

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7679014号
(P7679014)

(45)発行日 令和7年5月19日(2025.5.19)

(24)登録日 令和7年5月9日(2025.5.9)

(51)国際特許分類

F I

G 0 6 T 9/40 (2006.01)

G 0 6 T 9/40

H 0 4 N 19/85 (2014.01)

H 0 4 N 19/85

H 0 4 N 19/597(2014.01)

H 0 4 N 19/597

請求項の数 18 外国語出願 (全32頁)

(21)出願番号	特願2023-119205(P2023-119205)	(73)特許権者	507236292
(22)出願日	令和5年7月21日(2023.7.21)		ドルビー ラボラトリーズ ライセンシン
(62)分割の表示	特願2022-577293(P2022-577293)の分割		グ コーポレイション
原出願日	令和3年6月16日(2021.6.16)		アメリカ合衆国 9 4 1 0 3 カリフォル
(65)公開番号	特開2023-139163(P2023-139163 A)		ニア州 サンフランシスコ マーケット
(43)公開日	令和5年10月3日(2023.10.3)	(74)代理人	ストリート 1 2 7 5
審査請求日	令和6年6月12日(2024.6.12)		100107766
(31)優先権主張番号	20180179.2	(74)代理人	弁理士 伊東 忠重
(32)優先日	令和2年6月16日(2020.6.16)		100070150
(33)優先権主張国・地域又は機関	欧州特許庁(EP)	(74)代理人	弁理士 伊東 忠彦
(31)優先権主張番号	63/039,595		100135079
(32)優先日	令和2年6月16日(2020.6.16)	(74)代理人	弁理士 宮崎 修
最終頁に続く		(72)発明者	ワード, グレゴリー ジョン
			アメリカ合衆国 9 4 1 0 3 カリフォル
			ニア州 サンフランシスコ マーケット
			最終頁に続く

(54)【発明の名称】 ディスオクルージョンアトラスを用いたマルチビュービデオ動作のサポート

(57)【特許請求の範囲】

【請求項 1】

方法であって、

1 つまたは複数の基準ビューからの視覚シーンを描写する 1 つまたは複数の基準画像ではオクルードされ、前記 1 つまたは複数の基準ビューに隣接する非基準ビューでは少なくとも部分的にディスオクルードされるようになる画像フラグメントをサイズでソートするステップと、

前記画像フラグメントを記憶するために使用されるディスオクルージョンアトラスのためのレイアウトマスクを生成するステップであって、前記ディスオクルージョンアトラスは、複数のオーバーラップしていない画像フラグメントを含む最小総面積の結合画像である、ステップと、

前記レイアウトマスク内のノードに前記ソートされた画像フラグメントを降順で記憶するステップと、

前記 1 つまたは複数の基準画像で符号化されたボリュメトリックビデオ信号を生成するステップであって、前記ボリュメトリックビデオ信号は、前記ディスオクルージョンアトラス内の前記画像フラグメントを用いてさらに符号化される、ステップと

を含む方法。

【請求項 2】

前記 1 つまたは複数の基準画像の各々は、3 6 0 度画像、1 8 0 度画像、ビューポート画像、規則的な空間形状画像フレーム内の画像、または不規則な空間形状画像フレーム内

の画像のうちの 1 つを表す、請求項 1 に記載の方法。

【請求項 3】

前記画像フラグメントの各々は、前記 1 つまたは複数の基準ビューにおいてオクルードされた連続するピクセルによって形成される空間領域について、テクスチャ画像値および深度画像値を含む、請求項 1 または 2 に記載の方法。

【請求項 4】

前記 1 つまたは複数の基準画像は、複数の連続する時点についての複数の連続するマルチビュー画像を含むマルチビュー画像グループ内のマルチビュー画像に含まれ、前記ディスオクルージョンアトラスは、前記複数の連続する時点についての複数のディスオクルージョンアトラスを含むディスオクルージョンアトラスグループに含まれる、請求項 1 から 3 のいずれか一項に記載の方法。

10

【請求項 5】

前記レイアウトマスクは、前記複数のディスオクルージョンアトラスのために生成された複数の個別のレイアウトマスクに含まれ、グループレベルのレイアウトマスクは、和集合演算により前記複数の個別のレイアウトマスクから生成され、前記ボリュメトリックビデオ信号に符号化された前記ディスオクルージョンアトラスは、前記グループレベルのレイアウトマスクで表される、請求項 4 に記載の方法。

【請求項 6】

前記ディスオクルージョンアトラスグループは、アトラスフレームのグループとして前記ボリュメトリックビデオ信号に符号化され、前記アトラスフレームのグループは、アトラス I フレームから開始し、異なるアトラス I フレームの前に終了する、請求項 4 または 5 に記載の方法。

20

【請求項 7】

前記ディスオクルージョンアトラスは、マスクストライピングを含み、前記マスクストライピングは、前記ディスオクルージョンアトラスに記憶された画像フラグメントが画像セグメントの 1 つまたは複数の境界で接していることを示す、請求項 1 から 6 のいずれか一項に記載の方法。

【請求項 8】

前記レイアウトマスクは、拡張前のサイズの前記レイアウトマスク内に最良適合ノードが見つからないと決定したことに応答して拡張される、請求項 1 から 7 のいずれか一項に記載の方法。

30

【請求項 9】

前記画像フラグメントについての空間領域は、ビットマスクで識別され、前記画像フラグメントは、前記画像フラグメントについての前記空間領域のサイズを使用してソートされる、請求項 1 から 8 のいずれか一項に記載の方法。

【請求項 10】

前記ディスオクルージョンアトラスに記憶された前記画像フラグメントは、前記視覚シーンから識別される 1 つまたは複数の顕著性領域に位置する、請求項 1 から 9 のいずれか一項に記載の方法。

【請求項 11】

40

方法であって、

請求項 1 から 10 のいずれか一項に記載の方法にしたがって、1 つまたは複数の基準画像およびディスオクルージョンアトラス内の画像フラグメントで符号化されたボリュメトリックビデオ信号を受信するステップと、

前記ボリュメトリックビデオ信号から、前記 1 つまたは複数の基準画像を復号するステップと、

前記ボリュメトリックビデオ信号から、前記ディスオクルージョンアトラス内の前記画像フラグメントを復号するステップと、

前記 1 つまたは複数の基準画像から、非表現ビューにおける表示画像を合成するステップと、

50

前記ディスオクルージョンアトラス内の前記画像フラグメントを使用して、前記表示画像におけるディスオクルードされた空間領域内のディスオクルードされた画像データを充填するステップと、

画像ディスプレイ上に前記表示画像をレンダリングするステップとを含む方法。

【請求項 1 2】

前記画像フラグメントの各々は、前記 1 つまたは複数の基準ビューにおいてオクルードされた連続するピクセルによって形成される空間領域について、テクスチャ画像値および深度画像値を含む、請求項 1 1 に記載の方法。

【請求項 1 3】

前記表示画像を合成することは、前記 1 つまたは複数の基準ビューに利用可能なテクスチャ画像値および深度画像値を使用することを含む、請求項 1 1 または 1 2 に記載の方法。

【請求項 1 4】

前記合成された表示画像内の前記ディスオクルードされた空間領域は、前記 1 つまたは複数の基準ビューに利用可能な前記テクスチャ画像値および前記深度画像値が、前記 1 つまたは複数の基準ビューに隣接する前記非基準ビューについて取得可能でないと決定することによって識別される、請求項 1 3 に記載の方法。

