々な代替物、それらのURLアドレス、およびその他の特性のマニフェストを記述する。セグメントは、実際のマルチメディアビットストリームをチャンク形式で単一または複数のファイルに含んでいる。MPDは、クライアントがHTTPを介して動的適応型ストリーミングを確立するために必要な情報を提供する。MPDは、セグメントのGETリクエストを実行するための各セグメントのHTTP-URLなどの、メディアプレゼンテーションを記述する情報を含む。コンテンツを再生するために、DASHクライアントは、例えばHTTP、電子メール、Thumb Drive、ブロードキャスト、またはその他の転送方法を用いてMPDを取得できる。MPDを解析することによって、DASHクライアントは、プログラムのタイミング、メディアコンテンツの利用可能性、メディアタイプ、解像度、最小帯域幅と最大帯域幅、および、マルチメディアコンポーネントの様々な符号化された代替物の存在、アクセシビリティ機能、必要なデジタル著作権管理(Digital Rights Management: DRM)、ネットワーク上のメディアコンポーネントの場所、およびその他のコンテンツの特性を認識できる。この情報を用いて、DASHクライアントは、適切に符号化された代替物を選択し、例えば、HTTP GETリクエストを用いてセグメントをフェッチすることによってコンテンツのストリーミングを開始できる。ネットワークスループットの変動に対応するための適切なバッファの後、クライアントは後続セグメントのフェッチを継続でき、また、ネットワーク帯域幅の変動を監視できる。クライアントは、適切なバッファを維持するために、(より低いまたはより高いビットレートを有する)異なる代替物のセグメントをフェッチすることによって、使用可能な帯域幅に適応する方法を決定できる。

【0068】

DASHの状況では以下の定義を用いることができる。すなわち、メディアコンテンツコンポーネントまたはメディアコンポーネントは、メディアコンポーネントタイプが割り当てられた、メディアコンテンツの1つの連続したコンポーネントであり、個別にメディアストリームへと符号化できるコンポーネントとして定義できる。メディアコンテンツは、1つのメディアコンテンツ期間またはメディアコンテンツ期間の連続したシーケンスとして定義できる。メディアコンテンツコンポーネントタイプは、オーディオ、ビデオ、テキストなどの単一のメディアコンテンツタイプとして定義できる。メディアストリームは、メディアコンテンツコンポーネントが符号化されたものとして定義できる。

【0069】

DASHでは、以下のように、階層データモデルを用いてメディアプレゼンテーションを構造化している。メディアプレゼンテーションは1つ以上のピリオド(Period)のシーケンスで構成され、各ピリオドは1つ以上のグループ(Group)を含み、各グループは1つ以上のアダプテーションセット(Adaptation Set)を含み、各アダプテーションセットは1つ以上のリプレゼンテーション(Representation)を含み、各リプレゼンテーションは1つ以上のセグメントで構成される。グループは、同時に提示されないと想定されるアダプテーションセットの集合として定義できる。アダプテーションセットは、1つまたは複数のメディアコンテンツコンポーネントを符号化した互換性のあるもののセットとして定義できる。リプレゼンテーションは、一般的に、例えばビットレート、解像度、言語、コーデックなどの符号化の選択が異なる、メディアコンテンツまたはそのサブセットの代替選択肢の1つである。セグメントは、特定の持続時間のメディアデータと、含まれるメディアコンテンツを復号および提示するためのメタデータを含む。セグメントはURIで識別され、通常はHTTP GETリクエストによって要求できる。セグメントは、HTTP-URLに関連付けられたデータの単位として、また任意で、MPDによって指定さ

10

20

30

40

50

れるバイト範囲として、定義できる。

【 0 0 7 0 】

D A S HのM P DはX M L (Extensible Markup Language) に準拠しているため、X M Lに定義されている要素 (element) と属性 (attribute) によって指定できる。M P Dは以下の規約を用いて指定できる。X M L文書内の要素は先頭の大文字で識別され、Elementのように太字で表示されうる。要素Element1が別の要素Element2に含まれていることを表現するには、Element2.Element1と記述できる。要素の名前が2つ以上の語の組合せである場合、例えばImportantElementのようにキャメルケースを用いることができる。要素は1回のみ出現するか、または<minOccurs> ... <maxOccurs>によって最小出現回数と最大出現回数を定義することができる。X M L文書内の属性は、先頭の小さい文字によって識別され、その前に例えば「@」符号を付すことができる (例 : @attribute) 。要素Elementに含まれる特定の属性@attributeを指すには、Element@attributeと記述できる。属性の名前が2つ以上の語の組合せである場合、例えば@veryImportantAttributeのように、先頭の語の後にキャメルケースを用いることができる。X M Lでは、属性に必須 (M) 、任意 (O) 、任意かつデフォルト値あり (O D) 、および条件付きで必須 (C M) というステータスを割り当てることができる。

【 0 0 7 1 】

D A S Hでは、一般的に、すべての記述子要素は同じように構造化される。つまり、スキームを識別するためのU R Iを提供する@schemelDuri属性と、任意の属性@valueと、任意の属性@idとを含む。要素の意味は、採用するスキーム固有である。スキームを識別するU R Iは、U R NまたはU R Lでありうる。一部の記述子はM P E G - D A S H (I S O / I E C 2 3 0 0 9 - 1) で指定されるが、加えてまたは代わりに、記述子を別の仕様で指定することができる。M P E G - D A S H以外の仕様で指定する場合、M P Dは記述子要素の使用方法に関する具体的な情報を提供しない。適切なスキーム情報による記述要素のインスタンス化は、D A S Hフォーマットを採用するアプリケーションまたは仕様次第である。これらの要素の1つを用いるアプリケーションまたは仕様は、U R Iの形式でスキーム識別子を定義し、そのスキーム識別子を用いる際の要素に対する値空間を定義する。スキーム識別子は@schemelDuri属性内に示される。列挙された値の単純なセットが要求される場合、値ごとにテキスト文字列を定義し、この文字列を@value属性内に含めることができる。構造化データが要求される場合、別個の名前空間内に拡張要素または拡張属性を定義することができる。@id値は、一意の記述子または記述子のグループを指すために用いることができる。後者の場合、属性@idの値が同じである複数の記述子は同義であることを求められる場合がある。すなわち、同じ@id値を持つ記述子のうちの1つのみを処理すればよい。DescriptorTypeタイプの2つの要素は、要素名、@schemelDuriの値、および@value属性の値が等価である場合、等価である。@schemelDuriがU R Nである場合、R F C 2 1 4 1の第5項に定められているように、等価性は語彙的な等価性を意味しうる。@schemelDuriがU R Lである場合、R F C 3 9 8 6の第6 . 2 . 1項に定められているように、等価性は文字ごとの等価性を意味しうる。@value属性が存在しない場合、等価性は@schemelDuriの等価性のみによって決定されうる。拡張名前空間内の属性および要素は、等価性の決定に用いられない場合がある。@id属性は、等価性の決定の際に無視されてもよい。

【 0 0 7 2 】

M P E G - D A S Hは、記述子EssentialPropertyおよびSupplementalPropertyを規定している。要素EssentialPropertyの場合、この要素が別のEssentialProperty要素と同じ@idを共有しないかぎり、この記述子を含む親要素内の情報を適切に使用するには、この記述子を正しく処理することが必須であることをメディアプレゼンテーション作成者が表している。複数のEssentialProperty要素が同じ@idを有する場合、それらのEssentialProperty要素のうちの一つのみを処理すればよい。個別の@id値を持つ少なくとも一つのEssentialProperty要素が処理されると想定される。EssentialProperty記述子のスキームまたは値が認識されない場合、D A S Hクライアントは、この記述子を含む親要素を無視する

10

20

30

40

50

と想定される。同じ@id値を持つEssentialProperty要素と、異なる@id値を持つEssentialProperty要素とが、1つのMPD内に複数存在してもよい。

【0073】

要素SupplementalPropertyの場合、この記述子が、最適化された処理のためにDASHクライアントによって使用されうる補足情報を含むことをメディアプレゼンテーション作成者が表している。SupplementalProperty記述子のスキームまたは値が認識されない場合、DASHクライアントはこの記述子を無視すると想定される。SupplementalProperty要素は1つのMPD内に複数存在してもよい。

【0074】

MPEG-DASHは、プロパティ記述子としてフォーマットされるViewpoint要素を規定している。Viewpoint要素の@schemeIdUri属性は、採用されたビューポイントスキームを識別するために用いられる。等価ではないViewpoint要素値を含む複数のアダプテーションセットは、異なるメディアコンテンツコンポーネントを含む。Viewpoint要素は、ビデオ以外のメディアコンテンツタイプにも同様に適用できる。等価のViewpoint要素値を有する複数のアダプテーションセットは一緒に提示されることが意図されている。この処理は、認識された@schemeIdUri値と認識されない@schemeIdUri値に同様に適用されるべきである。

【0075】

空間関係記述子(Spatial Relationship Description: SRD)は、MPEG-DASHの規範附属文書Hに規定されている。以下は、SRD仕様の抜粋の一部を含んでいる。

【0076】

SRDスキームにより、MPD作成者は、空間オブジェクト間の空間関係を表すことができる。空間オブジェクトは、アダプテーションセットまたはサブプレゼンテーション(Sub-Representation)のいずれかによって表現される。例として、空間関係は、任意のビデオが別のフルフレームビデオの空間部分(例えば関心領域やタイル)を表現することを表すことができる。

【0077】

@schemeIdUriが「urn:mpeg:dash:srd:2014」であるSupplementalProperty記述子および/またはEssentialProperty記述子は、含まれる空間オブジェクトに関連付けられた空間関係情報を提供するために用いられる。SRDは、これら2つのMPD要素(アダプテーションセットおよびサブプレゼンテーション)のみに含まれるものとする。

【0078】

1つのリプレゼンテーション内で複数の空間オブジェクト(HEVCのタイルストリームなど)を表現するには、サブプレゼンテーションレベルのSRDを用いることができる。その場合、SRD記述子は、アダプテーションセットレベルとサブプレゼンテーションレベルに存在しうる。

【0079】

SRDスキームを用いるSupplementalProperty要素またはEssentialProperty要素の@valueは、SRDパラメータのカンマ区切りの値リストである。SRDパラメータsource_id、object_x、object_y、object_width、およびobject_heightの存在は必須であり、SRDパラメータtotal_width、total_height、およびspatial_set_idの存在は条件付きまたは任意である。

【0080】

source_idは、十進数で表現される負ではない整数であり、コンテンツのソースに対する識別子を提供する。source_idパラメータは、ピリオド内でコンテンツのソースに対する一意の識別子を提供する。このパラメータは、当該ソースに関連付けられた座標系を暗黙的に定義する。この座標系は任意の原点(0;0)を使用し、x軸は左から右に配置され、y軸は上から下に配置される。同じsource_id値を有するすべてのSRDは、原点および軸の配置が同じである。異なるsource_id値を有するSRDを用いる空間オブジェクトの空間関係は定義されない。

10

20

30

40

50

【 0 0 8 1 】

所定のsource_id値に対し、参照空間が定義される。この参照空間は、ソースコンテンツ全体を包含し、左上の角が座標系の原点である矩形領域に対応する。SRDのtotal_width値およびtotal_height値は、この参照空間のサイズを任意の単位で表す。total_widthは、十進数で表現される負ではない整数であり、参照空間の幅を任意の単位で表す。total_heightは、十進数で表現される負ではない整数であり、参照空間の高さを任意の単位で表す。コンテンツのソース全体をカバーする空間オブジェクトがMPD内になくてもよい。例えば、ソースコンテンツ全体が2つの個別のビデオによって表現される場合などである。

【 0 0 8 2 】

object_xは、十進数で表現される負ではない整数であり、空間オブジェクトの左上の角の水平位置を任意の単位で表す。object_yは、十進数で表現される負ではない整数であり、空間オブジェクトの左上の角の垂直位置を任意の単位で表す。object_widthは、十進数で表現される負ではない整数であり、空間オブジェクトの幅を任意の単位で表す。object_heightは、十進数で表現される負ではない整数であり、空間オブジェクトの高さを任意の単位で表す。object_xおよびobject_yパラメータは、ソースに関連付けられた座標系内における、関連付けられた空間オブジェクトの2D位置を表し、object_widthおよびobject_heightパラメータはその2Dサイズを表す。object_x、object_y、object_width、およびobject_heightパラメータの各値は、前述の定義のとおり、total_widthおよびtotal_heightパラメータの値と相対的である。同じsource_id値を有する複数のSRDの位置(object_x、object_y)およびサイズ(object_width、object_height)は、参照空間のサイズを考慮した後、すなわちobject_x値とobject_width値を対応する記述子のtotal_width値によって除算した後、およびobject_y値とobject_height値を対応する記述子のtotal_height値によって除算した後に比較できる。異なるtotal_width値とtotal_height値を異なる記述子で用いることによって、同じ参照空間に対して異なる単位で位置情報およびサイズ情報を提供できる。

【 0 0 8 3 】

spatial_set_idは、十進数で表現される負ではない整数であり、空間オブジェクトのグループに対する識別子を提供する。存在しない場合、この記述子に関連付けられた空間オブジェクトはどの空間セットにも属さず、空間セット情報が提供されない。MPD作成者は、spatial_set_idパラメータを用いて、所定のsource_id内のいくつかの空間オブジェクトが特定の空間関係を有することを表すことができる。例えば、MPD作成者は、同じ解像度レベルの複数のタイルに対応するすべてのアダプテーションセットをグループ化することができる。このように、DASHクライアントはspatial_set_idパラメータを用いて、空間的に関連付けられた空間オブジェクトを迅速に選択することができる。

【 0 0 8 4 】

初期化セグメント(Initialization Segment)は、メディアセグメント(Media Segment)にカプセル化されたメディアストリームを提示するために必要なメタデータを含むセグメントとして定義できる。ISOBMFFベースのセグメントフォーマットでは、初期化セグメントはムービーボックス(「moov」)を含むことができ、このムービーボックスはいかなるサンプルのメタデータも含まない場合がある。すなわち、サンプルの任意のメタデータは「moof」ボックスに提供される。

【 0 0 8 5 】

メディアセグメントには、通常速度の再生における特定の持続時間のメディアデータが含まれる。このような持続時間はメディアセグメント持続時間またはセグメント持続時間と呼ばれる。コンテンツプロデューサまたはサービスプロバイダは、サービスの望ましい特性に従ってセグメント持続時間を選択することができる。例えば、ライブサービスにおいて比較的短いセグメント持続時間を用いて、全体的な待ち時間を短くすることができる。その理由は、セグメントはDASH用のメディアデータ生成における個別の単位であるため、一般的にセグメント持続時間は、DASHクライアントが知覚する全体的な待ち時

10

20

30

40

50

間の下限であるからである。コンテンツは、一般的に、メディアデータの1つのセグメント全体がサーバに利用可能になるように生成される。また、多くのクライアント実装が、1つのセグメントをGETリクエストの単位として用いる。したがって、一般的なライブサービスの構成では、DASHクライアントがセグメントを要求できるのは、持続時間全体のメディアセグメントが利用可能であり、かつ符号化され1つのセグメントにカプセル化されている場合のみである。オンデマンドサービスの場合、セグメント持続時間を選択する様々な方式を用いることができる。

【0086】

例えば、セグメントを複数の部分としてダウンロードできるようにするために、1つのセグメントを複数のサブセグメントにさらに分割することができる。サブセグメントは、アクセス単位全体を含む必要がある。サブセグメントはセグメントインデックスボックスによってインデックス付けできる。セグメントインデックスボックスは、プレゼンテーション時間範囲とバイト範囲を各サブセグメントにマッピングするための情報を含んでいる。セグメントインデックスボックスは、セグメント内のサブセグメントとストリームアクセスポイントを、それらの持続時間とバイトオフセットをシグナリングすることで示すこともできる。DASHクライアントは、セグメントインデックスボックス(単数または複数)から取得した情報を用いてHTTP GETリクエストを実行し、バイト範囲HTTPリクエストを用いて特定のサブセグメントを取得できる。比較的長いセグメント持続時間を用いる場合、HTTPレスポンスのサイズをビットレート適応に対して妥当かつ柔軟にするために、サブセグメントを用いることができる。セグメントのインデックス付け情報は、そのセグメントの先頭で単一のボックス内に入れるか、またはそのセグメント内の多数のインデックス付けボックス内に分散させてもよい。例えば、階層、デージーチェーン、ハイブリッドなどの様々な分散方法が可能である。この技法により、セグメントの先頭に大きなボックスを追加することを回避できるため、初期のダウンロード遅延を防ぐことができる。

【0087】

サブリプレゼンテーションは通常のリプレゼンテーションに埋め込まれ、SubRepresentation要素によって記述される。SubRepresentation要素はRepresentation要素に含まれる。SubRepresentation要素は、リプレゼンテーションに埋め込まれた1つまたは複数のメディアコンテンツコンポーネントのプロパティを記述する。この要素は、例えば、埋め込まれたオーディオコンポーネントのプロパティ(例えば、コーデック、サンプリングレートなど)や埋め込まれた字幕のプロパティ(例えばコーデックなど)そのものを記述したり、埋め込まれたより低品質のビデオレイヤ(例えば、より低いフレームレートなど)を記述したりすることができる。サブリプレゼンテーションとリプレゼンテーションは、共通の属性および要素をいくつか有する。@level属性がSubRepresentation要素内に存在する場合、以下のようになる。

- ・ サブリプレゼンテーションにより、それらが含まれるリプレゼンテーションの低品質バージョンにアクセスする能力が提供される。この場合、サブリプレゼンテーションにより、例えば、多重化リプレゼンテーション内のオーディオトラックを抽出したり、より低いフレームレートで提供される場合の早送り操作と巻戻し操作を効率的にしたりすることができる。

- ・ 初期化セグメントおよび/またはメディアセグメントおよび/またはインデックスセグメントは、HTTPの部分GETリクエストによってデータに容易にアクセスできるように十分な情報を提供するものとする。そのような情報の提供の詳細は、使用するメディアフォーマットによって定義される。