【請求項 1 5】

前記ディスオクルージョンアトラスに記憶された前記画像フラグメントは、前記視覚シーンから識別される 1 つまたは複数の顕著性領域に位置し、前記ディスオクルージョンアトラスは、前記合成された表示画像において 1 つまたは複数の顕著性領域が識別されるように、前記 1 つまたは複数の顕著性領域から離れた空間領域をカバーするためのテクスチャ画像値または深度画像値を含まない、請求項 1 2 または 1 3 に記載の方法。

【請求項 1 6】

前記ボリュメトリックビデオ信号は、単射関数を指定する画像メタデータを含み、前記単射関数は、前記画像フラグメント内の各ピクセルを、画像フレーム内の前記ピクセルのピクセルロケーションから、前記視覚シーンが表される 3 次元座標系内の対応するロケーションにマッピングする、請求項 1 から 1 5 のいずれか一項に記載の方法。

【請求項 1 7】

1 つまたは複数のプロセッサに請求項 1 から 1 6 のいずれか一項に記載の方法を実行させる、コンピュータプログラム。

【請求項 1 8】

1 つまたは複数のプロセッサと、命令のセットを記憶する 1 つまたは複数の記憶媒体とを備えるコンピューティングデバイスであって、命令は、1 つまたは複数のプロセッサによって実行されると、請求項 1 から 1 6 のいずれか一項に記載の方法を実行させる、コンピューティングデバイス。

【発明の詳細な説明】

【技術分野】

【0001】

[関連出願の相互参照]

本出願は、いずれも 2020 年 6 月 16 日に提出された米国仮出願第 63 / 039, 595 号および欧州特許出願第 20180179. 2 号の優先権を主張し、これらの各々は、その全体が参照により組み込まれる。

【0002】

[技術分野]

本発明は、一般に、画像コーディングおよびレンダリングに関し、詳細には、マルチビュービデオ動作をサポートするためにディスオクルージョンアトラス (disocclusion atlas) を使用することに関する。

【背景技術】

【0003】

10

20

30

40

50

ビュー合成は、３次元（３Ｄ）ＴＶ、３６０度ビデオ、ボリュメトリックビデオ、仮想現実（ＶＲ）、拡張現実（ＡＲ）などのアプリケーションにおいて使用される。仮想ビューは、関連する深度情報を用いて既存のビューから合成される。既存のビューは、描写された３Ｄワールドにワープまたはマッピングされ、次いで、ターゲットビュー位置に逆投影され得る。

【０００４】

結果として、既存のビューにおいて前景オブジェクトによってオクルード（occlude）される背景領域は、ターゲットビュー位置からターゲットビューでは（既存のビューからの利用可能な画像データなしに）ディスオクルード（disocclude）され得、それによって、ターゲットビューにおいてギャップまたは穴が生じる。加えて、深度画像（複数可）における不連続性もまた、合成ビューにおいてギャップまたは穴を生じさせ得る。ビデオ信号に符号化または送信されるべきビューの総数が、実際のビデオディスプレイアプリケーションにおいて低減または最小化されるにつれて、低減または最小化された数のビューから生成される合成ビュー内の穴の領域は、比較的大きくなり、増加するので、容易に知覚可能な視覚的アーティファクトが発生する。

【０００５】

このセクションで説明される手法は、追求され得る手法であるが、必ずしも以前に考案または追求された手法であるとは限らない。したがって、別段の指示がない限り、このセクションで説明されるいずれの手法も、単にこのセクションに含まれるという理由で従来技術とみなされると仮定されるべきではない。同様に、１つまたは複数の手法に関して特定された問題は、別段の指示がない限り、このセクションに基づいて任意の従来技術において認識されていると仮定されるべきではない。

【図面の簡単な説明】

【０００６】

本発明は、添付の図面の図において、限定としてではなく、例として示され、同様の参照番号は同様の要素を指す。

【図１Ａ】例示的なボリュメトリックビデオストリーミングを示す図である。

【図１Ｂ】例示的なボリュメトリックビデオストリーミングを示す図である。

【図２Ａ】例示的なテクスチャおよび深度画像を示す。

【図２Ｂ】例示的なテクスチャおよび深度画像を示す。

【図２Ｃ】基準ビュー（reference view）ではオクルードされ、隣接するビューでは少なくとも部分的にディスオクルードされるようになる空間領域（spatial region）を識別するための例示的な画像マスクを示す。

【図３Ａ】例示的なディスオクルージョンアトラスを示す。

【図３Ｂ】連続するディスオクルージョンアトラスの例示的なシーケンスを示す。

【図３Ｃ】共通のグループレベルのレイアウトマスク（layout mask）を用いて生成された連続するディスオクルージョンアトラスの例示的なグループを示す。

【図４Ａ】例示的なプロセスフローを示す。

【図４Ｂ】例示的なプロセスフローを示す。

【図４Ｃ】例示的なプロセスフローを示す。

【図５】本明細書で説明されるコンピュータまたはコンピューティングデバイスが実装され得る例示的なハードウェアプラットフォームを示す。

【発明を実施するための形態】

【０００７】

マルチビュービデオ動作をサポートするためにディスオクルージョンアトラスを使用することに関する例示的な実施形態が本明細書で説明される。以下の説明では、説明の目的で、本発明の完全な理解を与えるために多数の具体的な詳細が記載される。しかしながら、本発明がこれらの具体的な詳細なしに実施され得ることは明らかであろう。他の事例では、本発明を不必要に隠蔽する、不明瞭にする、または難解にすることを回避するために、周知の構造およびデバイスは網羅的に詳細には説明されない。

【 0 0 0 8 】

例示的な実施形態は、以下の概要にしたがって本明細書で説明される：

- 1 . 概要
- 2 . ボリュメトリックビデオ
- 3 . 例示的なビデオストリーミングサーバおよびクライアント
- 4 . ディスオクルージョンデータ内の画像フラグメント
- 5 . ディスオクルージョンデータのための画像マスク
- 6 . ディスオクルージョンアトラスの生成
- 7 . 時間的に安定したグループレベルのレイアウトマスク
- 8 . 例示的なプロセスフロー
- 9 . 実装機構 - ハードウェアの概要
- 1 0 . 同等物、拡張物、代替物、およびその他

10

【 0 0 0 9 】

1 . 概要

この概要は、本発明の例示的な実施形態のいくつかの態様の基本的な説明を提示する。この概要は、例示的な実施形態の態様の広範なまたは網羅的な要約ではないことに留意されたい。さらに、この概要は、例示的な実施形態の特に重要な態様または要素を識別するものとして理解されることも、特に例示的な実施形態の範囲または概して本発明を定めるものとして理解されることも意図されていないことに留意されたい。この概要は、例示的な実施形態に関連するいくつかの概念を簡略化された形式で提示しているにすぎず、以下

に続く例示的な実施形態のより詳細な説明への単なる概念的な前置きとして理解されるべきである。本明細書では別個の実施形態について述べるが、本明細書で述べる実施形態および/または部分的な実施形態の任意の組合せを組み合わせ、さらなる実施形態が形成され得ることに留意されたい。

20

【 0 0 1 0 】

ボリュメトリックビデオを送信するための一般的な手法は、広視野（多くの場合360度）のキャプチャまたはレンダリングされた画像に、有限セットのビュー位置（「記録されたビュー（recorded view）」、「基準ビュー（reference view）」、または「表現ビュー（represented view）」とも呼ばれる）からの深度を付加すること（accompany）ことである。各ピクセルにおける深度値により、これらのピクセルは、典型的には、記録されたビュー位置（または基準ビュー）の間にある推定ビューに再投影（およびzバッファリング）されることができる。記録されたビュー位置における記録された画像から合成された変形画像（warped image）などの単一の再投影されたビュー画像は、記録された画像で表される元の視点からは見えないディスオクルードされた領域に対応する穴およびギャップを有することになる。より多くの周囲のソースの視点またはより多くの記録されたビュー位置を追加することによって、再投影されたビュー画像内に残る穴は少なくなり得るが、大量の冗長データ（例えば、追加された複数の記録されたビューの各々内で可視のピクセルなど）を犠牲にすることになる。

30

【 0 0 1 1 】

比較として、本明細書で説明される技法は、アトラス表現で比較的少量のディスオクルージョンデータ（disocclusion data）を送信するために使用され得る。ディスオクルージョンデータは、単一の（最も近い）基準ビュー位置からは見えないフラグメントについてのテクスチャおよび深度情報のみを含み、それにより、追加の記録されたビューにおける冗長性を回避し、ビデオストリーミングおよび復号の際のデータ量を大幅に低減する。これらの技法を使用して、（例えば、矩形、正方形などの）結合画像（combined image）内に画像フラグメント（image fragment）をレイアウトすることで、結合画像内にできるだけ空きスペースを残さないようにすることができる。

40

【 0 0 1 2 】

さらに、異なる連続するアトラスレイアウトにおけるフレーム間の時間的変動から生じるビデオ圧縮効率の問題が、本明細書で説明される技法によって効果的に対処され、動き

50

予測（例えば、インター予測など）が向上し得る。例えば、アトラス「Iフレーム」（動き予測なしで符号化または復号することができるフレーム）間の連続するディスオクルージョンアトラスのレイアウトを時間的に安定させて、比較的高い効率の圧縮率を達成することができる。

【0013】

いくつかの動作シナリオでは、マルチビュービデオの1つまたは複数の表現ビューに対応する（またはそれからの画像データを含む）1つまたは複数のビデオストリームは、1つまたは複数の表現ビューについてのディスオクルージョンデータと一緒にまたは別個に受信者ビデオデコーダに送信され得る。ディスオクルージョンデータは、ビデオストリーム内の表現ビューにおいて隠されているかまたはオクルードされる可能性のある画像詳細のためのテクスチャおよび/または深度画像データを含む。ディスオクルージョンデータによって描写されるオクルードされた画像詳細の一部は、ビデオストリーム内の表現ビューのうちの1つまたは複数の隣接するビューア（viewer）の現在のビュー（「仮想ビュー」または「ターゲットビュー」とも呼ばれる）では可視になり得る。

10

【0014】

上述したように、ディスオクルージョンデータは、ディスオクルージョンアトラスにパッケージ化または符号化することができる。ディスオクルージョンアトラスは、1つまたは複数の深度における可視の画像詳細および他の深度におけるオクルードされた画像詳細などのマルチ深度情報（場合によっては複数の表現ビューについての）を、表現ビューのビデオストリームを含むボリュメトリックビデオ信号（volumetric video signal）に符号化することをサポートするために、ビデオエンコーダによって使用され得る。ディスオクルージョンアトラスは、表現ビューのうちの1つまたは複数の隣接するビューアの現在のビューに固有の画像詳細などのビュー依存効果をレンダリングするために、ビデオ信号の受信者ビデオデコーダによって使用され得る。

20

【0015】

ボリュメトリックビデオ信号は、受信者ビデオデコーダが、ビデオストリーム内の表現ビューの画像データを使用してビューアの現在のビューに固有の画像をレンダリングするのに支援するために、画像メタデータの一部としてディスオクルージョンアトラスを含むことができる。ビデオストリームおよび画像メタデータは、限定はしないが、MPEG（Moving Picture Experts Group）ビデオ標準、H.264/AVC（H.264/Advanced Video Coding）、HEVC（High-Efficiency Video Coding）、MPEG-I、ドルビーのVIXファイルフォーマットなどを含むが、それに限定されない、ビデオコーディング標準またはプロプライエタリ規格に基づくコーディングシンタックスで符号化され得る。追加的に、オプションで、または代替的に、ディスオクルージョンアトラスは、表現ビューの画像データを含むビデオストリームに付随するサブストリームに符号化され、サブストリームから復号され得る。

30

【0016】

受信者ビデオデコーダは、ボリュメトリックビデオ信号によって搬送される画像メタデータ（またはサブストリーム）内のディスオクルージョンアトラスにパックされたディスオクルージョンデータと、ボリュメトリックビデオ信号に符号化されたビデオストリーム内の表現ビューの画像データとを復号することができる。ディスオクルージョンデータおよび画像データは、表現ビューのうちの1つまたは複数の隣接するビューアの現在のビューの画像を生成または構築する際に穴またはギャップを充填するためにビデオデコーダによって使用され得る。ビューアの現在のビューは、ビデオストリーム内のいずれの表現ビューとも一致しないことがあり、ビューアの現在のビュー（またはビュー位置）の画像は、画像ワーピング動作を通して、表現ビューの受信された画像から取得され得る。例示的な画像ワーピングおよび/または合成動作は、2017年6月12日に出願された米国仮特許出願第62/518,187号に記載されており、その内容全体は、本明細書に完全に記載されているかのように、参照によって本明細書に組み込まれる。

40

【0017】

50

変形画像内の穴またはギャップを充填するために、例えば、効率的なルックアップ動作またはインデックスによる探索動作により、ディスオクルージョンアトラス内のディスオクルージョンデータの一部または全部にアクセスして取り出し、表現ビューではオクルードされているが、ビューアの現在のビューではディスオクルードされている画像詳細を提供し得る。結果として、ビューアは、ポリュメトリックビデオ信号のビデオストリームに符号化された表現ビューの画像では提供されないビュー固有の画像詳細を、ビューアの現在のビューにしたがって、見ることができる。

【 0 0 1 8 】

本明細書で説明される例示的な実施形態は、ポリュメトリックビデオのストリーミングに関する。1つまたは複数の基準ビューからの視覚シーンを描写する1つまたは複数の基準画像 (reference image) ではオクルードされ、1つまたは複数の基準ビューに隣接する非基準ビューでは少なくとも部分的にディスオクルードされる画像フラグメントが、サイズでソートされる。画像フラグメントは、画像フラグメント内の任意の他の画像フラグメント以上のサイズである第1の画像フラグメントを含む。画像フラグメントを記憶するために使用されるディスオクルージョンアトラスのためのレイアウトマスクが生成される。レイアウトマスクは、第1の画像フラグメントに対して特にサイズ決定にされた第1の最良適合ノード (best fit node) を含む四分木でカバーされる。ソートされた画像フラグメントは、レイアウトマスク内で識別された最良適合ノードに降順で記憶される。ソートされた画像フラグメント内の各画像フラグメントは、最良適合ノード内のそれぞれの最良適合ノードに記憶される。最良適合ノードは、レイアウトマスクをカバーする四分木内の少なくとも1つのノードを反復的に分割することによって取得される少なくとも1つの最良適合ノードを含む。1つまたは複数の基準画像で符号化されたポリュメトリックビデオ信号が生成される。ポリュメトリックビデオ信号は、ディスオクルージョンアトラス内の画像フラグメントを用いてさらに符号化される。1つまたは複数の基準画像は、画像ディスプレイ上にレンダリングするための非表現ビューにおける表示画像を合成するためにポリュメトリックビデオ信号の受信者デバイスによって使用される。ディスオクルージョンアトラス内の画像フラグメントは、表示画像におけるディスオクルードされた空間領域内のディスオクルードされた画像データを充填するために受信者デバイスによって使用される。