- ・ ISO BMFFのセグメントを用いる場合、以下のようになる。
 - 初期化セグメントはレベル割当てボックスを含む。
 - サブセグメントインデックスボックス(「ssix」)が各サブセグメントに存在する。
 - 属性@levelは、サブセグメントインデックス内で前述のサブリプレゼンテーション

10

20

30

40

50

ンが関連付けられるレベルを指定する。リプレゼンテーション内、サブリプレゼンテーション内、およびレベル割当て（「leva」）ボックス内の情報は、レベルへのメディアデータの割当てに関する情報を含む。

- メディアデータは、各レベルがより低いレベルと比べて向上しているような順序であるべきである。

【0088】

@level属性が存在しない場合、リプレゼンテーションに埋め込まれたメディアストリームのより詳細な記述は、SubRepresentation要素のみを用いて提供される。

【0089】

ISOBMFFは、ファイルのサブセットを指定するための、いわゆるレベル機構を含む。各レベルは、レベルnにマッピングされたサンプルがレベルm ($m \leq n$)の任意のサンプルに依存し、レベルp ($p > n$)のサンプルには依存しないように、依存階層に従う。例えば、レベルは時間サブレイヤ（例えばHEVCのTemporalId）に従って指定することができる。レベルは、MovieExtends（「mvex」）ボックスに含まれるレベル割当て（「leva」）ボックス内で宣言することができる。レベルは最初のムービーには指定できない。レベル割当てボックスは、存在する場合、最初のムービーに後続するすべてのムービーフラグメントに適用される。レベル割当てボックスの状況では、フラクシオンは、1つ以上のムービーフラグメントボックスおよび関連付けられたメディアデータボックスを含み、場合によっては最後のメディアデータボックスはその最初の部分のみ含むように定義される。フラクシオン内では各レベルのデータは連続的に配置される。フラクシオン内の各レベルのデータは、レベル値の昇順に配置される。フラクシオン内のすべてのデータはレベルに割り当てられる。レベル割当てボックスにより、スケーラビリティレイヤや時間サブレイヤなどの機能からレベルへのマッピングが行われる。機能は、トラック、トラック内のサブトラック、またはトラックのサンプルグルーピングによって指定できる。例えば、時間レベルサンプルグルーピングを用いて、ピクチャから時間レベルへのマッピングを標示することができる。これらの時間レベルはHEVCの時間サブレイヤに相当する。すなわち、所定のTemporalId値のHEVCピクチャを、時間レベルサンプルグルーピングを用いて特定の時間レベルへとマッピングすることができる（かつ、すべてのTemporalId値に対して同様に繰り返すことができる）。レベル割当てボックスはその後、標示されたレベルへのマッピングについて時間レベルサンプルグルーピングを参照することができる。

【0090】

サブセグメントインデックスボックス（「ssix」）により、（レベル割当てボックスによって指定された）レベルから、インデックス付けされたサブセグメントのバイト範囲へのマッピングが行われる。換言すると、このボックスは、レベルに従ってサブセグメント内のデータを部分サブセグメントとして順序付けるための簡易なインデックスを提供する。これによりクライアントは、サブセグメント内のデータの範囲をダウンロードすることによって、部分サブセグメントのデータに容易にアクセスすることができる。サブセグメントインデックスボックスが存在する場合、サブセグメント内の各バイトは1つのレベルに割り当てられる。範囲がレベル割当て内のいかなる情報にも関連していない場合、レベル割当てに含まれていない任意のレベルを用いてもよい。セグメントインデックスボックスごとに0または1つのサブセグメントインデックスボックスが存在する。サブセグメントインデックスボックスは、リーフサブセグメントのみをインデックス付けする、すなわちサブセグメントのみをインデックス付けし、セグメントはインデックス付けしない。サブセグメントインデックスボックスは、存在する場合、関連付けられたセグメントインデックスボックスの次のボックスである。サブセグメントインデックスボックスは、直前のセグメントインデックスボックス内に標示されたサブセグメントについて記述する。各レベルは1つの部分サブセグメントだけに割り当てることができる。すなわち、1つのレベルの複数のバイト範囲は連続的である。部分サブセグメントのレベルは、サブセグメント内で昇順に割り当てられる。すなわち、部分サブセグメントのサンプルは、同じサブセグ

10

20

30

40

50

メント内の前の部分サブセグメントの任意のサンプルに依存しうるが、その逆はない。例えば、各部分サブセグメントは同一の時間サブレイヤを有する複数のサンプルを含み、部分サブセグメントはサブセグメント内で時間サブレイヤの昇順に配置される。このように部分サブセグメントがアクセスされると、最後のメディアデータボックスが不完全になる、つまり、メディアデータボックスが標示する長さより少ないデータがアクセスされる場合がある。メディアデータボックスの長さの調整が必要になるか、埋込みが用いられることがある。レベル割当てボックス内のpadding_flagは、この欠落したデータをゼロに置換えできるかどうかを標示する。置換えできない場合、アクセスされないレベルに割り当てられたサンプルのサンプルデータは提示されないため、注意する必要がある。

【0091】

MPEG-DASHは、ISOBMFFとMPEG-2の転送ストリームの両方に対するセグメントコンテナフォーマットを定義している。他の仕様では、他のコンテナフォーマットに基づいてセグメントフォーマットを規定している場合がある。例えば、Matroskaコンテナファイルフォーマットに基づくセグメントフォーマットが提案されており、これは以下のように要約することができる。MatroskaファイルがDASHセグメントなどとして保持される場合、DASH単位とMatroska単位との関連付けは以下のように指定されうる。(DASHの)サブセグメントは、Matroskaカプセル化コンテンツの1つ以上の連続したクラスタとして定義できる。DASHの初期化セグメントは、EBMLヘッダ、(Matroskaの)セグメントヘッダ、(Matroskaの)セグメント情報、およびトラックを含む必要がありうる。また、任意で他のレベル1要素および埋込みを含むことができる。DASHのセグメントインデックスは、MatroskaのCues要素を含むことができる。

【0092】

DASHは、様々なネットワーク帯域幅を一致させるために、アダプテーションセット内の異なるリプレゼンテーションからメディアセグメントを動的に要求することによってレート適応に対応する。DASHクライアントがリプレゼンテーションを上下に切り替える際、リプレゼンテーション内の符号化依存性を考慮する必要がある。リプレゼンテーションの切替えはランダムアクセスポイント(Random Access Point: RAP)で実行できる。これはH.264/AVCなどのビデオ符号化技法で一般的に用いられている。DASHでは、ストリームアクセスポイント(Stream Access Point: SAP)と呼ばれるより一般的な概念が導入され、リプレゼンテーションへのアクセスおよびリプレゼンテーション間の切替えのためにコーデックに依存しない解決策が提供される。DASHでは、SAPはリプレゼンテーション内の位置として規定され、その位置から始まるリプレゼンテーションデータ(存在する場合、初期化セグメント内の初期化データが先行する)に含まれる情報のみを用いてメディアストリームの再生を開始することを可能にする。したがって、リプレゼンテーションの切替えはSAPで実行することができる。

【0093】

以下を含む複数のタイプのSAPが規定されている。SAPタイプ1は、一部の符号化スキームで「クローズドGOPランダムアクセスポイント」(すべてのピクチャが復号順で正しく復号され、正しく復号されたピクチャの途切れのない連続的な時間シーケンスとなる)として知られるものに対応し、これに加えて、復号順の最初のピクチャが提示順の最初のピクチャでもある。SAPタイプ2は、一部の符号化スキームで「クローズドGOPランダムアクセスポイント」(すべてのピクチャが復号順で正しく復号され、正しく復号されたピクチャの途切れのない連続的な時間シーケンスとなる)として知られるものに対応し、復号順の最初のピクチャが提示順の最初のピクチャではないことがある。SAPタイプ3は、一部の符号化スキームで「オープンGOPランダムアクセスポイント」として知られるものに対応し、復号順の一部のピクチャが正しく復号されず、プレゼンテーション時間が、SAPに関連付けられたイントラ符号化ピクチャより短い。

【0094】

MPEG-2などの一部のビデオ符号化標準では、各イントラピクチャが符号化シーケ

10

20

30

40

50

ンス内のランダムアクセスポイントであった。H.264/AVCやH.265/HEVCなどの一部のビデオ符号化標準における、インター予測に複数の参照ピクチャを柔軟に使用できる機能では、イントラピクチャがランダムアクセスには十分ではない場合がある。したがって、符号化タイプからそのような機能を推測するのではなく、ランダムアクセスポイント機能に関してピクチャをマーキングする場合がある。例えば、H.264/AVC規格に規定されているIDRピクチャをランダムアクセスポイントとして用いることができる。クローズドGOP (Group of Picture) は、その中のすべてのピクチャを正しく復号できるピクチャグループである。例えば、H.264/AVCでは、クローズドGOPはIDRアクセス単位から（または、メモリ管理制御動作によって以前のすべての参照ピクチャが未使用としてマーキングされる、イントラ符号化ピクチャから）開始することができる。

10

【0095】

オープンGOPは、出力順で最初のイントラピクチャより前のピクチャは正しく符号化できない場合があるが、最初のイントラピクチャの後のピクチャは正しく符号化できるような、ピクチャのグループである。そのような最初のイントラピクチャは、例えばHEVCのCRANALユニットタイプ、またはH.264/AVCのリカバリポイントSEIメッセージを用いて、ビットストリーム内に標示、および/またはビットストリームの標示から判断できる。オープンGOPを開始する最初のイントラピクチャより前のピクチャは、先行ピクチャと呼ばれる場合がある。先行ピクチャには復号可能と復号不可の2つのタイプがある。復号可能な先行ピクチャは、オープンGOPを開始する最初のイントラピクチャから復号が開始される場合に正しく復号できるピクチャである。換言すると、復号可能な先行ピクチャは、インター予測における参照として、復号順で最初のイントラピクチャまたは後続のピクチャのみ用いる。復号不可の先行ピクチャは、オープンGOPを開始する最初のイントラピクチャから復号が開始される場合に正しく復号できないピクチャである。

20

【0096】

前述のように、クライアントまたはプレーヤーは、スケーラブルビデオビットストリームの伝送されるレイヤおよび/またはサブレイヤが決定される方法と同様に、セグメントまたはサブセグメントを異なるリプレゼンテーションから伝送するように要求してもよい。リプレゼンテーションの下方切替えまたはビットストリームの下方切替えという用語はそれぞれ、以前に要求または伝送されたものより低いビットレートのリプレゼンテーションの要求または伝送を指してもよい。リプレゼンテーションの上方切替えまたはビットストリームの上方切替えという用語はそれぞれ、以前に要求または伝送されたものより高いビットレートのリプレゼンテーションの要求または伝送を指してもよい。リプレゼンテーション切替えまたはビットストリーム切替えという用語は、リプレゼンテーションまたはビットストリームの上方切替えおよび下方切替えの総称であってもよく、加えてまたは代わりに、異なるビューポイントのリプレゼンテーションまたはビットストリームの切替えを含んでもよい。

30

【0097】

MPEG-DASHに類似したストリーミングシステムには、例えば、IETFのインターネットドラフトdraft-pantos-http-live-streaming-19（および同じインターネットドラフトの他のバージョン）に規定されたHTTP Live Streaming (HLSとも呼ばれる)がある。MPDに対応するマニフェストフォーマットとして、HLSは拡張M3Uフォーマットを用いる。M3Uはマルチメディアプレイリスト用のファイルフォーマットであり、元はオーディオファイル用に開発されたものである。M3Uプレイリストは、個々の行で構成されるテキストファイルであり、各行はURI、空白、または「#」文字で開始されタグまたはコメントを標示する。URIの行は、メディアセグメントまたはプレイリストファイルを識別する。タグは#EXTで開始される。HLSの仕様は多数のタグを規定しており、それらはキー/値ペアとみなすことができる。タグの値部分は属性リストでありうる。これは属性/値ペアのカン

40

50

マ区切りのリストであり、属性/値ペアはAttributeName=AttributeValue構文を有してもよい。したがって、HLS M3U8 ファイルのタグはMPDまたはXMLの要素に類似しているとみなすことができ、HLS M3U8 ファイルの属性は、MPDまたはXMLの属性に類似しているとみなすことができる。HLSの一部のバージョンでは、メディアセグメントはMP EG - 2のトランスポートストリームに従ってフォーマットされ、単一のMP EG - 2プログラムを含む。各メディアセグメントはPAT (Program Association Table) およびPMT (Program Map Table) で開始することが推奨されている。HLSの一部のバージョンでは、メディアセグメントは、DASHの(サブ)セグメントと同様に自己完結型のISOBMFFムービーフラグメントである。

【0098】

全体的なDASHシステムは以下のように構築することができる。メディアコンテンツがオリジンサーバによって提供される。オリジンサーバは一般的にはウェブ(HTTP)サーバである。このオリジンサーバはコンテンツデリバリネットワーク(Content Delivery Network: CDN)に接続されていてもよい。このネットワークを介して、ストリーミングされたコンテンツがエッジサーバに提供および格納される。MPDによって、そのコンテンツの複数のベースURLがシグナリングされる。これらのURLは、異なるエッジサーバにおけるそのコンテンツの利用可能性を宣言するために用いることができる。あるいは、コンテンツサーバがインターネットに直接接続されていてもよい。コンテンツの要求先のオリジンサーバまたはエッジサーバとDASHクライアントとの間のHTTPトラフィックをルーティングする経路上にウェブプロキシが存在してもよい。ウェブプロキシはHTTPメッセージをキャッシュできるため、クライアントの要求に対してキャッシュされたコンテンツを提供することができる。プロキシからオリジンサーバまたはエッジサーバまでに必要なネットワーク帯域幅が削減されるため、ウェブプロキシはネットワークサービスプロバイダによって一般的に用いられる。エンドユーザにとっては、HTTPキャッシングによって待ち時間が短くなる。DASHクライアントは、モバイルセルラネットワークなどのアクセスネットワークを介してインターネットに接続してもよい。

【0099】

ISO/IEC 23009-5はSAND (Server And Network assisted DASH)を規定している。SANDは、DASHクライアントとネットワーク要素との間、または様々なネットワーク要素間のメッセージを導入している。その目的は、ネットワーク、サーバ、プロキシ、キャッシュ、CDNのリアルタイム動作の特性、およびDASHクライアントのパフォーマンスとステータスに関する情報を提供することによってストリーミングの効率を向上することである。例えば、SANDのAnticipatedRequestsメッセージによって、DASHクライアントはDASH対応ネットワーク要素(DASH-Aware Network Element: DANE)に対し、関心のある特定のセグメントセットを宣言することができる。この意図は、DASHクライアントがまもなく選択および要求しそうなリプレゼンテーション内のセグメントセットをシグナリングすることである。メッセージペイロードには予想される要求のリストが含まれ、それぞれ当該要求のURLと、任意でその要求のバイト範囲(そのURLによって示されるコンテンツの一部のみが要求されると予想される場合)と、任意で、そのURLによって識別されるリソースに対してDASHクライアントが要求を発行すると予想される時間が含まれる。

【0100】

仮想現実ビデオコンテンツは、異なる複数の投影フォーマットを用いてもよい。「360°ビデオ」という用語は、「仮想現実ビデオ」という用語と区別なく使用されうる。360度の水平視野と180度の垂直視野をカバーする球状画像から、矩形の2次元画像平面への特定の投影は、正距円筒図法として知られている。この場合、変換や拡大縮小なしで、水平座標は経度に相当し、垂直座標は緯度に相当するとみなすことができる。場合によっては、360度の水平視野と180度未満の垂直視野を有するパノラマコンテンツは、正距円筒図法の特別なケースとみなすことができ、その球体の極領域が2次元画像平面上にマッピングされていない。正距円筒図法では、垂直線の直線度が維持されるが、天底

10

20

30

40

50

と天頂の領域が変形する。

【0101】

キューブマップ投影フォーマット（キューブマップとも呼ばれる）では、球体ビデオは立方体の六面（側面とも呼ばれる）上に投影される。キューブマップは、例えば、最初に視点から球体シーンを6回レンダリングすることによって生成することができる。この際、各ビューは、立方体の各面を表現する90度の視錐台によって画定される。この立方体の各側面は、同じフレームにフレームパッキングするか、または立方体の各側面を個別に（例えば符号化において）処理することができる。立方体の各側面を多様な順序でフレームに配置することができ、および/または、立方体の各側面を回転または鏡像化することができる。フレームパッキング用のフレームの幅と高さは、立方体の側面が例えば3×2の立方体側面グリッドに「ぴったりと」はめ込まれるように選択するか、または、例えば4×3の立方体側面グリッドに、未使用の構成フレームを含めることができる。

10

【0102】

一例において、仮想現実コンテンツは、図10に示す例示的プロセス1100、あるいはそのサブセットおよび/または変形に従って処理してもよい。図10に示すように、仮想現実コンテンツは、1つ以上の仮想現実カメラ、他のカメラ配列、および/または仮想現実コンテンツの撮影に適切な他のオーディオビジュアル機器などによって、ブロック1102において取得される。図10に示すように、ブロック1102で取得された画像はブロック1104に渡されてもよい。ブロック1104では、画像の貼合せ、投影、およびマッピングが行われてもよい。ブロック1102で取得されたオーディオは、一部の実装において、オーディオ符号化のためにブロック1106に渡されてもよい。ブロック1104からの貼合せ済み、投影済み、およびマッピング済みの画像は、ビデオ符号化のためにブロック1108に、および/または画像符号化のためにブロック1110に渡されてもよい。図10のブロック1112に示すように、ファイルエンキャプスレータは、ビデオ、画像、およびオーディオを含む符号化メディアコンテンツをブロック1106、1108、および1110から入力として取得し、それらを1つのコンテナファイルへとカプセル化する。ファイルエンキャプスレータは、復号されたパッキングVRフレームのレンダリングを支援する、投影およびマッピング情報などのメタデータを受信してファイルへと組み込んでよい。DASHに関連する実装では、図10に示すように、DASH MPDジェネレータがファイルを入力として取得し、ブロック1114においてMPDを生成する。このMPDには、ファイル内の相当する情報に基づいて生成することができる投影およびマッピングメタデータなどのVR固有のメタデータが含まれてもよい。ブロック1114においてDASH MPDが生成された後、ブロック1116においてDASHクライアント/サーバ転送が実行される。

20

30

【0103】

図10のブロック1118に示すように、ファイル再生中、ファイルデキャプスレータが当該ファイルおよび/または受信した（サブ）セグメントを処理し、トラックから符号化ビットストリームを抽出し、メタデータを解析する。オーディオ情報、ビデオ情報、および画像情報はそれぞれブロック1120、1122、および1124において復号される。ブロック1122および/または1124によって生成された、復号されたパッキングVRフレームはブロック1130においてレンダリングされてもよい。このレンダリングには、任意で、ファイルデキャプスレータから受信された投影および/またはマッピング情報が使用されてもよい。ブロック1132に示すように、レンダリングされた画像は、現在の視聴向き、およびファイルから解析した投影およびマッピングメタデータに基づいて、ヘッドマウントディスプレイまたは他の任意の表示デバイスの画面に投影される。同様に、ブロック1126においてレンダリングされたオーディオは、ブロック1128においてラウドスピーカおよび/またはヘッドフォンを介して出力されてもよい。

40

【0104】

画像の貼合せ、投影、およびマッピングのプロセス1200の例示的な詳細を図11に示しており、以下において説明する。VR画像またはビデオクリップは一般的に、複数の

50

カメラ、または複数のレンズとセンサを備えた1つのカメラデバイスを用いて撮影される。複数のカメラからの入力ピクチャ1204は、ブロック1206において貼り合わされ、球体や立方体などの3次元幾何学的構造に投影される。幾何学的構造上の画像データは、さらに2次元投影フレーム1208上に配置される。そのフォーマットは、ブロック1210においてVR投影フォーマット標示子によって標示されてもよい。一例において、このマッピングには、投影されたフレームの複数の矩形領域を、各領域の場所とサイズをパッキングVRフレーム1212内で標示することによって、パッキングVRフレーム1212上にマッピングすることが含まれる。一例において、このマッピングにはさらに、投影されたフレームの複数の矩形領域をパッキングVRフレーム1212上で鏡像化または回転させることのいずれかまたは両方が含まれる。鏡像化は水平および垂直の鏡像化に限定されてもよく、回転は90度単位に制限されてもよい。実際は、入力ピクチャ(単数または複数)は、中間ステップなしの1回の処理でパッキングVRフレーム1212に変換されてもよい。これを図11の点線の矩形1202で標示している。パッキングVRフレームは、ビデオ符号化1108および/または画像符号化1110への入力として提供される。パッキングVRフレームという用語は、投影されたフレームの単一の矩形領域のみがパッキングVRフレームにマッピングされる場合、またはパッキングVRフレームがそのような投影されたフレームを含む場合にも用いてよい。パッキングVRフレームという用語は、投影されたフレームのマッピングによって生じるフレームとして定義してもよい。

10

【0105】

20

投影構造は、VR画像/ビデオコンテンツが投影される1つ以上の表面で構成される3次元構造として定義されてもよい。投影されたフレームは、投影構造の表面(単数または複数)がマッピングされる2次元フレームとして定義されてもよい。投影されたフレームは、代わりにまたはさらに、VR投影フォーマット標示子によって定義される表現フォーマットを有するフレームとして定義されてもよい。例えば、キューブマップ投影の投影構造は立方体であり、キューブマップは、立方体の面を展開することによって形成される2次元の投影されたフレームである。VR投影フォーマット標示子は、例えば、投影されたフレームの表現フォーマットを標示する列挙型であってもよい。例えば、この標示子は、平面視の正距円筒パノラマ、立体視の正距円筒パノラマ、平面視のキューブマップ、および立体視のキューブマップのうちの1つを標示してもよい。立体視の投影フォーマットを標示する場合、特定のパッキング構成を事前定義するか、または別途標示してもよい。例えば、トップアンドボトムパッキング構成を事前定義してもよい。この構成では、例えば左側のビューが上に表示されるように定義することができる。

30

【0106】

いくつかの例において、異なる視聴向きのための複数のバージョンのVRビデオが符号化される。これにより、球体や立方体などの投影構造の向きが、目的とする視聴向きに従って回転される。投影構造や対応する投影されたフレームの向きを、グローバル座標系に対して標示するための様々な方法がありうる。例えば、正距円筒パノラマピクチャの中心点や、キューブマップの前面の中心点などの、主要点を投影フォーマットに定義してもよい。ヨーおよびピッチによってグローバル座標系内の主要点の場所を標示してもよい。投影構造または対応する投影されたフレームの向きをロールによって標示してもよい。この場合、参照方向に直交する主要平面がどのように回転されるかが標示される。

40

【0107】

グローバル向きオフセットという用語は、参照向きに対するヨー、ピッチ、およびロールとして定義されてもよい。参照向きは、レンダリングシステムまたはグローバル座標系において、(ヨー、ピッチ、ロール)が(0、0、0)に等しい場合に相当する。参照向きは、参照方向に直交し、ロール角がゼロ度である2次元表面の向きとして定義されてもよい。参照方向は、グローバル座標系のz軸またはカメラパラメータの座標系のz軸、あるいはマイクロフォン設定のゼロ方位角およびゼロ仰角の軸の方向として定義されてもよい。したがって、グローバル向きオフセットは、例えばコンテンツの符号化後にカメラま

50

たはコンテンツの向きを修正するために用いることができる。例えば、コンテンツの地平線が正確に水平ではない場合（例えば、向きがやや傾いたカメラでコンテンツが撮影された場合など）、VR向きメタデータによってこれを修正できる。

【0108】

グローバル向きオフセットは、例えば以下の1つ以上の方法でファイルに含まれてもよい。i)（例えば1つのトラック全体の）サンプルのセットに適用されるグローバル向きオフセットが、ISOBMFF対応ファイルのサンプルエントリに含まれてもよい。ii) グローバル向きオフセットに対してサンプルグループが定義されてもよく、各サンプルグループ記述エントリでヨー、ピッチ、およびロールの値の組合せが定義され、SampleToGroupボックスを用いてトラックのサンプルがサンプルグループ記述エントリにマッピングされる。iii) メタデータトラックとしてのVR向きが以下のように定義される。存在する場合、VR向きメタデータトラックは、同じグローバル向きオフセットデータを有する各ビデオトラックおよび各オーディオトラックへの、例えば「cdsc」タイプのトラック参照を含む。存在する場合、このメタデータはグローバル向きオフセットを指定する。このトラックが存在しない場合、グローバル向きオフセットのヨー、ピッチ、およびロールの値はそれぞれ(0、0、0)である。VR向きメタデータトラックのサンプルに提供されるグローバル向きオフセットは、「cdsc」タイプのトラック参照を用いて、VR向きメタデータトラックに関連付けられたトラックのすべての時間並列オーディオおよびビデオサンプルに適用される。特定のトラック内の特定のサンプルに対する時間並列サンプルは、同特定トラック内の同特定サンプルの復号時間と同じ復号時間を有する、または、同じ復号時間のサンプルがない場合は最も近い前の復号時間を有する、参照されるトラック内のサンプルとして定義されてもよい。

10

20

【0109】

グローバル向きオフセットは、VRオーディオビデオプレゼンテーション全体にわたって適用されてもよい。レンダリングにおいて、（ヘッドマウントディスプレイの最初の向きに対する）ヘッドマウントディスプレイの向きは、基本的に、復号コンテンツからの抽出に用いる向きを選択するために、その時点で有効なグローバル向きオフセットと合計される。例えば、ヨー、ピッチ、およびロールがそれぞれ(a、b、c)である向きのビデオが、例えばヘッドマウントディスプレイにレンダリングされ、ヨー、ピッチ、およびロールに対するグローバル向きオフセットがそれぞれ(i、j、k)である場合、ファイルに含まれる、向きに対応するビデオ情報のヨー、ピッチ、およびロールはそれぞれ(a - i、b - j、c - k)である。

30

【0110】

復号するトラックまたはリプレゼンテーションを選択する、および/または復号コンテンツから抽出するデータを選択するときに、投影構造または投影されたフレームの（グローバル座標系に対する）向きを考慮してもよい。例えば、投影構造が、45度のヨー角および0度のピッチ角とロール角を有すると標示される立方体であり、現在の視聴向きのヨー、ピッチ、およびロールが0である場合、復号キューブマップからのコンテンツは、レンダリングされたコンテンツの中心点がY軸周りに45度（すなわち投影されたフレーム内で水平に）オフセットされてレンダリングされるように選択される。

40

【0111】

いくつかの例において、グローバル向きオフセット（単数または複数）は、投影構造または投影されたフレームの向きの情報内に含まれるため、レンダリングする復号データを選択する際に別途考慮する必要はない。いくつかの例において、グローバル向きオフセット（単数または複数）は、投影構造または投影されたフレームの向きの情報とは別個であるため、レンダリングする復号データを選択する際に、基本的にそれらを適切な符号を用いて合計することによって一緒に考慮する必要がある。

【0112】

オーディオビジュアル仮想現実コンテンツなどのビジュアルコンテンツに対する観察点と観察向きの選択を制御するために、例示的实施形態に従って方法、装置、およびコンピ

50

ユーザプログラム製品が提供される。本明細書において説明または想定する例示的な実施形態および実装の多くは、仮想現実コンテンツを含むがこれに限定されないオーディオビジュアルコンテンツが、視聴者にストリーミングされる状況において生じる。仮想現実コンテンツの開発、伝送、および視聴に関する技術的課題の1つは、視聴者の向きや位置などによって、視聴者が、仮想現実コンテンツの最も顕著な部分ではないかもしれない部分を視聴しがちであるということである。コンテンツクリエイタまたは作成者は一般的に、最も顕著なおよび/または特に興味深いとみなされるコンテンツを、ユーザの視野候補内の選択された場所に提示する。これらの選択された場所は、最も取りうる視聴方向 (Most Probable Viewing Direction : M P V D) とみなすことができる。これは、一般的に視聴者は、提示された顕著なおよび/または興味深いコンテンツにより興味を引かれがちであるため、そのコンテンツが見やすくなるように自身を位置付ける傾向があるからである。しかしながら、仮想現実プレゼンテーションやその他の没入的コンテンツプレゼンテーションでは、視聴者やヘッドマウント視聴デバイスの位置に基づいて視聴者の遠近および/または視聴向きを変更できるようにしているため、ユーザがそのようなコンテンツを再生し始めると、ユーザの位置および/または向きにより、最も顕著なおよび/または興味深いコンテンツが視聴者からずれてレンダリングされ、M P V D と位置を合わせるには、視聴者は移動するか、あまり快適ではない視聴位置を取らなければならなくなる可能性が高い。例えば、視聴者がM P V Dを見つけるために真後ろを見なければならぬ場合がある。別の例では、ソファまたは椅子上の視聴者の位置によっては、視聴者はM P V D と位置を合わせるために物理的に不快な姿勢を取らなければならぬ場合がある。

10

20

【 0 1 1 3 】

そのような一部の状況において、V R ビデオのストリーミングビットレートの削減に向けたストリーミングプロトコルの最近の傾向として、現在のビューの向きをカバーする360度ビデオコンテンツのサブセットを最高の品質/解像度で伝送し、その360度ビデオの残りの部分はより低い品質/解像度で伝送するというものがある。そのような一部の状況、および本発明の例示的な実施形態を実装可能な他の状況では、仮想現実コンテンツ用にH T T Pを介した動的適応ストリーミング (「D A S H」) を用いることが想定されている。

【 0 1 1 4 】

D A S Hのいくつかの実装では、同じアダプテーションセット内の複数のリプレゼンテーション間で、例えば、幅と高さ (それぞれ@widthおよび@heightとして参照されうる) 、フレームレート (@frameRateとして参照されうる) 、ビットレート (@bandwidthとして参照されうる) 、および/またはリプレゼンテーション間の標示された品質順 (@quality Rankingとして参照されうる) 、に基づいて自動選択を実行できる。D A S Hのいくつかの例示的な実装において、@qualityRankingの意味は、同じアダプテーションセット内の他のリプレゼンテーションに対してリプレゼンテーションの品質ランキングを指定するものとして指定される。一般的に、値が低いほど高い品質のコンテンツを表す。D A S Hの実装では、@qualityRanking属性が存在しない場合、ランキングは定義されない。

30

【 0 1 1 5 】

仮想現実ビデオコンテンツの状況において、360度コンテンツの一部のビューポートをより高い品質で表現し、他のビューポートをより低い品質で表現することができる。しかしながら、前述の属性のいずれも、異なる主要ビューポート用に符号化された360度ビデオを区別するには十分ではないことを理解されたい。

40

【 0 1 1 6 】

D A S Hにおけるビューポートベースの適応を促進するために、リプレゼンテーションの主要ビューポートを標示するメタデータをM P Dに含めることができる。また、主要ビューポートの画質に基づいてリプレゼンテーションを選択できるようにするために、全体的な品質特性とは別途に主要ビューポートの品質を標示する手段をM P Dに含めることができる。1つ以上のプロパティ記述子または要素を用いて、主要ビューポート、および/または主要ビューポートの品質を標示することができる。これらの例を以下の段落で説明

50

する。

【 0 1 1 7 】

一例において、VRビデオ記述子 (VR Video Descriptor : VRD) には、1) どのビューポートが (サブ) リプレゼンテーションに提示されるかを標示する、2) ビューポート固有の品質ランキングを標示する、という2つの目的がある。VRビデオ記述子によって伝達される情報は、コンテンツによって表現されるビューポート (単数または複数)、ビューポート (単数または複数) 用の投影フォーマット、ビューポート (単数または複数) 用のコンテンツが平面視か立体視かの標示、立体視のコンテンツの場合、左、右、または両方のビューが提示されるかどうか、ビューポート (単数または複数) の品質ランキング値 (単数または複数) である。ビューポート固有の品質ランキング情報により、クライアントは、同じビューポート (単数または複数) を異なる品質で表現する複数のリプレゼンテーションおよびサブリプレゼンテーションを区別することができる。VRDスキームを用いるSupplementalProperty要素またはEssentialProperty要素の@valueは、以下の表に指定されているVRDパラメータのカンマ区切りの値リストである。

EssentialProperty@valueまたはSupplementalProperty@valueパラメータ	使用	説明	
vr_source_id	M	十進数で表現される負ではない整数であり、コンテンツのソースに対する識別子を提供する。	
view_idc	M	0は、ビューポートが平面視であることを標示し、1は、ビューポートが立体視のコンテンツの左側ビューであることを標示し、2は、ビューポートが立体視のコンテンツの右側ビューであることを標示し、3は、ビューポートが左側ビューと右側ビューの両方を含むことを標示する。その他の値は予備である。	10
projection_format	M	[C I C P] に規定されているとおり、ビューポートの投影フォーマットを指定する。	
viewport_yaw	M	[C I C P] に規定されている十進数の浮動小数点の値、または文字列「NA」。NAは、ビューポートが、同じコンテナ要素内の他のVRDのSupplementalPropertyまたはEssentialProperty要素によってカバーされない残りの視野をカバーすることを標示する。	20
viewport_pitch	M	[C I C P] に規定されている十進数の浮動小数点の値、または文字列「NA」。viewport_yawがNAである場合、viewport_pitchもNAである。	
viewport_hor_fov	M	[C I C P] に規定されている十進数の浮動小数点の値、または文字列「NA」。viewport_yawがNAである場合、viewport_hor_fovもNAである。	30
viewport_ver_fov	M	[C I C P] に規定されている十進数の浮動小数点の値、または文字列「NA」。viewport_yawがNAである場合、viewport_ver_fovもNAである。	
quality_ranking	O	同じアダプテーションセット内の(@qualityRankingが関連付けられている)他のリプレゼンテーション、および任意のアダプテーションセット内の同じvr_source_id値を有するビューポート、に対するビューポートの品質ランキングを指定する。値が低いほど高い品質のコンテンツを表す。存在しない場合、ランキングは定義されない。	40
凡例： M=必須、O=任意			

【 0 1 1 8 】

－実施形態において、（例えばISOベースメディアファイルフォーマット対応の）ファイルおよび/またはMPDのVR固有の記述子内の投影およびマッピングメタデータは

、 i) 投影されたフレームのVR投影フォーマット、 ii) グローバル座標系における投影されたフレームに対応する幾何学的構造の向き、 iii) 領域ごとのマッピング情報、 iv) 領域ごとの品質ランキング、 の1つ以上を含む。

【 0 1 1 9 】

一実施形態において、仮想現実ビデオ記述子 (VRD) は以下のように指定される。VRDスキームは、SupplementalPropertyおよび/またはEssentialProperty記述子を特定の@schemeldUri値と共に用いる。EssentialProperty記述子は、投影に対応した表示処理なしで復号ビデオコンテンツを従来の2次元ディスプレイに表示することが望ましくない場合に用いるべきである。VRビデオのSupplementalPropertyまたはEssentialProperty記述子は、AdaptationSet、Representation、またはSubRepresentationに存在することができる。VRDスキームを用いるSupplementalProperty要素またはEssentialProperty要素の@valueは、以下の表に指定されているVRDパラメータのカンマ区切りの値リストである。

EssentialProperty@valueまたはSupplementalProperty@valueパラメータ	使用	説明
vr_source_id	M	コンテンツのソースの識別子を指定する。同じvr_source_id値に関連付けられたすべてのリプレゼンテーションの座標系は同一である。
projection_format	M	VR投影フォーマット標示子。投影されたフレームの投影フォーマットを指定する。
global_yaw	M	グローバル座標系に対して投影のヨーを角度で指定する十進数の浮動小数点の値。
global_pitch	M	グローバル座標系に対して投影のピッチを角度で指定する十進数の浮動小数点の値。
global_roll	M	グローバル座標系に対して投影のロールを角度で指定する十進数の浮動小数点の値。
凡例： M=必須、O=任意		