【 0 0 1 9 】

本明細書で説明される例示的な実施形態は、ポリュメトリックビデオのレンダリングに関する。ポリュメトリックビデオ信号から1つまたは複数の基準画像が復号される。ポリュメトリックビデオ信号から、ディスオクルージョンアトラス内の画像フラグメントが復号される。1つまたは複数の基準画像に基づいて、非表現ビューにおける表示画像が合成される。ディスオクルージョンアトラス内の画像フラグメントを使用して、表示画像におけるディスオクルードされた空間領域内のディスオクルードされた画像データを充填する。表示画像は、画像ディスプレイ上にレンダリングされる。

【 0 0 2 0 】

いくつかの例示的な実施形態では、本明細書で説明される機構は、クラウドベースのサーバ、モバイルデバイス、仮想現実システム、拡張現実システム、ヘッドアップディスプレイデバイス、ヘルメットマウントディスプレイデバイス、CAVEタイプシステム、ウォールサイズディスプレイ、ビデオゲームデバイス、ディスプレイデバイス、メディアプレーヤ、メディアサーバ、メディア生成システム、カメラシステム、家庭用システム、通信デバイス、ビデオ処理システム、ビデオコーデックシステム、スタジオシステム、ストリーミングサーバ、クラウドベースのコンテンツサービスシステム、ハンドヘルドデバイス、ゲーム機、テレビ、シネマディスプレイ、ラップトップコンピュータ、ネットブックコンピュータ、タブレットコンピュータ、セルラー無線電話、電子書籍リーダー、POS (point of sale) 端末、デスクトップコンピュータ、コンピュータワークステーション、コンピュータサーバ、コンピュータキオスク、または様々な他の種類の端末およびメディア処理ユニットのうちのいずれかを含むが、それに限定されないメディア処理システムの

10

20

30

40

50

一部を形成する。

【 0 0 2 1 】

本明細書に記載された好ましい実施形態ならびに一般的な原理および特徴に対する様々な修正は、当業者には容易に明らかになるであろう。したがって、本開示は、示される実施形態に限定されることを意図するものではなく、本明細書で説明される原理および特徴と一致する最も広い範囲を与えられるべきである。

【 0 0 2 2 】

2. ボリュメトリックビデオ

本明細書で説明される技法は、最大で6自由度すべてのビューアの身体または頭部の動きに対応したビューアに対する完全な視差をもつビュー固有のビデオを提供するために使用され得る。本明細書で使用する場合、「ビュー固有の」ビデオ（画像）という用語は、ビューアの位置および／または向きに少なくとも部分的に基づいて（またはその決定に回答して）生成および／またはレンダリングされる、位置固有および／または向き固有のビデオ（画像）を意味し得る。

10

【 0 0 2 3 】

これを達成するために、空間内の異なる点のセットまたはサブセット（ビューアが自由に動くことができるビューボリューム（viewing volume）にわたる異なる位置および／または異なる向きのセットまたはサブセットに対応する）におけるビデオは、ビューアにレンダリングされるビュー固有の画像を生成するために、使用され得る。空間内のこれらの異なる点におけるビデオは、テクスチャビデオおよび深度ビデオを含み、ボリュメトリックビデオの基準ビュー（または基準視点）を形成し得る。

20

【 0 0 2 4 】

ビューアの所与の位置および／または向きについてのビューアの現在のビューなどの仮想ビュー（これは、これらの基準ビューのどれとも一致しないことがある）は、画像ベースのレンダリング技法を使用して、ボリュメトリックビデオで表されるこれらの基準ビューから合成され得る。

【 0 0 2 5 】

本明細書で使用される場合、テクスチャビデオは、複数の時点にわたるテクスチャ画像のシーケンスを指し、RGBピクセル値、YCbCrピクセル値、ルーマおよび／またはクロマピクセル値などの個々の色または輝度情報でそれぞれ指定されたピクセルの空間分布を含む。テクスチャビデオに対応する深度ビデオは、複数の時点にわたる深度画像のシーケンスを指し、z軸値、深度値、空間視差値、視差情報などの対応するテクスチャ画像の対応するピクセルの空間深度情報でそれぞれ指定されたピクセルの空間分布を含む。

30

【 0 0 2 6 】

ボリュメトリックビデオ内の1つまたは複数のビデオストリームで表される1つまたは複数の基準ビューに対するディスオクルージョンデータを含むディスオクルージョンアトラスは、ビュー依存効果のためのマルチ深度情報を符号化することをサポートするために使用され得る。例えば、ハイライトスペックルなどの画像詳細は、すべてではないがいくつかのビューに現れ得、可視であるとき、異なるビュー（例えば、異なる基準ビュー、異なる時点におけるビューアの現在のビューなどの異なる仮想ビューなど）において異なって現れる。基準ビューにおいて隠されているかまたはオクルードされるビュー依存画像詳細のマルチ深度情報は、ディスオクルージョンデータに含まれ、画像メタデータの一部として受信者ビデオデコードに配信され、それにより、ビューアの位置または向きの検出された変化に回答してビュー依存画像詳細（または効果）をビューアに正しくレンダリングまたは提示することができる。

40

【 0 0 2 7 】

追加的に、オプションで、または代替的に、画像メタデータは、本明細書で説明されるディスオクルージョンアトラス内のフラグメント、部分、パッチなどの記述を含み得る。画像メタデータは、ボリュメトリックビデオの一部として上流デバイスから受信者デバイスに配信され、受信者デバイスがビデオストリームおよびディスオクルージョンアトラス

50

から復号された画像データをレンダリングするのを助けるために使用され得る。

【0028】

3. 例示的なビデオストリーミングサーバおよびクライアント

図1Aは、マルチビューストリーム受信機132、視点プロセッサ134、ストリームコンポーザ136などを備えるビデオストリーミングサーバ100などの例示的な上流デバイスを示す。ビデオストリーミングサーバ(100)の構成要素の一部または全部は、1つまたは複数のデバイス、モジュール、ユニットなどによって、ソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどで実装され得る。

【0029】

マルチビューストリーム受信機(132)は、外部ビデオソースから直接または間接的に複数の基準ビューの基準テクスチャおよび/または深度ビデオ(106)を受信するように構成されたソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどを含む。

10

【0030】

視点プロセッサ(134)は、リアルタイムまたは準リアルタイムでビューアによって動作されるビデオクライアントデバイスからビューアの視点データを受信し、AR、VRまたはポリュメトリックビデオアプリケーションの時間間隔/持続時間にわたって複数の時点についてビューアの位置または向きを確立/決定するように構成されたソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどを含む。ビデオアプリケーションでは、基準テクスチャおよび/または深度ビデオ(106)から導出された表示画像は、ビデオクライアントデバイスなどと連動して動作する画像ディスプレイで提供されるように、ビューアのビューポート内の複数の時点でレンダリングされるべきである。ビューアのビューポートは、画像ディスプレイ上のウィンドウまたは可視領域のサイズを指す。

20

【0031】

ストリームコンポーザ(136)は、受信者デバイスなどから入力の一部として受信されたビューアの位置または向きを示す視点データ114に少なくとも部分的に基づいて、基準テクスチャおよび/または深度ビデオ(106)から、(例えば、リアルタイムなどの)ポリュメトリックビデオ信号112(1つまたは複数の基準ビューと、表現ビューに隣接するビューのディスオクルージョンアトラスを含むディスオクルージョンアトラスとを表す1つまたは複数のビデオストリームを含むが、それに限定されない)を生成するように構成されたソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどを含む。