【 0 1 2 0 】

一実施形態において、SRD記述子は以下のように拡張される。SRD記述子はリプレゼンテーションレベルに含めることもできる。同じRepresentation要素およびSubRepresentation要素内に多数のSRD記述子を含めることができる。同じコンテナ要素内の複数のSRD記述子は、例えば、投影されたフレーム内の複数の領域を標示するためにSRD記述子が使用され、そのうち少なくとも一部の領域が他の領域とは異なる品質ランキングを有することを標示する場合に役立つ。SRD記述子の構文と意味は、前述と同様に記述してもよい。しかしながら、object_x、object_y、object_width、およびobject_heightは、同じコンテナ要素内にそれらの値を有する別のSRD記述子がある場合は、任意として定義してもよい。object_x、object_y、object_width、およびobject_heightが存在しない場合、対応する領域は、同じレベル内の他の指定された領域を除いて投影されたフレームとして定義される。quality_rankingパラメータは以下のようにSRD内に、例えば最後のパラメータとして、定義してもよい。

quality_ranking	○	同じアダプテーションセット内の (@qualityRankingが関連付けられている) 他のリプレゼンテーション、および任意のアダプテーションセット内の同じvr_source_id値を有する (サブ) リプレゼンテーション、に対する領域の品質ランキングを指定する。値が低いほど高い品質のコンテンツを表す。存在しない場合、ランキングは定義されない。
-----------------	---	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

10

【 0 1 2 1 】

本発明の例示の実施形態のいくつかの実装は、DASHイベントを含む環境において想定される、および/または生じる。このイベントは、参照することによって本明細書に組み込まれるISO/IEC 23009-1において説明およびその他の手段で提示されているイベントを含むが、これに限定されない。

【 0 1 2 2 】

DASHイベントは、非周期的な情報をDASHクライアントまたはアプリケーションにシグナリングするために、メディアプレゼンテーション記述(MPD)またはリプレゼンテーション内に提供してもよいことを理解されたい。各イベントは特定のメディアプレゼンテーション時間に開始され、一般的に持続時間を有するという意味で、イベントは時間的である。イベントには、DASH固有のシグナリングまたはアプリケーション固有のイベントが含まれる。後者の場合、DASHクライアントがイベントを適切なアプリケーションに転送することができるように、適切なスキーム識別子によってアプリケーションが識別される。

20

【 0 1 2 3 】

いくつかの実装では、同じタイプのイベントはイベントストリームにおいてクラスタ化される。これによって、DASHクライアントは関心のあるイベントストリームをサブスクライブし、関係または関心のないイベントストリームは無視することができる。MPD内でのイベントシグナリングとセグメント内でのインバンドのイベントシグナリングという2つのイベントシグナリング方法が指定されていることを理解されたい。メディアプレゼンテーション時間に割り当てられたイベントのシーケンスは、MPD内のピリオドレベルで提供してもよい。同じタイプのイベントは、Period要素内のEventStream要素によって指定される1つのイベントストリームにまとめられる。一般的に、イベントは、その開始時間がピリオド境界の後、またはその持続時間がピリオド境界を越えていても、そのピリオドの終了時に終了する。

30

【 0 1 2 4 】

DASHに基づくほとんどの状況では、EventStream要素は、スキームを識別するためのURIを提供する@schemeIdUri属性と、任意の属性@valueとを含むという意味で、DASHのプロパティ記述子と同様に構築される。要素の意味は、採用するスキーム固有である。スキームを識別するURIは、URN (Uniform Resource Name) またはURL (Uniform Resource Locator) でありうる。

40

【 0 1 2 5 】

同様に、DASHに基づくほとんどの状況では、1つのピリオドは、同じ@schemeIdUri属性値と@value属性値を有する最大1つのEventStream要素を含む。例えば、1つのタイプのすべてのイベントは1つのイベントストリームにクラスタ化されうる。イベントストリームは時限イベントを含むため、イベントをピリオド内の特定のメディアプレゼンテーション時間に割り当てるために時間尺度属性@timescaleも提供されることを理解されたい。時限イベント自体はEvent要素によって記述される。

【 0 1 2 6 】

DASHに精通している方は、EventStream要素内の属性の規定された意味を認識され

50

るであろう。これらを以下に挙げる。

要素名または属性名	使用	説明
EventStream		イベントストリームを指定する。
@xlink:href	O	外部の EventStream 要素への参照を指定する。
@xlink:actuate	OD デフォルト： onRequest	処理命令を指定する。「onLoad」または「onRequest」のいずれかを指定できる。 この属性は、@xlink:href属性が存在しない場合は存在しない。
@schemeIdUri	M	メッセージスキームを識別する。文字列にはURN構文またはURL構文を用いることができる。URLを用いる場合、mm y y yの形式で月日も含めることを推奨する。URLの割当ては、そのURLのドメイン名の所有者によってその日に、またはそれに極めて近い日に許可されていなければならない。URLはインターネット上の場所に変換されてもよく、変換される場所にメッセージスキームの仕様を格納してもよい。
@value	O	EventStream要素の値を指定する。値空間と意味は、@schemeIdUri属性において識別されるスキームの所有者が定義しなければならない。
@timescale	O	Event 要素に異なるリアルタイム持続時間値を提供するために用いる、時間尺度を秒単位で指定する。どのレベルにも存在しない場合、1に設定される。
Event	0...N	1つのイベントを指定する。 イベントストリーム内の複数のイベントは、プレゼンテーション時間が降順にならないように順序付けされる。
<p>凡例： 属性：M=必須、O=任意、OD=任意かつデフォルト値あり、CM=条件付きで必須 要素：<minOccurs>...<maxOccurs> (N=無制限) 要素は太字であり、属性は太字なしで先頭に@が付いている。</p>		

【 0 1 2 7 】

DASHに精通している方は、Event要素内の属性の規定された意味を認識されるであろう。これらを以下に挙げる。

要素名または属性名	使用	説明	
Event		イベントを指定し、文字列としてフォーマットされたイベントのメッセージを含む。この要素のコンテンツは、イベントスキームに応じて異なる。	
@presentationTime	OD デフォルト：0	ピリオドの開始時に対するイベントのプレゼンテーション時間を指定する。 プレゼンテーション時間の値（秒単位）は、この属性の値と@timescale属性の値を除算したものである。 存在しない場合、プレゼンテーション時間の値は0である。	10
@duration	O	イベントのプレゼンテーション持続時間を指定する。 持続時間の値（秒単位）は、この属性の値と@timescale属性の値を除算したものである。 存在しない場合、持続時間の値は不明である。	
@id	O	イベントのこのインスタンスの識別子を指定する。 Event 要素に同等のコンテンツおよび属性値を有する複数のイベントは、この属性に同じ値を有するものとする。 各イベントの@idは、同じ@schemeIdUriと@valueのペアと共に適用される。	20
@messageData	O	EventStream要素の値を指定する。値空間と意味は、@schemeIdUri属性において識別されるスキームの所有者が定義しなければならない。 注記：この属性は、イベントに完全なXML要素（単数または複数）を指定することの代替手段である。イベントが簡潔な文字表現をとる傾向がある場合は有益である。	30
凡例： 属性：M=必須、O=任意、OD=任意かつデフォルト値あり、CM=条件付きで必須 要素：<minOccurs>...<maxOccurs>（N=無制限） 要素は 太字 であり、属性は太字なしで先頭に@が付いている。			40

【 0 1 2 8 】

DASHに精通している方は、セグメントの一部としてイベントメッセージを追加することで、リプレゼンテーションによってイベントストリームを多重化できることを認識されるであろう。イベントストリームは、複数の選択されたリプレゼンテーション、1つの（または複数の）選択されたアダプテーションセット、またはすべてのリプレゼンテーションに存在することができる。例えば、可能な構成の1つとして、オーディオアダプテーションセットだけがインバンドイベントを含みうる構成がある。2つ以上のリプレゼンテーションが同じ@schemeIdUriおよび同じ@valueを有するイベントストリームを実行する場

合、それらのストリームは意味的に等価であり、1つのリプレゼンテーションを処理するだけでよい。

【0129】

DASH環境では、リプレゼンテーションに存在するインバンドイベントストリームは、アダプテーションセットレベルまたはリプレゼンテーションレベルのInbandEventStream要素によって標示される。InbandEventStream要素の構文および意味は、前述のEventStream要素の構文および意味と同じであってもよい。1つのリプレゼンテーションは、それぞれ個別のInbandEventStream要素によって標示される複数のインバンドイベントストリームを含んでもよい。

【0130】

DASHでは、イベントメッセージボックス(「emsg」)が、メディアプレゼンテーション時間に関連する一般的なイベントへのシグナリングを提供する。前述の、MPDに定義されるイベントに関する同じ意味が適用される。イベントメッセージボックスのフィールドの意味は、Event要素の各属性の意味に類似している。ISOベースメディアファイルフォーマット(ISOBMFF)にカプセル化される場合、メディアセグメントは1つ以上のイベントメッセージ(「emsg」)ボックスを含むことができる。存在する場合、「emsg」ボックスは「moof」ボックスの前に配置される。イベントメッセージボックスの構文は以下のように指定できることを理解されたい。

```
aligned(8) class DASHEventMessageBox extends FullBox('emsg', version = 0, flags = 0){
```

```
    string          scheme_id_uri;
    string          value;
    unsigned int(32) timescale;
    unsigned int(32) presentation_time_delta;
    unsigned int(32) event_duration;
    unsigned int(32) id;
    unsigned int(8)  message_data[];
}
```

```
}
```

【0131】

本発明の例示的实施形態が実装される正確なプロトコルおよび/または環境にかかわらず、いくつかの技術的課題は、視聴者および/またはコンテンツクリエイタの期待および/または要件を満たす品質レベルでコンテンツを提示する必要性、および視聴者および/または視聴者のディスプレイと視聴者に提示されるコンテンツの向きが合っていない可能性から生じる。

【0132】

VRコンテンツに対する初期観察向きまたは初期ビューポートをシグナリングする能力は、VR再生セッションの開始時にコンテンツ作成者の好みは反映されるようにする、望ましい機能である。初期観察向きまたは初期ビューポートは、VRプレゼンテーションの開始時だけでなく、任意のランダムアクセスポイントまたはVRプレゼンテーション内の任意のポイントにも割り当てることができる。しかしながら、初期観察向きまたは初期ビューポートのシグナリングの「強さ」を標示する能力が求められている。例えば、視聴者がコンテンツの一部を以前に見ており、そのコンテンツの他の部分をシークする場合、(1)シーク後のコンテンツ再生をユーザの頭部の向き(より一般的には、ユーザが以前にコンテンツを見ていた向き)を用いて継続するか、(2)シグナリングされた初期視聴向きまたは初期ビューポートを適用するか、をコンテンツクリエイタ/作成者が制御できるようにする必要がある。

【0133】

前者は、例えば、コンテンツが固定のVRカメラまたはカメラリグを用いて実際のシーンカットなしで生成された場合、または以前の視聴位置とシーク位置との間にシーンカッ

10

20

30

40

50

トがない場合に、有効に用いることができる。後者は、例えば、カメラの位置が変更されたか、以前の視聴位置とシーク位置との間にシーンカットがある場合、または同じビデオのコンテンツが以前に視聴されていない場合に、有効に用いることができる。また、プレゼンテーション内の一部のポイント（例えばシーンカットなど）は、ヘッドマウントディスプレイの以前の向きに関係なく、コンテンツ作成者の希望どおりに観察向きを選択してもよい。

【0134】

さらに、DASHでは、DASHクライアントが正しいアダプテーションセットおよびリプレゼンテーションから（サブ）セグメントを要求できるように、クライアントに対して初期観察設定をシグナリングできるようにするべきである。アダプテーションセット内の各リプレゼンテーションが同じビューポートを含み、このビューポートが品質の異なる複数の構成ビューポートを含む場合、DASHクライアントはシグナリングによって、初期観察向きまたは初期ビューポートに正確にまたはほぼ一致する高い品質の構成ビューポートを含むリプレゼンテーションを選択することができる。アダプテーションセット内の各リプレゼンテーションが、比較的狭く、典型的に構成ビューポートを含まない同じビューポートを含み、複数のアダプテーションセットそれぞれが同じ全方位コンテンツの異なるビューポートをカバーする場合、DASHクライアントはシグナリングによって、初期観察向きまたは初期ビューポートに正確にまたはほぼ一致するアダプテーションセットを選択し、次にそのアダプテーションセットから高品質のリプレゼンテーションを選択することができる。

【0135】

現在の技術はこのような技術的ニーズに対応していないと考えられる。例えば、（参照によって本明細書に組み込まれる）MPEG M38689では、DASH MPDにおける初期ビューポートのシグナリングを以下のように説明している。

『最初の視点を設定し、すべての角度位置の計算の基準となる原点軸を定義するために、中心点を配置する新しい追加プロパティをアダプテーションセットに導入する。中心点の位置は、中心点が配置されるグリッドセルの左上角から画素単位で指定される。

このような追加プロパティは、urn（例：urn:mpeg:dash:vrorigin:2016）、およびx-y座標を画素単位で含む値（例："640, 360"）によって定義される。

あるいは、VR原点の追加プロパティは、空間オブジェクト全体に関連付けられたアダプテーションセットに設定できることに注意されたい。ただしその場合、MPDオーサリング時に、より多くの計算が必要となる（VR原点と空間オブジェクト全体の左上角との間のすべてのセルの幅と奥行きを合計する必要があるため）。』

【0136】

MPD記述子は、経時的または動的エンティティではないという意味で静的であることを理解されたい。したがって、M38689では、初期ビューポートを例えば時間の関数としてまたはSAPに基づいて標示できない。このため、M38689はDASHクライアントが適切なアダプテーションセットおよび/またはリプレゼンテーションを選択する助けにならない。

【0137】

図1は、本発明の一例示的实施形態による実装を実行しうる例示的システム環境100を示す図である。環境100の描写は、本明細書において説明または想定する実施形態を、いかなる特定の要素またはシステムの構成にも限定または制限することを意図するものではなく、本発明の実施形態に関連して用いることができる一連の構成またはシステムに対する代替の構成またはシステムを除外することを意図するものでもない。むしろ、図1およびそれに開示する環境100は、単に、本明細書において開示および想定する方法、装置、およびコンピュータプログラム製品の特徴、態様、および使用のいくつかを促進す

るための例示的な基礎または状況として提示するものである。図1に示す態様およびコンポーネントの多くは、個々の異なる要素として示されているが、本明細書において説明する方法、装置、およびコンピュータプログラムと関連して他の構成を用いてもよく、これには態様および/またはコンポーネントを組み合わせる、省略する、および/または追加する構成も含まれることを理解されたい。

【0138】

図1に示すように、システム環境100は少なくとも1つのカメラ102を含む。システム環境100の多くの実装では、仮想現実コンテンツの作成に用いる360°ビデオ画像の撮影に適切な1つ以上のカメラ(NokiaのOZOシステムなど)、および/または360°ビデオ画像および/または他のパノラマビューの作成に用いることができる他のカメラまたはカメラ配列を用いることが想定されている。図1では、1つ以上のメディアソースメディアソース104があることも想定されている。このメディアソースは、以前に撮影または生成されたオーディオビジュアルコンテンツの伝送および/またはアクセスのためのデータベース、他のデバイスおよび/または他のシステムであってもよい。

10

【0139】

図1に示すように、カメラ102およびメディアソース104は、画像および/または他のオーディオビジュアルコンテンツ(360°ビデオ画像など)をデータストリームとして伝送できるおよび/または伝送するように構成されている。そのような伝送は、カメラから1つ以上のデバイスに画像データを伝送するのに適切な任意の手法および/またはプロトコルに従って実行できる。いくつかの実装では、画像データの伝送は、ビデオ画像を受信および/または処理するように構成された1つ以上のデバイスへと、無線または有線接続を介してリアルタイムまたはほぼリアルタイムで行われる。

20

【0140】

本明細書のいくつかの例示的実装では、360°画像内の点や領域などの顕著点または顕著領域を想定している。これらは、注目が向けられる、画像内の最も顕著な点または領域とみなされる。本明細書のいくつかの例示的実装では、画像内に1つ以上の関心点または関心領域があると想定している。これは、コンテンツクリエイターおよび/または1人以上の視聴者が関心を持ちうる画像要素とみなされる。多くの状況において、画像の顕著点は関心点となり、画像の顕著領域は関心領域となる。また、画像の顕著点または顕著領域は変わってもよく、変更されてもよい。例えば、システムまたはシステム要素および/または外部アクター(ディレクタなど)によって自動的に変更されてもよい。そのようないくつかの状況では、顕著点はある関心点から別の関心点へと、また顕著領域はある関心領域から別の関心領域へと、切り替えられてもよい。以下において、顕著点という用語に言及して実施形態を説明しているが、そのような例示的実施形態および他の例示的実施形態は、顕著点の代わりに顕著領域を用いる場合にも同様に当てはめることができることを理解されたい。

30

【0141】

図1に示すように、カメラ102およびメディアソース104はそれぞれビデオ画像ストリームをビデオプロセッサ106に伝送してもよい。ビデオプロセッサ106は、スタンドアロンデバイスとして、および/または他のデバイスやコンポーネントに統合されるデバイスとして実装される任意のデバイス類の代表的なものである。図1に示すように、ビデオプロセッサ106は、カメラ102およびメディアソース104のそれぞれから画像データストリームおよび任意の関連情報を受信するように構成される。いくつかの例示的実装において、ビデオプロセッサ106は、ビデオストリーム内の1つ以上の顕著点を選択および/または識別できるようにも構成される。いくつかの例示的実施形態において、ビデオプロセッサ106は顕著点を標示する情報を、当該ビデオストリームに、または当該ビデオストリームに関連付けられた別のストリーム(または、MPDなどのシグナリング構造)に埋め込む。いくつかの例示的実施形態において、ビデオプロセッサ106は、その顕著点を、再生デバイスの意図された動作に関連付けられた標示とみなし、その再生デバイスの意図された動作を特定し、その特定に応じて制御信号を生成させる。こ

40

50

の制御信号は、再生デバイス上でのオーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。前述の制御信号は、例えばビデオストリーム、またはビデオストリームの記述内に含まれてもよい。

【0142】

ディレクタ108はビデオプロセッサ106の任意の操作者として示されており、いくつかの実装では、画像データストリームの作成および/またはストリーミング中に1つ以上の画像データストリームを監視および/または制御することができる。いくつかの例示の実施形態において、ディレクタ108は、顕著点を標示する情報をビデオストリーム内の特定の場所に埋め込ませる。いくつかの例示の実施形態において、ディレクタ108は、再生デバイスの意図された動作を特定し、制御信号を生成させる。この制御信号は、再生デバイス上でのオーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。前述の制御信号は、例えばビデオストリーム内、またはビデオストリームの記述内に含まれてもよい。ディレクタ108は、加えてまたは代わりに、ビデオストリームに提示されたコンテンツ、ならびにその作品内の主題、背景要素、およびその他のオブジェクトの相対的な配置に関する創作的な決定を行ってもよい。前述のとおり、ディレクタ108は環境100において任意であり、ビデオプロセッサ106、他のなんらかのデバイスの動作、またはその他の手段で、ディレクタや他のエンティティの存在や動作なしで、1つ以上の顕著点がビデオストリームに埋め込まれる実装が可能である。

10

【0143】

図1に図示するように、ビデオプロセッサ106はオーディオビジュアルコンテンツをネットワーク110を介して送信する。実際の送信装置はビデオプロセッサエンティティとは異なるエンティティであってもよいが、これらのエンティティは動作可能に接続されているため、単一のビデオプロセッサ106として図示していることを理解されたい。送信装置は、いくつかの実施形態において、例えばHTTPサーバ(ウェブサーバなど)であってもよい。ネットワーク110は、360°ビデオおよび関連する向き情報を、ビデオプロセッサ106などの1つ以上のデバイスから、仮想現実ヘッドセット114などの視聴デバイスへと直接的および/または間接的に伝送するのに適切な任意のネットワークであってもよい。視聴デバイスは図1において単一の装置として図示されているが、視聴デバイスは通常は動作可能に接続された複数のデバイスを含んでもよいことを理解されたい。例えば、仮想現実ヘッドセットは、ネットワーク110を介してオーディオビジュアルコンテンツを受信するコンピュータに接続されてもよい。別の例において、仮想現実ヘッドセットはその表示デバイスとして、ヘッドセットに取り付けられてネットワーク110を介してオーディオビジュアルコンテンツを受信するスマートフォンを用いる。いくつかの実装では、ネットワーク110は公衆インターネットを含むおよび/または組み込まれている。

20

30

【0144】

図1には、仮想現実ヘッドセット114などの視聴デバイスに関連付けられたユーザ112も図示している。一般的に、仮想現実ヘッドセット114は、1つ以上の360°画像データストリームなどの1つ以上のデータストリームを(対応する向き情報と共に)受信し、ユーザ112に表示可能な可視画像をレンダリングすることができる。いくつかの実装では、仮想現実ヘッドセット114は、ユーザ112が頭部を回した角度(度)などの、ユーザ112に関する位置情報、およびユーザ112またはその頭部の動きに関する他の情報を確認することもできる。図1ではユーザ112が仮想現実ヘッドセット114を介して視聴コンテンツを見ているように描いているが、ユーザは、自身に対して伝送されたビデオのすべてまたは一部を表示するように構成された任意の視聴システムを介してコンテンツを視聴してもよい。例えば、ユーザは1つ以上のモニタ、モバイルデバイス、および/またはその他の携帯型ディスプレイやデスクトップディスプレイを用いてコンテンツを視聴してもよい。ディスプレイが、任意の一時点で360°コンテンツの一部を表示するように構成されている場合、ユーザ112はそのコンテンツのどの部分を表示するか制御できてもよい。例えば、ユーザ112は、例えばキーボード、ジョイスティック、

40

50

マウス、またはその他に入力機器を用いて、あるいはスマートフォンなどの表示デバイスを回転または旋回させることで、視聴方向を制御できてもよい。

【0145】

一実施形態において、ユーザによるVRビデオクリップの視聴動作の統計が収集される。例えば、あるプレーヤーは、視聴方向または視聴向き（例えば、クリップの再生開始時の仮想現実ヘッドセット114の最初の向きに対するその向き）をクリップのメディア時間の関数として、統計を収集するサーバに報告する。報告された視聴方向を収集することで、クリップのメディア時間の関数として、最も取りうる視聴方向(MPVD)が形成されてもよい。MPVDは、統計的にユーザに対して最もレンダリングされがちな方向または領域を標示すると理解されてもよい。MPVDは、創作的な決定を支援するためのイン
10 プットとしてディレクタ108に提供されてもよい。あるいは、MPVDは、ビデオストリーム内の特定の場所に埋め込まれる顕著点として、ビデオプロセッサ106によって用いられてもよい。あるいは、ビデオプロセッサ106は、そのMPVDを再生デバイスの意図された動作に関連付けられた標示とみなし、その再生デバイスの意図された動作を特定し、その特定に応じて制御信号を生成させる。この制御信号は、再生デバイス上でのオーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。この実施形態により、最初の一連のユーザの視聴動作によって顕著点の選択を支援または決定
15 できるようになり、後続のユーザの視聴体験が改善される。

【0146】

オーディオビジュアルコンテンツの一部に関連付けられた初期観察設定、およびその初期観察設定の一式の条件に少なくとも部分的に基づいて、ユーザの位置、およびコンテンツを視聴者にレンダリングする際にコンテンツクリエイターが行った創作的選択を考慮して、オーディオビジュアルコンテンツをレンダリングできる。この点において、オーディオ
20 ビジュアルコンテンツの観察点および観察向きの選択は、図2に示す装置200によって制御できる。この装置は、カメラ102、メディアソース104、または図1に関連して説明した、ビデオプロセッサ106、および/またはネットワーク110に組み込まれうるまたはその他の手段で関連付けられうるデバイスなどの他の任意のデバイスのいずれかによって、具体化してもよい。あるいは、装置20は、そのようなデバイスの外部にある別のコンピューティングデバイスによって具体化されてもよい。例えば、この装置は、パーソナルコンピュータ、コンピュータワークステーション、サーバなどによって、または
30 、携帯端末（例えば、スマートフォン、タブレットコンピュータ、ビデオゲームプレーヤー）などの様々なモバイルコンピューティングデバイスのいずれかによって具体化されてもよい。あるいは、この装置は、仮想現実ヘッドセット114などのヘッドマウントディスプレイといった仮想現実システムによって具体化されてもよい。

【0147】

装置200が具体化される方法にかかわらず、一例示的实施形態の装置は、プロセッサ202とメモリデバイス204、および任意でユーザインタフェース206および/または通信インタフェース208を備える、またはそれらと通信するように構成される。いくつかの実施形態において、このプロセッサ（および/またはコプロセッサ、あるいはこの
40 プロセッサを支援する、またはその他の方法でそれに関連付けられた処理回路構成）は、この装置のコンポーネント間で情報を渡すためのバスを介してメモリデバイスと通信してもよい。メモリデバイスは非一時的であってもよく、例えば、1つ以上の揮発性および/または不揮発性メモリを含んでもよい。換言すると、例えば、このメモリデバイスは、マシン（例えば、プロセッサなどのコンピューティングデバイス）によって取得されうるデータ（例えば、ビット）を記憶するように構成されたゲートを備える電子記憶デバイス（例えば、コンピュータ可読記憶媒体）であってもよい。このメモリデバイスは、この装置が本発明の一例示的实施形態による様々な機能を実行できるようにするための情報、データ、コンテンツ、アプリケーション、命令などを記憶するように構成されてもよい。例えば、このメモリデバイスは、プロセッサが処理する入力データをバッファするように構成
50 することができる。加えてまたは代わりに、このメモリデバイスはプロセッサが実行する

命令を記憶するように構成することができる。

【0148】

前述のように、装置200はコンピューティングデバイスによって具体化されてもよい。しかしながら、いくつかの実施形態では、この装置はチップまたはチップセットとして具体化されてもよい。換言すると、この装置は構造アセンブリ（例えば、ベースボード）上の材料、コンポーネント、および/またはワイヤを含む1つ以上の物理パッケージ（例えば、チップ）を含んでもよい。この構造アセンブリは、物理的強度、サイズの維持、および/またはそれに含まれるコンポーネント回路構成に対する電氣的相互作用の制限を提供してもよい。したがってこの装置は、場合によっては、本発明の実施形態を単一のチップ上に実装、または単一の「システムオンチップ（system on a chip）」として実装する

10

【0149】

プロセッサ202は多数の異なる方法で具体化することができる。例えば、このプロセッサは、コプロセッサ、マイクロプロセッサ、コントローラ、デジタル信号プロセッサ（Digital Signal Processor：DSP）、DSPが付属するまたは付属しない処理要素などの様々なハードウェア処理手段、あるいは、例えば、特定用途向け集積回路（Application Specific Integrated Circuit：ASIC）、フィールドプログラマブルゲートアレイ（Field-Programmable Gate Array：FPGA）、マイクロコントローラユニット（Micro Controller Unit：MCU）、ハードウェアアクセラレータ、専用コンピュータチップなどの集積回路を含むその他様々な処理回路構成、の1つ以上として具体化されてもよい。したがって、いくつかの実施形態において、このプロセッサは個別に動作するように構成された1つ以上の処理コアを含んでもよい。マルチコアプロセッサにより、単一の物理パッケージ内で多重処理を可能にしてもよい。加えてまたは代わりに、このプロセッサは、命令、パイプライン処理、および/またはマルチスレッド処理を個別に実行できるように、バスを介して連携するように構成された1つ以上のプロセッサを含んでもよい。

20

【0150】

一例示の実施形態において、プロセッサ202は、メモリデバイス204に記憶されている命令、または別の方法でプロセッサにアクセス可能な命令を実行するように構成されてもよい。代わりにまたは加えて、プロセッサは、ハードコードされた機能を実行するように構成されてもよい。したがって、ハードウェアまたはソフトウェアの方法による構成か、それらの組合せによる構成かにかかわらず、プロセッサは、本発明の一実施形態に従って構成され、それによる動作を実行可能な（例えば、回路構成において物理的に具体化される）エンティティを表してもよい。したがって、例えば、ASIC、FPGAなどとして具体化される場合、このプロセッサは、本明細書で説明する動作の実行専用構成されたハードウェアであってもよい。あるいは、別の例として、このプロセッサがソフトウェア命令の実行部として具体化される場合、これらの命令が実行されると、本明細書で説明したアルゴリズムおよび/または動作を実行するようにプロセッサが専用構成されてもよい。ただし場合によっては、プロセッサは、本明細書で説明したアルゴリズムおよび/または動作を実行するように命令によって当該プロセッサがさらに構成されることによ

30

40

【0151】

いくつかの実施形態において、装置200は任意でユーザインタフェース206を備えてもよく、ユーザインタフェース206はプロセッサ202と通信し、出力をユーザに提供、およびいくつかの実施形態ではユーザ入力の標示を受信してもよい。したがって、ユーザインタフェースはディスプレイを備えてもよく、いくつかの実施形態では、キーボード、マウス、ジョイスティック、タッチスクリーン、タッチエリア、ソフトキー、マイク

50

ロフォン、スピーカ、またはその他の入出力機構を備えてもよい。代わりにまたは加えて、プロセッサは、ディスプレイなど、およびいくつかの実施形態ではスピーカ、リンガー、マイクロフォンなどの、1つ以上のユーザインタフェース要素の少なくとも一部の機能を制御するように構成される、ユーザインタフェース回路構成を備えてもよい。プロセッサ、および/またはプロセッサを含むユーザインタフェース回路構成は、プロセッサにアクセス可能なメモリ（例えば、メモリデバイス204など）に記憶されたコンピュータプログラム命令（例えば、ソフトウェアおよび/またはファームウェア）を介して、1つ以上のユーザインタフェース要素の1つ以上の機能を制御するように構成されてもよい。

【0152】

装置200は、任意で通信インタフェース208を備えてもよい。この通信インタフェースは、ネットワークと、および/または装置と通信する他の任意のデバイスまたはモジュールとデータを送受信するように構成された、ハードウェア、またはハードウェアとソフトウェアの組合せのいずれかとして具体化される、デバイスや回路構成などの任意の手段であってもよい。この点において、通信インタフェースは、例えば、無線通信ネットワークとの通信を可能にする、アンテナ（単数または複数）と支援ハードウェアおよび/またはソフトウェアを含んでもよい。加えてまたは代わりに、通信インタフェースは、アンテナ（単数または複数）を介して信号を伝送するため、またはアンテナ（単数または複数）を介した信号の受信を処理するために、アンテナ（単数または複数）と相互作用する回路構成を備えてもよい。いくつかの環境では、通信インタフェースは、代わりにまたはさらに有線通信に対応する。したがって、例えば通信インタフェースは、通信モデム、および/または、ケーブル、デジタル加入者回線（Digital Subscriber Line: DSL）、ユニバーサルシリアルバス（Universal Serial Bus: USB）、またはその他の機構を介して通信を支援するその他のハードウェア/ソフトウェアを含んでもよい。

【0153】

図3は、本発明の一実施形態の例示的実装を示しうるビデオプレゼンテーション300を描いている。図3に示すように、ビデオプレゼンテーション300は2つの異なるシーン316および318からのコンテンツを結合して形成されるVRビデオプレゼンテーションである。1つのシーン内の複数のピクチャは、通常は同様のコンテンツを有し、一般的に同じカメラによって連続的に撮影される。2つの異なるシーンのピクチャは、一般的に2つの異なるカメラで撮影されるか、同じカメラで時間的に不連続的に撮影される。図3に示すように、ビデオプレゼンテーション300内にシーンカット320がある。例示的ビデオプレゼンテーション300に示すように、ビデオプレゼンテーション300は、伝送のために（サブ）セグメント302~314に分割されている。ただし、実施形態はセグメント化されていないファイルにも同様に当てはまる。図3に示す例では、各（サブ）セグメント302~314はランダムアクセスポイント（DASHおよびISOBMFのSAPなど）によって開始される。各SAPでは、連続再生時、および各（サブ）セグメント302~314へのランダムアクセス時の意図されるプレーヤー動作の標示と共に、初期観察設定が（例えば、ファイルまたは転送エンキャプスレータによって）標示される。

【0154】

図3に示すように、最初の4つの（サブ）セグメント302、304、306、および308は、同じシーンまたはカメラからのものである。これらの（サブ）セグメント302~308に対し、例えば、初期観察設定はランダムアクセス時に条件付きで適用され、連続再生時には適用されないことが標示されてもよい。このような標示により以下のような結果となる。

【0155】

(1) 以前に他の（サブ）セグメントが処理されておらず、4つの（サブ）セグメント302~308のいずれかから処理が開始される場合、その初期観察設定がレンダリングの開始時に用いられる。換言すると、再生開始時のヘッドマウントディスプレイの向きに関係なく初期観察設定が用いられる。(2) 4つの（サブ）セグメント302~308が

10

20

30

40

50

、観察設定がリセットされない範囲内にあることが標示されてもよい。このような標示は一般的に、例えば、固定カメラによって撮影されたコンテンツに適切である。これら4つの(サブ)セグメント302~308のうち1つ目の(サブ)セグメントが以前に処理されて少なくとも部分的に表示され、その後2つ目の(サブ)セグメントがランダムアクセスされて表示された場合、2つ目の(サブ)セグメントの初期観察設定は用いられず、1つ目の(サブ)セグメントの初期観察設定によるヘッドマウントディスプレイの向きがレンダリングに用いられる。

【0156】

図3に示すように、5つ目の(サブ)セグメント310は、それより前の(サブ)セグメント302~308とは異なるシーンまたはカメラからのビデオコンテンツを含んでいる。5つ目の(サブ)セグメント310の初期観察設定は、連続再生とランダムアクセス時の両方に無条件に適用される。換言すると、5つ目の(サブ)セグメント310が連続再生時に(つまり、例えば4つ目の(サブ)セグメント308の処理が完了した後に)アクセスされたか、ランダムアクセス後にアクセスされたかに関係なく、視聴者には常に同じ観察向きのコンテンツが表示される。

【0157】

図3に示すように、6つ目の(サブ)セグメント312および7つ目の(サブ)セグメント314は、5つ目の(サブ)セグメント310と同じシーンまたはカメラからのものである。6つ目の(サブ)セグメント312および7つ目の(サブ)セグメント314に対し、例えば、初期観察設定はランダムアクセス時に無条件で適用され、連続再生時には適用されないことが標示されてもよい。そのような標示は一般的に、以下のようなコンテンツに適切である。例えば、カメラが動いているため、例えばシーク後に關心領域を見失う場合があるので、コンテンツクリエイタの意見では、以前の(サブ)セグメントの初期観察設定による観察向きを維持することが望ましくないコンテンツである。

【0158】

いくつかの例において、グローバル向きオフセット(単数または複数)は初期観察向き情報内に含まれているため、要求するアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションを選択する際に別途考慮する必要はない。いくつかの例において、グローバル向きオフセット(単数または複数)は初期観察向き情報とは個別であるため、要求するアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションを選択する際に一緒に考慮する必要がある。

【0159】

いくつかの例示的実装では、リセット範囲または維持範囲、あるいはその両方も想定する。そのようないくつかの例示的実装では、コンテンツクリエイタは、ユーザが、例えば(1)常に特定の關心領域を見ること、および(2)頭部の動きに正確に従い空間的および時間的に連続した体験を得ること、を望むことがある。ユーザが關心領域がある場所とは完全に反対の向きを見る場合など、状況によってはこれら2つの目的は矛盾することがある。両方の機能を得るために、初期観察設定にリセット条件および/または維持条件を加えてもよい。リセット範囲は、現在の視聴方向がその中にいる場合に観察設定をリセットさせる一連のビューポートとして定義されてもよい。維持範囲は、現在の視聴方向がその中にいる場合に、観察設定をリセットせずに現在の視聴方向を維持する一連のビューポートとして定義される。

【0160】

そのような一例示的実装において、維持範囲および/またはリセット範囲は、角度幅または角度高さを参照して標示または解析される。例えば、維持条件は、例えば關心領域の水平視野および垂直視野を含んでもよい。レンダリングされたビューポートが(初期観察向き、および關心領域の水平FOVと垂直FOVで標示される)關心領域全体をカバーする場合、観察設定はリセットされない。そうでない場合、観察設定は、ヘッドマウントディスプレイの以前の向き(例えば、現在の観察設定など)を考慮せずにリセットされる。別の例示的実装において、維持範囲および/またはリセット範囲は、特定の投影および特