30

【0032】

ビデオストリーミングサーバ(100)は、ARアプリケーション、VRアプリケーション、360度ビデオアプリケーション、ポリュメトリックビデオアプリケーション、リアルタイムビデオアプリケーション、準リアルタイムビデオアプリケーション、非リアルタイム全方向ビデオアプリケーション、自動車エンターテインメント、ヘルメットマウントディスプレイアプリケーション、ヘッドアップディスプレイアプリケーション、ゲーム、2Dディスプレイアプリケーション、3Dディスプレイアプリケーション、マルチビューディスプレイアプリケーションなどをサポートするために使用され得る。

40

【0033】

図1Bは、リアルタイムストリーム受信機142、視点トラッカ144、ポリュメトリックビデオレンダラ146、画像ディスプレイ148などを備えるビデオクライアントデバイス150などの例示的な受信者デバイスを示す。ビデオクライアントデバイス(150)の構成要素の一部または全部は、1つまたは複数のデバイス、モジュール、ユニットなどによって、ソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどで実装され得る。

【0034】

視点トラッカ(144)は、1つまたは複数のビューア位置/向きトラッキングセンサ(例えば、動きセンサ、位置センサ、アイトラッカなど)とともに動作して、ビューアに

50

関連するリアルタイムまたは準リアルタイムの視点データ 114 を収集し、視点データ (114) または視点データから決定されたビューアの位置 / 向きをビデオストリーミングサーバ (100) に送信することなどを行うように構成されたソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどを含む。視点データ (114) は、比較的細かい時間スケール (例えば、ミリ秒毎、5 ミリ秒毎など) でサンプリングまたは測定され得る。視点データは、所与の時間分解能 (例えば、ミリ秒毎、5 ミリ秒毎など) でビューアの位置また向きを確立 / 決定するために使用されることができる。

【0035】

リアルタイムストリーム受信機 (142) は、(例えば、リアルタイムなどの) ボリュームメトリックビデオ信号 (112) を受信し、復号するように構成されたソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどを含む。

10

【0036】

ボリュームメトリックビデオレンダラ (146) は、ボリュームメトリックビデオ (112) から復号された画像データに対して画像ワーピング、画像ワーピング、ブレンディング (例えば、複数のカメラソースからの複数の変形画像のブレンディングなど)、画像合成、穴埋めなどを実行して、ビューアの予測または測定された位置または向きに対応するビュー固有の画像を生成し、レンダリングのためにビュー固有の画像を画像ディスプレイ (148) に出力することなどを行うように構成されたソフトウェア、ハードウェア、ソフトウェアとハードウェアの組合せなどを含む。

【0037】

20

本明細書で使用される場合、本明細書で説明されるビデオストリーム内のビデオコンテンツは、視聴覚プログラム、映画、ビデオプログラム、TV 放送、コンピュータゲーム、拡張現実 (AR) コンテンツ、仮想現実 (VR) コンテンツ、自動車エンターテインメントコンテンツなどのいずれかを含み得るが、必ずしもそれらに限定されない。例示的なビデオデコーダは、ディスプレイデバイス、ニアアイディスプレイを有するコンピューティングデバイス、頭部装着型ディスプレイ (HMD)、モバイルデバイス、ウェアラブルディスプレイデバイス、テレビなどのディスプレイを有するセットトップボックス、ビデオモニタなどのうちのいずれかを含み得るが、必ずしもそれらに限定されない。

【0038】

本明細書で使用される場合、「ビデオストリーミングサーバ (video streaming server)」は、ビデオコンテンツの少なくとも一部分を 1 つまたは複数のディスプレイ上にレンダリングするために、ビデオコンテンツを準備し、ビデオデコーダなどの 1 つまたは複数のビデオストリーミングクライアントにストリーミングする 1 つまたは複数の上流デバイスを指し得る。ビデオコンテンツがレンダリングされるディスプレイは、1 つまたは複数のビデオストリーミングクライアントの一部であってもよいし、1 つまたは複数のビデオストリーミングクライアントと連動して動作していてもよい。

30

【0039】

例示的なビデオストリーミングサーバは、ビデオストリーミングクライアント (複数可) から遠隔に位置するクラウドベースのビデオストリーミングサーバ、ローカルのワイヤードまたはワイヤレスネットワーク上でビデオストリーミングクライアント (複数可) に接続されたローカルビデオストリーミングサーバ、VR デバイス、AR デバイス、自動車エンターテインメントデバイス、デジタルメディアデバイス、デジタルメディア受信機、セットトップボックス、ゲーム機 (例えば、Xbox)、汎用パーソナルコンピュータ、タブレット、Apple TV または Roku ボックスなどの専用デジタルメディア受信機などのうちのいずれかを含み得るが、必ずしもそれらに限定されない。

40

【0040】

4. ディスオクルージョンデータ内の画像フラグメント

ディスオクルージョンアトラス内のディスオクルージョンデータは、ボリュームメトリックビデオ信号内の表現 (基準) ビューにおいてオクルードされる画像フラグメントを含み得る。本明細書で説明される画像フラグメントは、ピクセルごとの画像テクスチャ情報 (例

50

例えば、色、輝度／クロミナンス値、RGB値、YCbCr値など）とピクセルごとの深度情報とを有するピクセルの連続した非凸（またはオクルードされた）領域を指す。ディスオクルージョンアトラス内の画像フラグメントに対して指定されたピクセルごとの画像テクスチャおよび深度情報は、ポリュメトリックビデオ信号の表現ビューでは、隠されているかまたはオクルードされる画像特徴／オブジェクト／構造を視覚的に描写し得るが、表現ビューに隣接するビューでは、少なくとも部分的にディスオクルードされるようになるかまたは見えるようになり得る。

【0041】

画像テクスチャおよび深度情報が欠けている穴を含んでいない所与の基準ビューについて、深度画像ベースレンダリング（DIBR: depth image based rendering）と、基準ビューに利用可能な画像テクスチャ／深度情報とを使用して、表現ビューの周りの隣接するビューについての合成画像（synthesized image）が生成され得る。合成画像は、基準ビューに利用可能な画像テクスチャ／深度情報から画像テクスチャ情報および深度情報を取得することができない穴を有し得る。合成画像を用いて、画像マスクを生成して、隣接するビューについての合成画像内の穴を識別することができる。

10

【0042】

いくつかの動作シナリオでは、画像マスクは、隣接するピクセル間の深度の比較的滑らかな遷移を有する所与の基準ビュー内の他の画像領域から、隣接するピクセル間の深度の大きいギャップを含む所与の基準ビュー内の画像領域（またはエリア）を識別することによって、所与の基準ビューについて少なくとも部分的に生成され得る。

20

【0043】

画像マスクにおいて識別された穴（またはピクセルの非凸領域）内の画像フラグメントについての画像テクスチャ情報および深度情報は、空間的に異なる基準ビューから、または時間的に異なる基準ビューから取得され得る。例えば、所与の基準ビューと同じ時点のものであるが、所与の基準ビューとは空間的に異なる空間的に異なる基準ビューは、隣接するビューにおける合成画像内の穴についての画像テクスチャおよび深度情報を含んでおり、提供し得る。所与の基準ビューを含むこれらの空間的に異なる基準ビューは、同じ時点についてのマルチビュー画像を集合的に形成し得る。