10

20

30

40

50

定のマッピングの2D画像平面内の2次元領域（例えば矩形）を参照して標示または解析される。

【0161】

いくつかの例示的実装は、観察点が複数ある状況を想定する、および/またはその状況で生じる。そのような一部の状況では、VRコンテンツが複数のカメラで作成され、同じコンテンツに複数の代替の観察点が提供される。例えば、コンテンツは、ステージ上の1つまたは複数の固定VRカメラ、ステージ上でカメラマンが持っている1つ以上のVRカメラ、および/または移動可能なクレーン上に取り付けられた1つ以上のVRカメラによって撮影されてもよい。初期観察設定は、レンダリングに用いる初期観察点の選択を含んでもよい。ユーザが、コンテンツを視聴する際にある観察点から別の観察点へと切り替えることができるようにしてもよい。場合によっては、例えばカメラリグを用いた場合など、観察点同士が互いに近くてもよい。一実施形態では、観察点の維持範囲またはリセット範囲が標示または解析される。例えば、特定の隣接した観察点間で切り替える場合、初期観察設定を適用しないことが標示されてもよい（すなわち、観察点の維持範囲が標示される）。DASHに関連する実装では、観察点は、例えばDASH仕様ですでに定義されているViewpointプロパティ記述子によって識別されてもよい。VR観察点標示がViewpointプロパティ記述子またはその他のプロパティ記述子と共に用いられるように、@schemeIdUriの特定の値を定義してもよい。@valueを用いて観察点の識別子を指定してもよい。

10

【0162】

いくつかの例示的実装では、再生モード条件も想定する。再生モード条件は、例えば標準において事前に定義しても、例えばビデオプロセッサ106によってビデオストリーム内、またはビデオストリームの記述内で標示してもよい。再生モード条件は、初期観察設定が適用される再生モード、および/または初期観察設定が適用されない再生モードを標示してもよい。再生モードは、連続再生、連続再生を開始するためのシークまたはランダムアクセス、例えばイントラ符号化ピクチャのみが再生される早送り再生、例えばイントラ符号化ピクチャのみが再生される早戻し再生、などを含む場合があるがこれらに限定されない。再生モード条件により、観察設定は連続再生開始のためのシーク時にリセットされるが、早送りまたは早戻し時はリセットされないこと、またはその逆をコンテンツ作成者が標示できてよい。

20

【0163】

いくつかの例示的実装では、追加のシグナリング選択肢も想定する。そのような一例示的実装において、ユーザ主導の選択的レンダリングを意図した時限オーディオビジュアルコンテンツの提供の手法は、ある構文構造内で初期観察設定を標示することと、ランダムアクセス時に初期観察設定が無条件で適用されるかどうかをその構文構造内に標示することと、を特徴としてもよい。別の例示的実装において、ユーザ主導の選択的レンダリングを意図したオーディオビジュアルコンテンツへのアクセスの手法は、ある構文構造から初期観察設定を解析することと、ランダムアクセス時に初期観察設定が無条件で適用されるかどうかをその構文構造から解析することと、を特徴としてもよい。さらに別の例示的実装において、構文構造は、SMILやHTML5、またはそれらに含まれるプレゼンテーション情報（例えばカスケードスタイルシート）などの、プレゼンテーションレイヤに含まれても、またはそこから解析されてもよい。一実施形態では、構文構造はHTML5などのマークアップ言語の要素である。別の例示的実装では、構文構造は、DASH MPDやSDP（Session Description Protocol）などのプレゼンテーション記述やストリーミングマニフェストに含まれても、そこから解析されてもよい。

30

40

【0164】

DASHに関連する別の例示的実装では、構文構造はイベントであってもよい。イベントおよびそのイベントを含むEventStream要素を、例えば以下のように用いてもよい。

【0165】

観察設定識別子はEvent@idである。

【0166】

50

Event@messageDataは、連続再生時に初期観察設定を適用しないか、無条件で適用するか、条件付きで適用するかの標示、ランダムアクセス時に初期観察設定を適用しないか、無条件で適用するか、条件付きで適用するかの標示、初期観察点の標示、および/または初期観察向きの標示、の1つ以上を含む。

【0167】

EventStream@schemeIdUri内の特定のURIによって、含まれるイベントが初期観察設定に関する情報を提供することを識別する。

【0168】

そのような一例示的実装では、EventStream@valueによって、含まれるイベントに関する観察点を識別してもよい。別の例示的実装では、EventStream@valueによって、含まれるイベントに関するアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションを識別してもよい。

10

【0169】

別の例示的実装において、構文構造は、コンテナファイル、セグメント、またはサブセグメントに含まれる、またはそれから解析される。コンテナファイル、セグメント、またはサブセグメントがISOBMFFに準拠している場合、以下のような実施形態が可能である。すなわち構文構造は、インバンドイベントであってもよく、ISOBMFFの時限メタデータトラックのサンプルであってもよく、ISOBMFFのサンプルグループ記述エントリであってもよく、および/またはサンプルエントリ内のボックスであってもよい。

20

【0170】

別の例示的実装において、構文構造はメディアビットストリーム内にインバンドで含まれてもよい。例えば、構文構造は、ビデオビットストリーム内のSEIメッセージであっても、オーディオビットストリーム内の補助データユニットであってもよい。H.264/AVCやH.265/HEVCなどの多くのビデオ符号化標準では、ビデオビットストリーム内に付加拡張情報(Supplemental Enhancement Information: SEI)を含めることを可能にしている。SEIは、H.264/AVCおよびH.265/HEVCのSEI NALユニットなどの、特定のデータ構造にカプセル化されてもよい。このデータ構造は1つ以上のSEIメッセージを含んでもよい。SEIメッセージは出力ピクチャの復号には必要ないが、ピクチャ出力タイミング、レンダリング、エラー検出、エラー隠蔽、リソース予約などの関連プロセスを支援することができる。いくつかのSEIメッセージはH.264/AVCおよびHEVCで規定されており、ユーザデータSEIメッセージによって、組織や企業は独自の用途にSEIメッセージを指定することができる。H.264/AVCおよびHEVCは、規定されたSEIメッセージの構文と意味を含んでいるが、受領者においてそれらのメッセージを処理するプロセスについては定義されていない。したがって、エンコーダはSEIメッセージを作成する際にH.264/AVC標準またはHEVC標準に従う必要があり、それぞれ対応するH.264/AVC標準またはHEVC標準に準拠しているデコーダは、出力順を順守するためにSEIメッセージを処理する必要はない。H.264/AVCおよびHEVCにSEIメッセージの構文と意味が含まれている理由の1つは、異なるシステム仕様において付加情報を同様に解釈し、それによって相互作用できるようにするためである。システム仕様において、符号化側と復号側の両方で特定のSEIメッセージを用いるように求めることができ、さらに、受領者において特定のSEIメッセージを処理するためのプロセスを指定できるようにすることを意図している。

30

40

【0171】

いくつかの例示的実装は、DASHクライアント動作の状況を想定する、および/またはその状況で生じる。そのような一例示的実装では、シグナリングが前述のイベントストリームの手法によって実行されると想定されるが、この説明は他のシグナリングの選択肢にも同様に当てはまる。まず、DASHクライアントは、MPDから、初期観察設定用のイベントストリームが使用可能であることを解析する。そのようなイベントストリームが

50

使用可能でない場合、DASHクライアントは以下の処理を行うことができない。

【0172】

DASHクライアントは、次に、再生が開始される1つ目の(サブ)セグメントに一致またはそれをカバーするイベントを解析してもよい。そのイベントは、初期観察点(MPDが複数の観察点に対するコンテンツを宣言している場合)および初期観察向きの標示を含む。初期観察点イベントに含まれている場合、クライアントは、初期観察点に一致するアダプテーションセットを選択する。これには、例えば、観察点を標示する@schemeIdUriと、イベントに含まれている初期観察点識別子に等しい@valueとを有するViewpointプロパティ記述子を含むのはどのアダプテーションセットかを調べる。初期観察向きがイベントに含まれている場合、DASHクライアントは、その向きを含むアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションを選択する。これには、例えば、(参照によって本明細書に組み込まれるMPEG M38613に記載されているような)VRプロパティ記述子に標示されているビューポート(単数または複数)がその初期観察向きをカバーするかを調べる。その初期観察向きに一致するアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションを特定する際に、適用可能なグローバル向きオフセット(単数または複数)、および投影構造または投影されたフレームの向きを前述のように考慮してもよい。初期観察向きをカバーするアダプテーションセットが複数ある場合、DASHクライアントは、例えば、初期観察向きを(例えば、VRプロパティ記述子のquality_ranking値で標示される)最高品質で含むアダプテーションセット、および/またはアダプテーションセットがカバーするビューポート内で初期観察向きが最も中央にあるアダプテーションセットを選択してもよい。初期観察向きをカバーするアダプテーションセット内にリプレゼンテーションまたはサブリプレゼンテーションが複数ある場合、DASHクライアントは、例えば、初期観察向きを(例えば、VRプロパティ記述子のquality_ranking値で標示される)最高品質でカバーするビューポートがあるリプレゼンテーションまたはサブリプレゼンテーションを選択してもよい。アダプテーションセット、およびそのアダプテーションセットからリプレゼンテーションまたはサブリプレゼンテーションを選択した後、クライアントはそのリプレゼンテーションまたはサブリプレゼンテーションから1つ目の(サブ)セグメントを要求してもよい。

【0173】

連続再生中、DASHクライアントは初期観察設定についてイベントストリーム内のイベントを解析してもよい。連続再生に適用されるイベントのプレゼンテーション時間が一致した場合、クライアントはイベントを適用するかどうかを決定する。無条件で適用されるイベントは、現在の観察設定を(イベントに含まれる)初期観察設定と等しくなるようにリセットする。条件付きで適用されるイベントの場合、それらの条件が処理され、条件が満たされた場合、クライアントは現在の観察設定を(イベントに含まれる)初期観察設定と等しくなるようにリセットする。そのようリセットの後、それに従って後続の(サブ)セグメントが要求され、クライアントはそのコンテンツのレンダリングにもこの初期観察設定を用いる。それ以外の場合(現在の観察点がリセットされない場合)、クライアントは現在の観察点を続けて用いる。

【0174】

DASHクライアントは、ユーザにシークまたはランダムアクセスの機能を提供してもよい。シークの後、クライアントは、前述と同様に動作してもよいが、加えて、ランダムアクセス時に条件付きで適用される初期観察設定の処理について考慮してもよい。DASHクライアントは、初期観察設定を適用する条件の標示を処理してもよい。このために、DASHクライアントは、例えばヘッドマウントディスプレイから現在の視聴向きを取得してもよい。条件が満たされるか、またはランダムアクセス時に初期観察点が無条件で適用される場合、クライアントは現在の観察設定を初期観察設定と等しくなるようにリセットし、それに従って(サブ)セグメントを要求する。クライアントは、そのコンテンツのレンダリングにもこの初期観察設定を用いる。条件が満たされない場合、クライアントは現在の観察設定を続けて用いる。MPDから(例えば、前述の1つ以上の記述子から)の

10

20

30

40

50

投影およびマッピングメタデータを解析することで、DASHクライアントは、最高品質、かつ現行の推定ネットワークスループットで対応可能なビットレートで現在の視聴向きをカバーするアダプテーションセットおよびプレゼンテーションを判断する。これに従って、DASHクライアントは(サブ)セグメント要求を発行する。

【0175】

一例示の実装において、クライアントは、初期観察設定に関連付けられた標示を、それらが適用されるメディアデータより前に受信して解析する。例えば、DASHクライアントは、対応する(サブ)セグメントへの要求を発行する前に、イベントストリームを受信して解析することができる。クライアントは、少なくとも、連続再生時に無条件で適用される初期観察設定の標示はどれかを解析する。クライアントは、連続再生時に無条件で適用される初期観察設定表示に基づいて、次に発行される可能性のある(サブ)セグメント要求を判断する。クライアントは、可能性のある(サブ)セグメント要求を特定する際に、現行の推定ネットワークスループットおよびその他の側面(ディスプレイの視野など)も考慮してもよい。

10

【0176】

一例示の実装において、可能性のある後続の(サブ)セグメント要求は、URLとして、および場合によってはバイト範囲(単数または複数)として標示される。当該URLおよび関連付けられたバイト範囲(単数または複数)へのHTTP GETリクエストにより、可能性のある後続の(サブ)セグメント要求に変換される。

【0177】

一例示の実装において、前述のような通知はDASH SANDメッセージによって行われる。一実施形態において、前述のURL、および場合によってはバイト範囲(単数または複数)を伝達するために、SANDのAnticipatedRequestsメッセージが用いられる。

20

【0178】

一実施形態において、前述のような通知はHTTPヘッダを用いて行われる。このHTTPヘッダは、例えばGETリクエストの一部として含めることができる。一実施形態において、このHTTPヘッダはDASH SANDの仕様に準拠する。

【0179】

前述のように、いくつかの例示の実装において、ビデオプロセッサ106は顕著点を標示する情報を、当該ビデオストリームに、または当該ビデオストリームに関連付けられた別のストリーム(または、MPDなどのシグナリング構造)に埋め込む。

30

【0180】

一例示の実装において、顕著点を標示する情報は、ストリーミングの初期化時およびランダムアクセス時に、様々な実施形態における初期観察設定として解釈される。

【0181】

一実施形態において、顕著点情報は、時間的および空間的に正確であると理解されうる復号コンテンツ内の顕著点そのものではなく、要求される可能性のある(サブ)セグメントを標示する。

【0182】

一例示の実装による方法は、顕著点を含む(サブ)セグメントを標示する情報を受信することと、当該(サブ)セグメントの要求を示す信号を生成させることと、当該信号を伝送することと、を含む。

40

【0183】

一例示の実装において、顕著点情報は、(サブ)セグメント単位の時間の関数として生成される。換言すると、顕著点情報は、重複しない各(サブ)セグメント持続時間において、クライアントによって要求される可能性のある(サブ)セグメントを標示する。一実施形態において、(サブ)セグメント単位の顕著点情報は、それぞれがクライアントによって要求される可能性のある(サブ)セグメントを標示するURLのシーケンスとして標示される。一実施形態において、(サブ)セグメント単位の顕著点情報は、URLテンプレ

50

レート参照して、ならびにクライアントによって要求される可能性のある複数の（サブ）セグメントのURLを取得するためにURLテンプレートに挿入される属性値のリストおよび/または範囲を参照して、標示される。一実施形態において、（サブ）セグメント単位の顕著点情報は、MPDなどを参照して、ならびに（例えば、DASHの）階層データモデルに従うまたはそれと同様の、事前定義されたまたは標示された識別子階層の識別子値のリストまたは範囲を参照して、標示される。例えば、ピリオド識別子のリストが標示されてもよく、ピリオドごとに、アダプテーションセット識別子のリストが標示されてもよく、アダプテーションセットごとに、リプレゼンテーション識別子、および（例えば、セグメントタイムラインにおける）プレゼンテーション時間またはセグメント番号としてのそれらの有効期間が標示されてもよい。

10

【0184】

一例示的実装において、（サブ）セグメント単位の顕著点情報は、例えばヨー、ピッチ、およびロールとして、可能性のある視聴方向または視聴向きを標示することによって標示される。一実施形態において、可能性のある視聴方向または視聴向きは、グローバル向きオフセット（単数または複数）を含む。したがって、クライアントは、可能性のある視聴方向または視聴向きをカバーするアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションはどれかを判断する前にグローバル向きオフセット（単数または複数）をフェッチする必要はない。一例示的実装において、可能性のある視聴方向または視聴向きは、グローバル向きオフセット（単数または複数）を含まない。したがって、クライアントは、可能性のある視聴方向または視聴向きをカバーするアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションはどれかを判断する前にグローバル向きオフセット（単数または複数）をフェッチし、基本的に、可能性のある視聴方向または視聴向きをカバーするアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションはどれかを判断する際に、グローバル向きオフセット（単数または複数）と、可能性のある視聴方向および視聴向きとの合計を考慮する。可能性のある視聴方向または視聴向きの標示は、以下の意味を有するものとして理解することができる。すなわち、可能性のある視聴方向または視聴向きをカバーし、その可能性のある視聴方向または視聴向きに対して比較的高い品質を表示する品質ランキングを有するアダプテーションセット、リプレゼンテーション、またはサブリプレゼンテーションが、クライアントによって要求される可能性がある、という意味である。カバーされた視聴方向または視聴向き、および当該品質ランキングは、前述のように、例えば仮想現実ビデオ記述子（VRD）および/または空間関係記述子（SRD）によって標示されてもよい。

20

30

【0185】

一例示的実装において、顕著点情報は、例えばDASHEventMessageBoxなどのイベントを用いてビデオストリーム内に標示される。これらのイベントは、（サブ）セグメントの先頭で、その（サブ）セグメントの任意の「moof」ボックスの前に指定することができる。

【0186】

一例示的実装において、顕著点情報は、MPDなどにおいてイベントストリームなどとして標示される。いくつかの実施形態において、イベントのプレゼンテーション時間（@presentationTime）および持続時間（@duration）は、（サブ）セグメント境界に一致するように選択される。メッセージデータ（@messageData）は、クライアントによって要求される可能性のある（サブ）セグメントを標示する構造を含んでもよい。この構造に関する様々な選択肢は上述のとおりである。

40

【0187】

一例示的実装において、クライアントは、例えば、DASH MPDのEventStream要素などの前述の手段の1つによって、顕著点情報を受信する。クライアントは、前述のような顕著点情報に基づいて、次に発行される可能性のある（サブ）セグメント要求を判断する。クライアントは、プロキシキャッシュなどのネットワーク要素に、可能性のある後続の（サブ）セグメント要求について通知する

【0188】

50

一例示の実装において、クライアントは(サブ)セグメント単位の顕著点情報を受信する。

【0189】

一例示の実装において、可能性のある後続の(サブ)セグメント要求は、URLとして、および場合によってはバイト範囲(単数または複数)として標示される。当該URLおよび関連付けられたバイト範囲(単数または複数)へのHTTP GETリクエストにより、可能性のある後続の(サブ)セグメント要求に変換される。