【0044】

30

追加的に、オプションで、または代替的に、所与の基準ビューのものとは異なる時点についてのものである時間的に異なる基準ビューは、隣接するビューにおける合成画像内の穴についての画像テクスチャおよび深度情報を含んでおり、提供し得る。所与の基準ビューを含むこれらの時間的に異なる基準ビューは、同じ視覚シーン、同じピクチャグループ（GOP）などに属し得る。

【0045】

追加的に、オプションで、または代替的に、人工知能（AI）または機械学習（ML）は、画像をトレーニングすることによってトレーニングされ、次いで、隣接するビューにおける合成画像内の穴についての画像テクスチャおよび深度情報の一部または全部を生成または予測するために適用され得る。

40

【0046】

所与の時点についてのディスオクルージョンアトラスに含まれる画像フラグメントは、異なる基準ビューについての画像フラグメントの異なるサブセットに区分され得る。異なるサブセット内の画像フラグメントの各サブセットは、異なる基準ビューのそれぞれの基準ビュー内の（オクルードされた）画像フラグメントを含み得る。

【0047】

本明細書で説明されるディスオクルージョンアトラス技法は、これらの画像フラグメントを、最小総面積（minimal total area）をカバーし、かつ、オーバーラップするフラグメントを有しない結合画像（または「アトラス」）へと、（例えば、適応的に、最適に、など）パックするために使用され得る。ディスオクルージョンアトラスを表す結合画像

50

内の各フラグメントは、ディスオクルージョンアトラスに含まれる他のフラグメントによってオーバーラップされることなく専用の領域（またはエリア）を有する。

【 0 0 4 8 】

ボリュメトリックビデオ信号は、連続するマルチビュー画像のシーケンスから生成され得る。連続するマルチビュー画像のシーケンスは、連続する時点のシーケンスを形成する複数の時点についての複数のマルチビュー画像を含む。複数のマルチビュー画像内の各マルチビュー画像は、複数の時点内のそれぞれの時点についての複数の基準ビューに対する複数の単一ビュー画像を含む。

【 0 0 4 9 】

連続するディスオクルージョンアトラスのシーケンスは、連続する時点のシーケンスに対して生成され得る。連続するディスオクルージョンアトラスのシーケンスは、連続する時点のシーケンス内の複数の時点についての複数のディスオクルージョンアトラスを含む。複数のディスオクルージョンアトラス内の各ディスオクルージョンアトラスは、複数の時点内のそれぞれの時点についてボリュメトリックビデオ信号で表される、複数の基準ビュー内の1つまたは複数の基準ビューに対する画像フラグメントの1つまたは複数のサブセットを含む画像フラグメントを含む。

10

【 0 0 5 0 】

連続する時点のシーケンスによってカバーされる時間間隔（例えば、30分、1時間以上など）内の部分間隔（例えば、1秒の数分の1、1秒以上など）について、ボリュメトリックビデオ信号は、この信号で表される1つまたは複数の基準ビューに対するピクチャグループ（GOP）の1つまたは複数のサブシーケンスで符号化され得る。GOPの1つまたは複数のサブシーケンス内のGOPの各サブシーケンスは、ボリュメトリックビデオ信号で表される1つまたは複数の基準ビュー内のそれぞれの基準ビューについてのテクスチャ画像のサブシーケンスおよび深度画像のサブシーケンスを含む。

20

【 0 0 5 1 】

GOPの各サブシーケンスは、1つまたは複数のGOPを含む。各GOPは、Iフレームによって区切られるか、または開始Iフレームで始まり、次の開始Iフレームの直前のフレームで終わる。いくつかの実施形態では、開始Iフレームおよび次の開始Iフレームは、それらの間に他のIフレームがない2つの最も近いIフレームであってもよい。いくつかの実施形態では、開始Iフレームおよび次の開始Iフレームは、近くのIフレームであり得るが、必ずしも2つの最も近いIフレームであるとは限らない。GOP内のIフレームは、他のフレームからの画像データに依存することなく復号され得るが、GOP内のBフレームまたはPフレームなどの非Iフレームは、GOP内の他のフレームから少なくとも部分的に予測され得る。GOP内のIフレーム（複数可）および/または非Iフレーム（複数可）は、時間的に安定したまたは時間的に類似したソース/入力画像から生成され得る。これらの時間的に安定したソース/入力画像は、GOP内のIフレーム（複数可）および/または非Iフレーム（複数可）を生成する際に、比較的効率的なインター予測またはイントラ予測、およびデータ圧縮または符号化を容易にすることができる。

30

【 0 0 5 2 】

連続する時点のシーケンスによってカバーされる間隔内の同じ部分間隔について、ボリュメトリックビデオ信号は、信号で表される1つまたは複数の基準ビューに対するディスオクルージョンアトラスのグループの1つまたは複数のサブシーケンスで符号化され得る。ディスオクルージョンアトラスのグループの1つまたは複数のサブシーケンス内のディスオクルージョンアトラスのグループの各サブシーケンスは、ボリュメトリックビデオ信号で表される1つまたは複数の基準ビュー内のそれぞれの基準ビューに隣接するビュー内の穴についてのテクスチャ画像のサブシーケンスおよび深度画像のサブシーケンスを含む。

40

【 0 0 5 3 】

ディスオクルージョンアトラスのグループの各サブシーケンスは、ディスオクルージョンアトラスの1つまたは複数のグループを含む。ディスオクルージョンアトラスの各グループは、アトラスIフレームによって区切られるか、または開始アトラスIフレームで始

50

まり、次の開始アトラスIフレームの直前のアトラスフレームで終わる。いくつかの実施形態では、開始アトラスIフレームおよび次の開始アトラスIフレームは、それらの間に他のアトラスIフレームがない2つの最も近いアトラスIフレームであってもよい。いくつかの実施形態では、開始アトラスIフレームおよび次の開始アトラスIフレームは、近くのアトラスIフレームであってもよいが、必ずしも2つの最も近いアトラスIフレームである必要はない。ディスオクルージョンアトラスのグループ内のアトラスIフレームは、他のアトラスフレームからのディスオクルージョンデータに依存することなく復号され得るが、ディスオクルージョンアトラスのグループ内のアトラスBフレームまたはアトラスPフレームなどのアトラス非Iフレームは、ディスオクルージョンアトラスのグループ内の他のアトラスフレームから少なくとも部分的に予測され得る。ディスオクルージョンアトラスのグループ内のアトラスIフレーム（複数可）および/またはアトラス非Iフレーム（複数可）は、時間的に安定したまたは時間的に類似したディスオクルージョンアトラスから生成され得る。これらの時間的に安定したディスオクルージョンアトラスは、ディスオクルージョンアトラスのグループ内のアトラスIフレーム（複数可）および/またはアトラス非Iフレーム（複数可）を生成する際に、比較的効率的なインター予測またはイントラ予測、およびデータ圧縮または符号化を容易にすることができる。

10

【0054】

5. ディスオクルージョンデータのための画像マスク

図2Aは、基準ビューにおける例示的なテクスチャ画像（例えば、360度の「ベースボールカバー」ビューなど）を示す。テクスチャ画像は、画像フレーム内のピクセルのアレ

20

【0055】

図2Bは、図2Aのテクスチャ画像と同じ基準ビューにおける例示的な深度画像（例えば、360度の「ベースボールカバー」ビューなど）を示す。図2Bの深度画像は、図2Aのテクスチャ画像内のピクセルのアレ

30

【0056】

図2Cは、ビットのアレ

40

【0057】

例示的な穴埋め動作は、Wenhui Jiaらによる、2019年4月1日に出願された「HOLE FILLING FOR DEPTH IMAGE BASED RENDERING」と題する米国仮特許出願第62/811,956号に記載されており、その全体の内容は、本明細書に