【0190】

一例示の実装において、前述のような通知はDASH SANDメッセージによって行われる。一実施形態において、前述のURL、および場合によってはバイト範囲(単数または複数)を伝達するために、SANDのAnticipatedRequestsメッセージが用いられる。

10

【0191】

一例示の実装において、前述のような通知はHTTPヘッダを用いて行われる。このHTTPヘッダは、例えばGETリクエストの一部として含めることができる。一実施形態において、このHTTPヘッダはDASH SANDの仕様に準拠する。

【0192】

一例示の実装において、プロキシキャッシュやエッジサーバなどのネットワーク要素は、例えば前述の手段の1つによって顕著点情報を受信する。ネットワーク要素は、前述のような顕著点情報に基づいて、次に発行される可能性のある(サブ)セグメント要求を判断する。ネットワーク要素はその(サブ)セグメントをプリフェッチする。したがって、クライアントがこれらの(サブ)セグメントに対する要求を発行するとき、これらの(サブ)セグメントはネットワーク要素内ですぐに利用可能であるため、要求に対してより高速な応答が得られる。

20

【0193】

一例示の実装において、プロキシキャッシュやエッジサーバなどのネットワーク要素は、例えばクライアントから、可能性のある後続の(サブ)セグメントに関する情報を受信する。ネットワーク要素はその(サブ)セグメントをプリフェッチする。したがって、クライアントがこれらの(サブ)セグメントに対する要求を発行するとき、これらの(サブ)セグメントはネットワーク要素内ですぐに利用可能であるため、要求に対してより高速な応答が得られる。

30

【0194】

一例示の実装において、グローバル向きオフセット情報は、(サブ)セグメント単位の時間の関数として生成される。グローバル向きオフセットは(サブ)セグメントから導出されてもよいため、(サブ)セグメント単位のグローバル向きオフセット情報は、その(サブ)セグメント内のグローバル向きオフセットの変動をカバーするグローバル向きオフセットの範囲を標示してもよい。一実施形態において、(サブ)セグメント単位のグローバル向きオフセット情報は、インバンドのイベントストリームとしての、またはMPD内の、DASHイベントに含まれる。

【0195】

図4は、仮想現実環境内のオーディオビジュアルコンテンツの向きの変更に関連する技術的課題に対処するために、本明細書で説明する解決策をいかに用いることができるかを示す別の図である。図4に示すように、オーディオビジュアルコンテンツ400は8つのビュー402~416を含んでいる。これらのビューは、360度の視野を表現するために用いることができ、NokiaのOZOカメラシステムやその他の複数カメラシステムなどによって撮影されている。図4に示す例示の実装において、MPVDは、ビュー402~416に含まれるコンテンツに関して常に既知であり、点418、420、422、424、および426で示されている。視聴者がMPVDからずれて(または、MPVDに合わせるような物理的移動が困難および/または不快となるように)配向および/または配置される可能性を低減するために、ビュー402~416は、特定のカメラビューに

40

50

対してMPVDが最も中心に配置されているビュー（単数または複数）が視聴者の正面に提示されるように視聴者に対してレンダリングされてもよい。例えば、図4に示すように、MPVD点424および426は、それらの対応するビュー410および412の中心に比較的近い。囲み428によって示すように、ビュー410および412が、視聴者の正面に提示されるように選択されレンダリングされる。これにより、視聴者の位置または向きによって通常は別のビュー（単数または複数）がその視聴者の正面に提示される場合でも、その視聴者にMPVDが提示される。残りのビュー402、404、406、408、414、および416も、視聴者の正面になるように移動するビュー410および412に関連付けられたシフトを考慮してレンダリングすることができる。

【0196】

図5は、仮想現実環境内のオーディオビジュアルコンテンツの向きの変更に関連する技術的課題に対処するために、本明細書で説明する解決策をいかに用いることができるかを示す別の図である。図4からのオーディオビジュアルコンテンツ400を、ビュー402～416およびMPVD点418～426を含めて図5に示している。図4では、MPVD点から対応するビューの中心までの距離を用いて、どのビュー（単数または複数）を視聴者の正面に提示するかを確認するのに対し、図5では、ビューのスケラブル符号化を想定する実装を示している。1つのビュー内の複数のレイヤによって、そのビュー内の異なるビューポートをカバーしてもよい。図5は、1つのビューの複数のレイヤに対するビューポートが、向きは同じであるが範囲が増大する視野を有する場合を描いている。しかしながら、別のレイヤ状符号化構成も可能であり、例えば、複数のレイヤが部分的に重複し、向きも異なるが、ビュー内の視野は同じである構成を含むが、これに制限されない。図5では、レイヤによってビューの中心の近接領域が定義される。図5に示すように、レイヤ402'および402''がビュー402内にマークされている。ここで、その他のビュー404～416内にもそれぞれ対応するレイヤが示されていることを理解されたい。図5に示す例示の実装では、特定のビューの中心からMPVD点までの絶対距離を計算する必要はない。MPVD点が含まれる特定のレイヤに基づいて、ビューをソートおよび選択することができる。図5に示すように、MPVD点424および426は両方とも、それぞれビュー410および412内の最も内側のレイヤ内に配置されており、他のMPVD点は、それぞれ対応するビューの外側のレイヤ内に少なくとも部分的に配置されているように示されている。したがって、囲み428によって示すように、ビュー410および412を視聴者の真正面に提示するように選択してレンダリングすることができる。

【0197】

図4および図5に示す例示の実装は、立体視ビューの状況において、ならびに、MPVDに存在するコンテンツが、視聴者（または、ユーザが用いる視聴デバイス）の位置および/または向きに関係なく、可能なかぎり頻繁にユーザの正面に提示されるようにレンダリングされる状況において、特に有効でありうる。図4および図5に関して本明細書で説明する例では、特定のストリーム内に存在する任意のビュー候補を選択することを想定しているが、視聴者に提示されるビューの選択には追加の制約が課せられる場合があることを理解されたい。例えば、ビューの選択は、コンテンツの開始時のヘッドマウントディスプレイの方向に少なくとも部分的に依存しうる。例えば、ヘッドマウントディスプレイが上を向いている場合、選択されるビュー候補の集合は、（垂直軸の観点から）上方のビューに制限されうる。同様に、ディスプレイが、選択可能なビューの（垂直軸の観点から）中央に概して向けられている場合、選択されるビュー候補の集合は、同様の垂直位置に沿って存在するビューに制限されうる。

【0198】

ここで図6Aを参照すると、本発明の一例示の実施形態による、図2の装置200によって実行される動作が、例示のプロセスフロー600として示されている。この点において、この装置は、オーディオビジュアルプレゼンテーションの伝送単位のセットにおける初期観察設定に関連付けられた標示を受信し、再生デバイスの意図された動作に関連付けられた標示を受信し、その再生デバイスの意図された動作を特定し、その特定に応じて制

10

20

30

40

50

御信号を生成させるための、プロセッサ202、メモリ204、通信インタフェース208などの手段を備えている。この制御信号は、再生デバイス上でのオーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。このように、この装置は一般的に、本明細書において説明あるいは想定しているオーディオビジュアルコンテンツの制御された観察および向きを選択を有効にすることができる。

【0199】

この装置は、オーディオビジュアルプレゼンテーションの伝送単位のセットにおける初期観察設定に関連付けられた標示を受信するための、プロセッサ202、メモリ204、通信インタフェース208などの手段を備えている。図6Aを参照すると、プロセスフロー600はブロック602において、初期観察設定に関連付けられた標示の受信を開始する。プロセスフロー600のいくつかの例示的実装では、観察設定は観察点および観察向きを含んでもよい。DASH環境において生じる例示的実装において、初期観察設定は、視聴者に提示されるオーディオビジュアルコンテンツの各セグメントおよび/またはサブセグメントのストリームアクセスポイント(SAP)に標示されてもよい。

【0200】

この装置は、再生デバイスの意図された動作に関連付けられた標示を受信するための、プロセッサ202、メモリ204、通信インタフェース208などの手段も備えている。図6Aを参照すると、プロセスフロー600は続いてブロック604において、再生デバイスの意図された動作に関連付けられた標示を受信する。一般的に、多くの例示的実装では、再生デバイスの意図された動作に関連付けられた標示では、特定のコンテンツを特定の方法でレンダリングする、コンテンツクリエイタの好みの「強さ」を特定することを可能にし、視聴者の位置(例えば、ユーザの頭部の位置や、視聴デバイスのその他の向きなど)がコンテンツクリエイタの好みに優先しうる条件を想定している。ブロック604のいくつかの例示的実装において、再生デバイスの意図された動作には、ある条件を満たしたうえで、観察設定を初期観察設定に設定することが含まれる。ブロック604のいくつかの例示的実装において、再生デバイスの意図された動作に関連付けられた標示は、再生デバイスの連続再生モードにおける再生デバイスの意図された動作に関連付けられた標示と、再生デバイスのランダムアクセスモードにおける再生デバイスの意図された動作に関連付けられた標示とを含む。したがって、ブロック604のいくつかの例示的実装では、連続再生モードの際はコンテンツがある方法でレンダリングされ、コンテンツのセグメントまたはサブセグメントがランダムアクセスされる際はコンテンツが別の方法でレンダリングされうるといふ点で、再生デバイスにおける状況依存の動作の程度を想定している。

【0201】

この装置は、再生デバイスの意図された動作を特定するための、プロセッサ202、メモリ204、通信インタフェース208などの手段を備えている。図6Aを参照すると、プロセスフロー600は続いてブロック606において、再生デバイスの意図された動作を特定する。ブロック606のいくつかの例示的実装において、再生デバイスの意図された動作を特定することは、再生デバイスの意図された動作に関連付けられた条件が満たされたかどうかを判断することを含む。いくつかの例示的実装では、連続再生時に初期観察設定を(1)適用しない、(2)条件付きで適用する、または(3)無条件で適用する、という標示を含む状況を想定している。同様に、そのようないくつかの例示的実装および他の例示的実装では、セグメントまたはサブセグメントのランダムアクセスが発生したときに、初期観察設定を(1)適用しない、(2)条件付きで適用する、または(3)無条件で適用する、という標示を含む状況を想定している。

【0202】

いくつかの例示的実装において、条件は、初期観察設定に関連付けられた少なくとも1つのリセット条件を含む。例えば、満たされた場合に初期観察設定を適用するというリセット条件であってもよい。そのような例示的実装において、リセット条件は、初期観察設定を適用させる観察点および/または観察向きのリセット範囲という形式であってもよい。また、いくつかの実装では、リセット条件は現在の観察設定に少なくとも部分的に基づ

10

20

30

40

50

いてもよい。

【0203】

そのようないくつかの例示の実装、およびその他の例示の実装において、条件は、例えば初期観察設定を適用させない維持条件などの、初期観察設定に関連付けられた少なくとも1つの維持条件を含んでもよい。そのような例示の実装において、維持条件は、初期観察設定を適用させない観察点および/または観察向きを標示する維持範囲を含んでもよい。

【0204】

この装置は、その再生デバイスの意図された動作の特定に応じて制御信号を生成させるための、プロセッサ202、メモリ204、通信インタフェース208などの手段を備えている。この制御信号は、再生デバイス上でのオーディオビジュアルプレゼンテーションのレンダリング動作に関連付けられている。図6Aを参照すると、プロセスフロー600はブロック608へと続く。ブロック608は、その再生デバイスの意図された動作の特定に応じて、レンダリング動作に関連付けられた制御信号を生成させることを含む。ブロック608のいくつかの例示の実装において、制御信号は、再生デバイスの意図された動作を標示する。ブロック608のいくつかの例示の実装において、再生デバイス上でのオーディオビジュアルプレゼンテーションのレンダリング動作は、オーディオビジュアルプレゼンテーションの部分を選択することを含む。本明細書において、かつ図3、図4、および図5を参照して説明したように、プロセス600の実装および本発明の他の実施形態では、コンテンツのセグメントまたはサブセグメントに関連付けられた初期観察設定、視聴デバイスの向き、および/またはセグメントまたはサブセグメントに到達する方法（例えば、連続再生またはランダムアクセス）に基づいて、視聴者に対するコンテンツの相対位置を再配置および/またはシフトするように、視聴者に提示されたコンテンツをレンダリングするかどうか、およびどの程度そのようにするか、が想定される。ブロック608のいくつかの例示の実装において、再生デバイスの意図された動作を（例えば、初期観察設定の適用に関連付けられた条件が満たされたという判断に基づいて）特定すると、意図された方法でコンテンツを直接的または間接的にレンダリングさせ、視聴者に提示させる制御信号が生成される。

【0205】

プロセス600のいくつかの実装、および本明細書に記載した本発明の他の実施形態では、再生デバイスの意図された動作を繰り返し特定することを想定していることを理解されたい。例えば、いくつかの例示の実装では、オーディオビジュアルコンテンツの伝送単位の第1セットに関連付けられた第1観察設定識別子と、伝送単位の第2セットに関連付けられた第2観察設定識別子とを想定している。そのようないくつかの例示の実装では、伝送単位は1つ以上のアダプテーションセット、リプレゼンテーション、サブリプレゼンテーション、セグメントセット、サブセグメントセット、および/または時間範囲によって定義されてもよい。これらの異なる伝送単位は、個々の伝送単位の意図された動作がそれぞれ異なるように、異なるリセット条件および/または維持条件に関連付けられてもよい。また、個々の伝送単位にそれぞれ関連付けられた条件は、ある伝送単位に関連付けられた条件および/または意図された動作が、別の伝送単位に関連付けられた条件および/または意図された動作に少なくとも部分的に基づくように、互いに関連付けられるか、あるいはリンクされてもよい。

【0206】

ここで図6Bを参照すると、本発明の一例示の実施形態による、図2の装置200によって実行される別の一連の動作が、例示のプロセスフロー1000として示されている。プロセスフロー1000のいくつかの例示の実装は、例えば、仮想現実ヘッドセット114などの視聴デバイス内で生じるおよび/またはそれによって実行される実施形態において特に有利でありうることを理解されたい。この点において、この装置は、オーディオビジュアルプレゼンテーションの伝送単位のセットに関連付けられた観察設定を検出するための、プロセッサ202、メモリ204、通信インタフェース208などの手段を備えてい

10

20

30

40

50

る。図6Bを参照すると、プロセスフロー1000はブロック1002において、オーディオビジュアルプレゼンテーションの伝送単位に関連付けられた観察設定の検出を開始する。観察設定を検出することは、観察設定が受信される方法に応じて、多数の手法の中から任意のものに従って実行されてもよい。観察設定が受信される方法は、オーディオビジュアルプレゼンテーションおよびその伝送に関連付けられたプロトコルおよびフォーマットに基づいてもよい。本明細書で説明および/または想定するメディアフォーマットおよびプロトコルのいずれも、プロセス1000およびブロック1002の実装に用いることができることを理解されたい。いくつかの例示的実施形態において、観察設定は観察点と観察向きいずれかまたは両方を含む。いくつかの例示的実装において、観察設定は最も取りうる視聴方向の標示を含んでもよい。コンテンツクリエイタによる識別によって、および/またはコンテンツの複数回の視聴および/または複数の視聴者に関連付けられた利用データの収集および処理によってMPVDが特定されている場合、MPVDについて説明している実装は特に有利でありうる。

10

【0207】

この装置は、観察設定に関連付けられた条件が満たされたかどうかを判断するための、プロセッサ202、メモリ204、通信インタフェース208などの手段も備えている。図6Bを参照すると、プロセスフロー1000は続いてブロック1004において、観察設定に関連付けられた条件が満たされたかどうかを判断する。本明細書において説明したように、本発明の実施形態の多くの例示的実装では、コンテンツの態様(向き設定、MPVD、および/または、コンテンツに関する他の情報など)、および/または他の要因(視聴者および/または視聴デバイスの位置、ユーザがコンテンツの特定のセグメントまたはサブセグメントに到達する状況、および/または他の要因)に基づいて、オーディオビジュアルコンテンツに関連付けられた観察点および観察点向きの選択を制御し、それに応じて、特定の仕方方でコンテンツを視聴者に表示することが想定される。ブロック1004のいくつかの例示的実装において、観察設定に関連付けられた条件が満たされたかどうかを判断することは、再生デバイスに関連付けられた向きを特定することを含む。例えば、仮想現実ヘッドセット114などの再生デバイスは、例えば、再生デバイス、ユーザ、またはその両方の絶対位置および/または相対位置(ロール、ピッチ、ヨー、および/または視角度を含むがこれに限らない)に関連付けられた情報を検出して提供するように構成されてもよい。いくつかの例示的実装において、再生デバイスに関連付けられた条件が満たされたかどうかを判断することは、その再生デバイスが連続再生モードであるかどうか、および/またはその再生デバイスがランダムアクセスモードであるかどうかを判断することを含む。本明細書において説明したように、視聴者に望ましい視聴体験を提供するために、所定の状況で特定の観察設定を適用するかどうかは、ユーザがオーディオビジュアルプレゼンテーションの特定の部分にランダムアクセスしたかどうか、または、それより前の部分から継続的に視聴することによってその同じ部分に到達したか、によって決定してもよい。

20

30

【0208】

この装置は、観察に関連付けられた条件が満たされたかどうかの判断に応じて、オーディオビジュアルプレゼンテーションの伝送単位のサブセットを選択するための、プロセッサ202、メモリ204、通信インタフェース208などの手段も備えている。図6Bを参照すると、プロセスフロー1000は続いてブロック1006において、伝送単位のサブセットを選択する。図3、図4、図5、およびそれらに関連する説明を参照すると、観察設定を受信し、その観察設定に関連付けられた単数または複数の条件が満たされたかどうかを判断すると、オーディオビジュアルプレゼンテーションに関連付けられた1つ以上の部分またはビューを、ユーザに表示するために選択することができる。本明細書において説明および/または想定した、提示するビューを識別および/または選択する手法のいずれも、ブロック1006の例示的実装において用いることができる。