50

完全に記載されているかのように参照により本明細書に組み込まれる。

【 0 0 5 8 】

画像ワーピングおよび穴埋め動作は、基準ビューに隣接するビューであり得る、ビューアの現在のビューのための合成画像を生成するために使用され得る。ディスオクルージョンアトラスにおいて提供されるディスオクルードされたピクセルテクスチャ値および/またはディスオクルードされたピクセル深度値は、図 2 A のテクスチャ画像および/または図 2 B の深度画像ではオクルードされるが、基準ビューに隣接するビューでは部分的に見えるようになり得る画像詳細を描写する。ディスオクルージョンアトラスは、連続する時点のシーケンスによってカバーされる時間間隔内の同じ時点に対応し得るか、またはそれによってインデックス付けされ得、例えば、ビデオストリームまたは別個の付随のビデオストリーム内のディスオクルージョンアトラスのグループ内のアトラス I フレームまたはアトラス非 I フレームとして、基準ビューのためのビデオストリームまたは別個の付随のビデオストリームに符号化され得る。

10

【 0 0 5 9 】

図 2 C に示されるような画像マスクは、マスクが、基準ビューに隣接する 1 つまたは複数の隣接するビューから見えないテクスチャ画像および/または深度画像の部分をカバーするので、図 2 A の対応するテクスチャ画像または図 2 B の対応する深度画像とアラインするようには見えない。画像マスクを用いて生成されたディスオクルージョンアトラスの目的は、ビューアの現在のビューなどの合成ビュー内の穴を充填するためにテクスチャおよび深度画像データを提供することであり、穴は、合成ビュー（または選択された「基準」ビュー）の再投影におけるディスオクルージョンに起因して生じる。様々な動作シナリオでは、ディスオクルージョンアトラス内のテクスチャおよび深度データは、合成ビュー内の穴よりも多いまたは少ない、または同じ空間領域をカバーし得る。

20

【 0 0 6 0 】

いくつかの動作シナリオでは、ディスオクルージョンアトラスによってカバーされる空間領域は安全マージンを含み得、それにより、ディスオクルージョンアトラスは、ディスオクルージョンアトラス内のディスオクルードされたテクスチャおよび深度データが基準ビューに隣接するビュー内の穴を完全に充填するために利用可能であることを保証することができる。

【 0 0 6 1 】

いくつかの動作シナリオでは、ディスオクルージョンによってカバーされる空間領域は安全マージンを含まなくてもよく、その結果、ディスオクルージョンアトラスは、ディスオクルージョンアトラス内のディスオクルードされたテクスチャおよび深度データが基準ビューに隣接するビュー内の穴を完全に充填するために利用可能であることを保証しない可能性がある。これらの動作シナリオでは、受信者ビデオデコードは、穴埋めアルゴリズムを適用して、ビデオストリームで表される基準ビューに隣接または近接する合成ビュー内の穴の一部についてのテクスチャおよび深度情報の少なくとも一部を生成し得る。

30

【 0 0 6 2 】

追加的に、オプションで、または代替的に、ディスオクルージョンアトラスにおいてカバーされるマスクされた空間領域は、基準ビューにおいて描写される視覚シーンから顕著な視覚オブジェクトを選択するために使用され得る。例えば、ディスオクルージョンアトラスは、顕著な視覚オブジェクトから離れた空間領域をカバーするためのテクスチャまたは深度情報を受信者ビデオデコードに搬送または提供しないことがある。ディスオクルージョンアトラスが受信者ビデオデコードテクスチャまたは深度情報を搬送または提供する空間領域は、空間領域が顕著な視覚オブジェクトを含むことを受信者ビデオデコードに示し得る。

40

【 0 0 6 3 】

6 . ディスオクルージョンアトラスの生成

図 3 A は、1 つまたは複数の基準ビューについてのオクルードされた領域を表す画像フラグメントを含むか、またはそれとともにパッケージ化された例示的な（出力）ディスオ

50

クルージョンアトラスを示す。画像メタデータは、ディスオクルージョンアトラス内のこれらの画像フラグメントがそれぞれどの基準ビューに対応するかを示すために生成され得る。

【 0 0 6 4 】

例として、ポリュメトリックビデオ信号は、マルチビュー画像のシーケンスから生成される。マルチビュー画像のシーケンス内の各マルチビュー画像は、N個の基準ビューについてのN個の単一ビュー（入力／ソース）テクスチャ画像の集合と、連続する時点のシーケンス内のある時点についてのN個の基準ビューについてのN個の単一ビュー（入力／ソース）深度画像の集合とを含み得る。

【 0 0 6 5 】

ビューパラメータが受信され、画像（ピクセル）座標（例えば、ピクセル位置、ピクセルの行および列など）および深度をワールド（3D）座標系などの座標系にマッピングする単射関数を指定または定義するために使用され得る。ビューパラメータは、隣接するビュー内の画像を合成し、基準ビューではオクルードされ得るが隣接するビューでは少なくとも部分的にディスオクルードされるようになり得る穴または領域を識別し、基準ビューの一部または全部について基準ビューごとにこれらの穴または領域についてのディスオクルージョンテクスチャデータおよびディスオクルージョン深度データを決定、推定、または予測するために使用され得る。

【 0 0 6 6 】

基準ビューおよび所与の時点についての単一ビューテクスチャ画像および単一ビュー深度画像ごとに、ビットマスクなどの画像マスクが基準ビューについて生成され得、図3Aに示されるように、所与の時点についてのディスオクルージョンアトラスにおいてディスオクルージョンテクスチャおよび深度データが提供されるべき空間領域を識別する。

【 0 0 6 7 】

図3Bは、受信されたまたは入力のマルチビュービデオ内のマルチビュー画像のシーケンスについて作成され得る連続するディスオクルージョンアトラスの例示的なシーケンスを示す。ディスオクルージョンアトラスのシーケンスは、ディスオクルージョンアトラスのグループに符号化され得る。ディスオクルージョンアトラスのそのような各グループは、時間的に安定したディスオクルージョンアトラスを含み、比較的効率的にビデオストリームに符号化され得る。

【 0 0 6 8 】

図4Aは、時間間隔をカバーするマルチビュー画像のシーケンス内のマルチビュー画像について、図3Aに示されるようなディスオクルージョンアトラスを生成するための例示的な処理フローを示す。いくつかの例示的な実施形態では、1つまたは複数のコンピューティングデバイスまたは構成要素がこのプロセスフローを実行し得る。

【 0 0 6 9 】

マルチビュー画像は、時間間隔内の時点に対応するか、または時間間隔内の時点にインデックス付けされ、N個の基準ビューに対するN個の（ソース／入力）単一ビューテクスチャ画像と、N個の基準ビューに対するN個の（ソース／入力）単一ビュー深度画像とを含む。N個の単一ビューテクスチャ画像内の各単一ビューテクスチャ画像は、N個の単一ビュー深度画像内のそれぞれの単一ビュー深度画像に対応する。

【 0 0 7 0 】

ブロック402において、本明細書で説明されるシステム（例えば、図1Aの100など）は、N個の基準ビューに隣接するビュー内の合成／変形画像内に存在し得る空間領域または穴についての画像フラグメントを記憶する（例えば、コピー、スタンプ、配置などを行う）ためにディスオクルージョンアトラスが使用される前に、ディスオクルージョンアトラスに対して初期化動作を実行する。