40

【0209】

この装置は、制御信号を生成させるための、プロセッサ202、メモリ204、通信イ

50

ンタフェース 208 などの手段も備えている。この制御信号は、再生デバイス上でのオーディオビジュアルプレゼンテーションの伝送単位の選択されたサブセットのレンダリング動作に関連付けられている。図 6 B を参照すると、プロセスフロー 1000 は続いてブロック 1008 において、伝送単位の選択されたサブセットのレンダリングに関連付けられた制御信号を生成させる。ブロック 1008 のいくつかの例示的実装において、制御信号は、再生デバイスの意図された動作の標示を含む。例えば、観察設定に関連付けられた任意の条件が満たされたかどうかを判断し、表示する 1 つ以上のビューを選択したら、この装置は、仮想現実ヘッドセット 114 などの視聴デバイスに、例えば、選択されたコンテンツを特定の意図された方法でレンダリングするように命令および / またはその他の方法でそうさせる、制御信号を生成してもよい。例えば、受信された観察設定を適用する状況では、制御信号は、その観察設定に従ってコンテンツをレンダリングするために、レンダリングプロセスおよび / または視聴デバイスが認識および処理できるように生成されてもよい。同様に、受信された観察設定を適用しない状況では、制御信号は、受信された観察設定が、視聴者にレンダリングされたビューに影響を与えないことをレンダリングプロセスおよび / または視聴デバイスに対して認めるように生成されてもよい。

10

【0210】

前述のように、図 6 A および図 6 B は、本発明の例示的実施形態による装置、方法、およびコンピュータプログラム製品のフローチャートを示している。これらのフローチャートの各ブロック、およびフローチャート内のブロックの組合せは、ハードウェア、ファームウェア、プロセッサ、回路構成、および / または 1 つ以上のコンピュータプログラム命令を含むソフトウェアの実行に関連付けられたその他のデバイスなどの、様々な手段によって実装されてもよいことを理解されたい。例えば、前述の手順の 1 つ以上は、コンピュータプログラム命令によって具体化されてもよい。この点において、前述の手順を具現化するコンピュータプログラム命令は、本発明の実施形態を用いる装置のメモリデバイス 204 によって記憶され、この装置のプロセッサ 202 によって実行されてもよい。理解されるように、そのような任意のコンピュータプログラム命令は、コンピュータまたはその他のプログラム可能な装置（例えば、ハードウェア）上にロードされてマシンを構成し、それによってそのコンピュータまたはその他のプログラム可能な装置がフローチャートのブロックに指定された機能を実行するようにしてもよい。これらのコンピュータプログラム命令は、コンピュータ可読メモリに記憶され、コンピュータまたはその他のプログラム可能な装置を特定の方法で機能させてもよく、このコンピュータ可読メモリに記憶された命令によって製造品を作成し、それを実行することでフローチャートのブロックに指定された機能が実行されるようにしてもよい。また、コンピュータプログラム命令は、コンピュータまたはその他のプログラム可能な装置にロードされて、一連の動作をそのコンピュータまたはその他のプログラム可能な装置上で実行させてコンピュータ実装プロセスを作成してもよく、そのコンピュータまたはその他のプログラム可能な装置上で実行される命令により、フローチャートのブロックに指定された機能を実行するための動作を提供するようにしてもよい。

20

30

【0211】

したがって、フローチャートのブロックは、指定された機能を実行するための手段の組合せと、指定された機能を実行するための動作の組合せに対応している。フローチャートの 1 つ以上のブロック、およびフローチャート内のブロックの組合せは、指定された機能を実行する専用ハードウェアベースのコンピュータシステム、または専用ハードウェアとコンピュータ命令の組合せによって実装できることも理解されたい。

40

【0212】

いくつかの実施形態において、前述の動作のうち特定のものは、変更またはさらに拡張してもよい。また、いくつかの実施形態では、任意の動作を追加で含めてもよい。前述の動作の変更、追加、または拡張は、任意の順序または組合せで行ってもよい。

【0213】

本明細書において、いくつかの実施形態は 360 ° ビデオを参照して説明している。本

50

明細書で用いる360°ビデオという用語は、任意の投影フォーマットをカバーするものとして理解されるべきである。また、いくつかの実装では360°視野を想定しているが、例示的な実装は、本明細書に記載した実施形態の範囲から逸脱することなく、360°以外の視野を含むがそれに限定されず、一般的にそれより狭い他の視野に関連して用いられ

【0214】

前述において、いくつかの実施形態は、ISO BMFFに関連して、および/またはISO BMFFから派生したフォーマットに関連して説明されている。しかしながら、多くの例示的な実施形態は、Matroskaファイルフォーマットを含むがそれに限定されない、他のファイルおよびセグメントのフォーマットに同様に当てはまる。

10

【0215】

前述において、いくつかの実施形態は、HTTPおよび/またはHTTP GETリクエストに関連して説明されている。実施形態はHTTPの利用に限定されず、加えてまたは代わりにWeb Socketsなどの他のプロトコルを用いてもよいことを理解する必要がある。同様に、HTTP/1.1やHTTP/2などの異なるバージョンのHTTPを用いてもよい。同様に、HTTPSを用いてもよい。本発明は、クライアントによってHTTP GETなどのリクエストが発行されない場合、例えば、クライアントが3GPPのMBMS (Multimedia Broadcast/Multicast Service) に従ったブロードキャストサービスなどの、ブロードキャストを介してデータを受信する場合に適用してもよいことも理解する必要がある。

20

【0216】

前述において、いくつかの実施形態は、MPEG-DASHまたはDASHに関連して説明されている。しかしながら、例示的な実装および実施形態は、例えばApple HTTP Live Streaming (HLS) などのHTTPを介したストリーミングの他の形式にも同様に当てはまる。実施形態におけるDASH固有の用語は、他のストリーミングのフォーマットおよびシステムにおける同様の用語に調整できることを理解されるべきである。

【0217】

前述において、いくつかの実施形態は、MPEG-DASHのMPD (Media Presentation Description) に関連して説明されている。しかしながら、例示的な実装および実施形態は、例えばHLS M3Uフォーマットなどの他のストリーミングマニフェストフォーマットや、SDP (Session Description Protocol) などの他のストリームまたはプレゼンテーション記述フォーマットにも同様に当てはまる。

30

【0218】

前述において、いくつかの実施形態は、(サブ)セグメントという用語に関連して説明されている。この表現にある括弧は、これらの実施形態はサブセグメントという用語とセグメントという用語に等しく当てはまることを示すことを意味している。また、実装は、セグメントまたはサブセグメントと同様の単位に同様に適用してもよい。例えば、実施形態は自己完結型のムービーフラグメントに適用することができる。

【0219】

前述において、いくつかの実施形態は、MPEG-DASHのイベントおよび/またはイベントストリームに関連して説明されている。実施形態は、イベントと同様のエンティティ、およびイベントストリームと同様のエンティティストリームにも同様に当てはまることを理解する必要がある。例えば、実施形態は、イベントストリームではなく、(オーディオ/ビデオのリプレゼンテーションとは別個のリプレゼンテーションで伝達される) 時限メタデータトラックによって実現されてもよく、その場合イベントは、いくつかの実施形態において、時限メタデータトラックのサンプルに対応する。

40

【0220】

前述において、いくつかの実施形態は、ストリーミングという用語に言及して説明されている。しかしながら、例示的な実装および実施形態は、プログレッシブダウンロード、

50

ファイル配信、および会話的ビデオ通信（例えばビデオ電話など）などの、ビデオ伝送の他の形式にも同様に当てはまる。

【0221】

前述において、いくつかの実施形態は、顕著点という用語に言及して説明されている。しかしながら、同様または同じ意味を有する、例えば関心点などの用語を代わりに用いることができる。また、実装は、単一の点（例えば、単一の画素として解釈することができる）ではなく、顕著点の代わりに顕著領域または関心領域に言及しても同様に実現することができる。

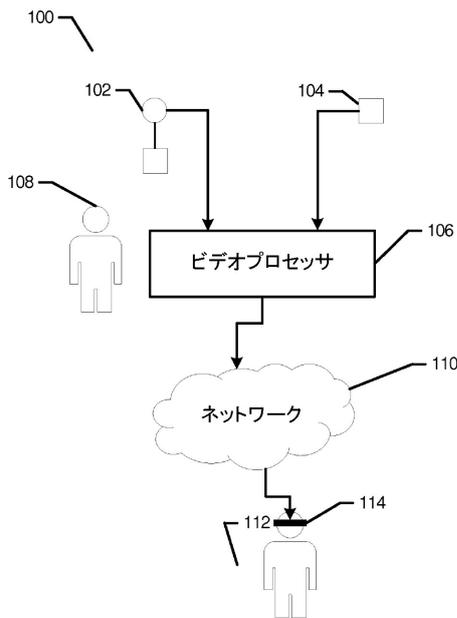
【0222】

前述において、方向および向きという表現は区別なく用いられる場合があるが、向きという用語に含まれている回転「成分」が、方向という用語には含まれていないことがある。実施形態はいずれの解釈でも実装することができ、いずれの用語も用いることができることを理解する必要がある。

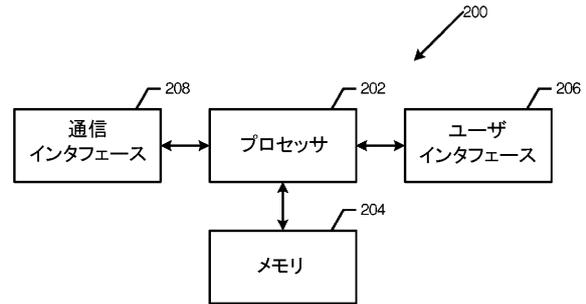
【0223】

本明細書において定める本発明の多くの変形例および他の実施形態を、これらの発明に関連し、前述の説明および関連する図面に提示された教示から利益を得る当業者は想到されるであろう。したがって、本発明は、開示した特定の実施形態に限定されず、変形例および他の実施形態は付随する請求の範囲に含まれると意図されていることを理解されたい。また、前述の説明および関連する図面は、特定の例示的な要素および/または機能の組合せの状況における例示的な実施形態を説明しているが、代替的な実施形態によって、付随する請求の範囲から逸脱することなく、異なる要素および/または機能の組合せを提供できることを理解されるべきである。この点において、付随する請求項のいくつかに定めようように、例えば、明示的に前記したものは異なる要素および/または機能の組合せも想定される。本明細書において具体的な用語を用いているが、これらの用語は、限定するためではなく、一般的小および説明的な意味でのみ用いられる。

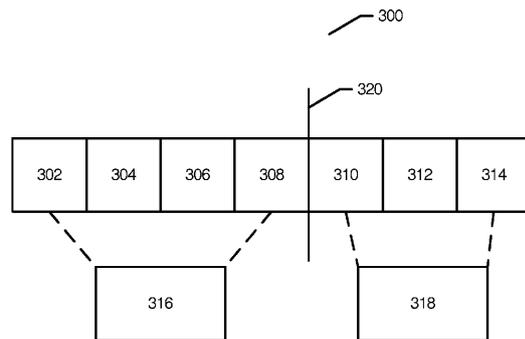
【図1】



【図2】



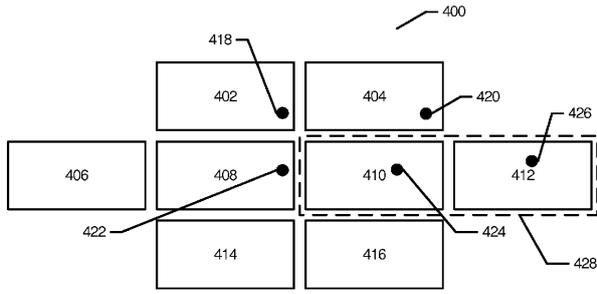
【図3】



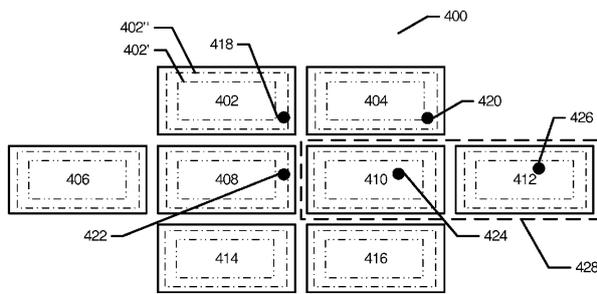
10

20

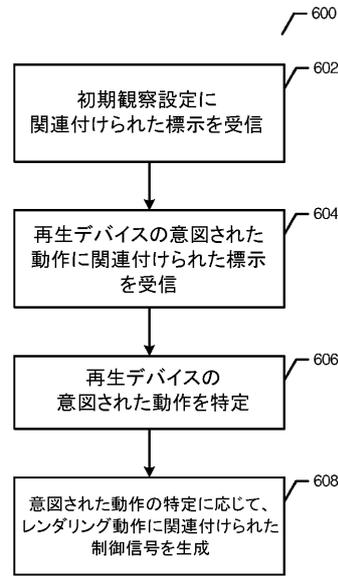
【図4】



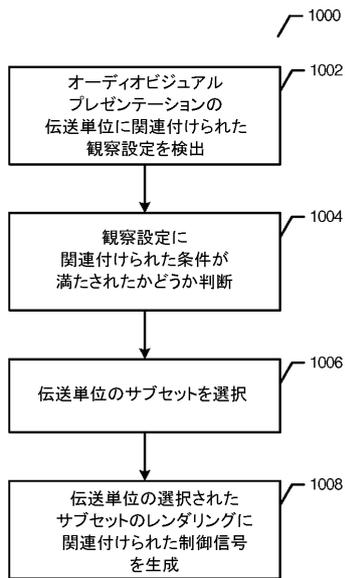
【図5】



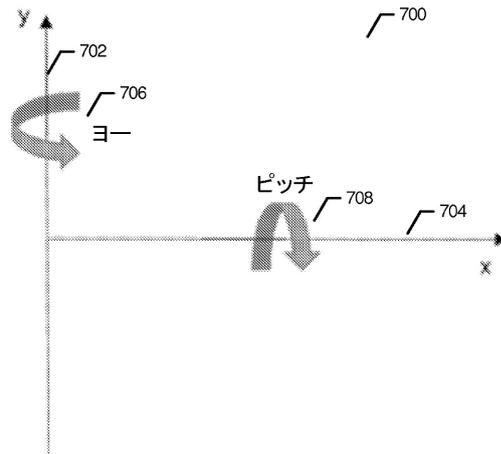
【図6A】



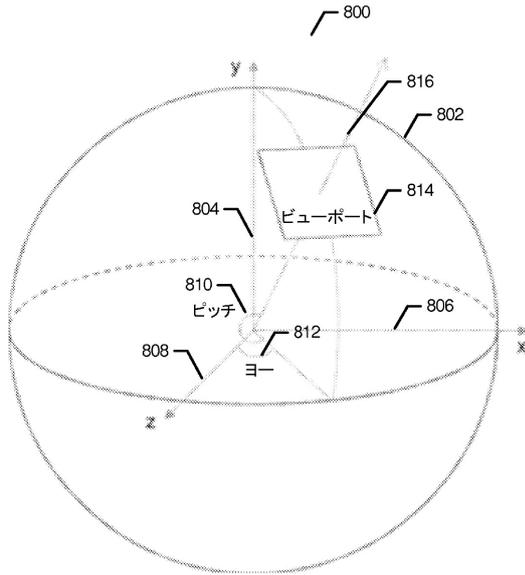
【図6B】



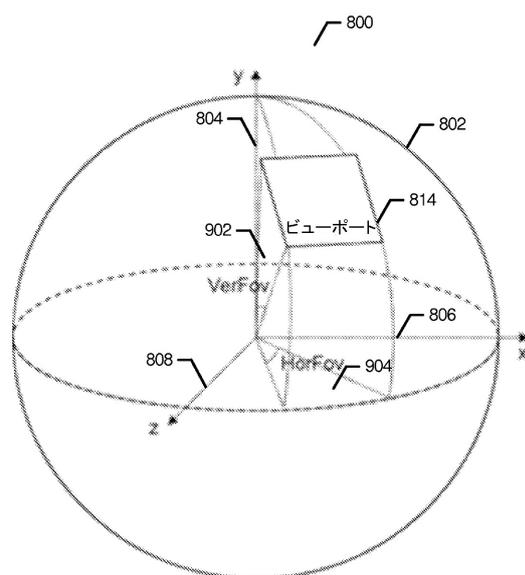
【図7】



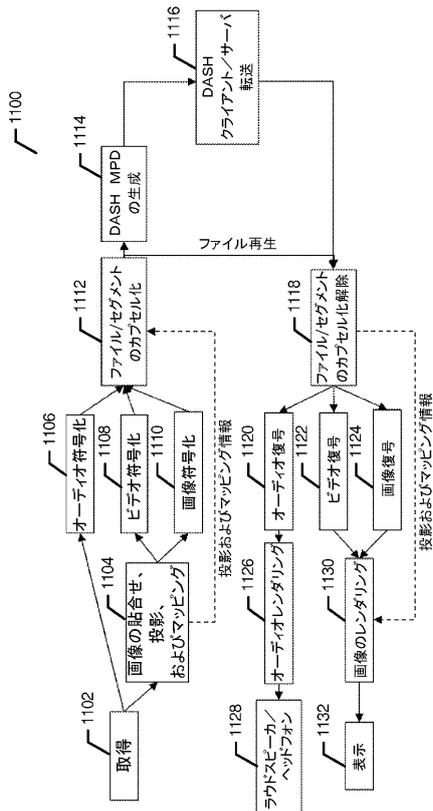
【図8】



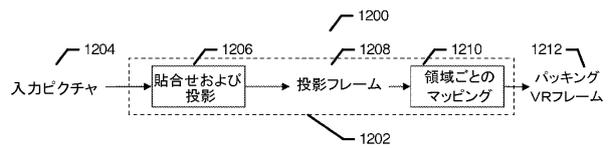
【図9】



【図10】



【図11】



フロントページの続き

(56)参考文献 国際公開第2016/009864(WO, A1)
特開2016-105593(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F3/01
3/048 - 3/0489
13/00
G06T1/00
11/60 - 13/80
17/05
19/00 - 19/20
G09F9/00
G09G5/00 - 5/36
5/377 - 5/42
H04N7/10
7/14 - 7/173
7/20 - 7/56
13/00 - 17/06
21/00 - 21/858