【 0 0 7 1 】

ブロック402の初期化動作は、以下を含み得る：（a）N個の基準ビューに隣接するビューで、合成／変形画像においてテクスチャまたは深度データが欠けている可能性のあ

10

20

30

40

50

るN個の基準ビュー内の空間領域または穴を識別するN個の画像マスクを受信またはロードすること；(b)N個の画像マスクにおいて識別された画像フラグメントのためのテクスチャおよび深度情報を受信またはロードすること；(c)サイズによって画像フラグメントをソートして画像フラグメントのリストにすること；など。

【0072】

ここで、「サイズ」は、画像フラグメントの空間寸法を測定するためのメトリックを指す。画像フラグメントの空間寸法を測定するために様々なメトリックが使用され得る。例えば、画像フラグメントを完全に囲む最小の矩形が決定され得る。水平サイズ(「xsize」と表される)、垂直サイズ(「ysize」と表される)、水平サイズと垂直サイズとの組合せなどが、画像フラグメントのサイズを測定するためのメトリック(複数可)として個々にまたは集合的に使用され得る。

10

【0073】

いくつかの動作シナリオでは、画像フラグメントのサイズは、 $64 * \max(xsize, ysize) + \min(xsize, ysize)$ として計算され得、ここで、xsizeおよびysizeの各々は、ピクセルの単位で、または 2×2 ピクセルブロックの2ピクセル、 4×4 ピクセルブロックの4ピクセルなど、特定のサイズのピクセルブロックの水平または垂直寸法(これは2の非負の整数乗であり得る)の単位で表され得る。

【0074】

N個のロードされた画像マスク内の各画像マスクは、N個の基準ビュー内のそれぞれの基準ビューに対応する。画像マスクは、基準ビューではオクルードされるが、基準ビューに隣接するビューでは少なくとも部分的に見えるようになる画像フラグメントのための画像マスク部分を含む。画像マスクの画像マスク部分内の各画像マスク部分は、画像マスクが対応する基準ビューではオクルードされるが、基準ビューに隣接するビューでは少なくとも部分的に見えるようになる画像フラグメント内のそれぞれの画像フラグメントを空間的に画定または定義する。画像マスクで表される各ピクセルについて、ピクセルが画像フラグメントのうちの1つに属する場合、ビットインジケータは真または1に設定され、ピクセルが画像フラグメントのいずれにも属さない場合、偽または0に設定される。

20

【0075】

いくつかの動作シナリオでは、ディスオクルージョンアトラスは、マルチビュー画像の(例えば、すべてなどの)画像フラグメントの空間配置を示し、ディスオクルージョンアトラスにおいてディスオクルージョンデータが記憶または維持される画像フラグメントを識別または追跡するために使用されるレイアウトマスクを含む。レイアウトマスクは、矩形形状などの空間形状内に配置されたピクセルのアレイを含み得る。ディスオクルージョンアトラスのレイアウトマスクにおいて空間的に画定または定義された画像フラグメントは、互いに排他的であり、レイアウトマスクにおいて(例えば、完全には、など)互いにオーバーラップしない。

30

【0076】

ブロック402の初期化動作はさらに以下を含み得る：(d)単一の四分木ルートノードを作成すること。このルートノードは、最大の画像フラグメントのサイズをちょうどカバーする最良のサイズに初期化される。四分木は、対応するレイアウトマスクを可能な限り小さく保つために、必要に応じて各次元において2倍ずつ漸増的に成長させられる；(e)画像フラグメント(例えば、画像フラグメントの画像マスク部分など)を、ディスオクルージョンアトラスのレイアウトマスクにおける、第1のノードのための指定された領域にスタンプすることによって、最大の画像フラグメントを四分木の第1のノードにリンクすること；その他。四分木の第1のノードは、ここでは、レイアウトマスク全体を表すルートノードの下第1のレベルの四分木ノードの中の第1の四分木ノードを指す。ここで、「スタンプ」とは、画像フラグメントまたはその画像マスク部分をディスオクルージョンアトラスのレイアウトマスクにコピー、転写、または嵌め込むことを指す。ここで、「四分木」とは、各内部ノードが4つの子四分木ノードを有するツリーデータ構造を指す。

40

【0077】

50

四分木は、最初、同じサイズの長方形など、同じサイズの空間形状の4つのノードを含む。本明細書で説明される四分木のノードの空間形状は、2の非負の整数乗であるピクセルのカウントを有する特別な次元を有し得る。

【0078】

ディスオクルージョンアトラスのレイアウトマスクに最大の画像フラグメントをスタンプするのに続いて、最大の画像フラグメントが（サイズでソートされた）画像フラグメントのリストから除去され、最初の四分木ノードの後の次の四分木ノードが現在の四分木ノードとして設定される。現在の四分木ノードは、画像フラグメントをホストするために次に使用されるべき空の四分木ノードまたは候補四分木ノード（画像フラグメントまたはそれぞれの画像マスク部分によってまだポピュレートされていない）を表す。

10

【0079】

ブロック404において、システムは、サイズでソートされた画像フラグメントのリストが、ディスオクルージョンアトラスのレイアウトマスクにスタンプまたは空間的に配置される必要が依然としてある任意の画像フラグメントを含むかどうかを決定する。いくつかの実施形態では、最小フラグメントサイズしきい値を下回る任意の画像フラグメントは、リストから除去されてもよいし、リスト内で無視されてもよい。例示的な最小フラグメントサイズしきい値は、水平寸法および垂直寸法的一方または両方に4ピクセル、水平寸法および垂直寸法的一方または両方に6ピクセルなどのうちの1つであり得る。

【0080】

（サイズでソートされた）画像フラグメントのリストが、ディスオクルージョンアトラスのレイアウトマスクにスタンプまたは空間的に配置される必要が依然としてある画像フラグメント（複数可）を含まないと決定したことに応答して、処理フローは終了する。

20

【0081】

そうでなければ、（サイズでソートされた）画像フラグメントのリストが、ディスオクルージョンアトラスのレイアウトマスクにスタンプまたは空間的に配置される必要が依然としてある画像フラグメント（複数可）を含むと決定したことに応答して、システムは、（サイズでソートされた）画像フラグメントのリストから、次に大きい画像フラグメントを現在の画像フラグメントとして選択する。

【0082】

ブロック406において、システムは、四分木内の現在の四分木ノードが、現在の画像フラグメント、または現在の画像フラグメントのための対応する画像マスク部分をホストするのに十分に大きいかどうかを決定する。

30

【0083】

四分木内の現在の四分木ノードが、現在の画像フラグメントをホストするのに十分に大きくないと決定したことに応答して、処理フローはブロック410に進む。

【0084】

そうでなければ、四分木内の現在の四分木ノードが現在の画像フラグメントをホストするのに十分に大きいと決定したことに応答して、処理フローはブロック408に進む。

【0085】

ブロック408において、システムは、現在の四分木ノードが現在の画像フラグメントのための「最良」適合四分木ノードであるかどうかを決定する。「最良」適合四分木ノードは、画像フラグメントまたはその画像マスク部分をホストするのにちょうど十分な大きさである四分木ノードを指す。言い換えると、最良「適合」四分木ノードは、ディスオクルージョンアトラスのレイアウトマスク内の画像フラグメントを完全に囲むか、またはホストするための最小サイズの四分木ノードを表す。

40

【0086】

現在の四分木ノードが現在の画像フラグメントのための「最良」適合四分木ノードではないと決定したことに応答して、システムは、「最良」適合四分木ノードが見つかるまで、現在の四分木ノードを（例えば、繰り返し、反復的に、再帰的に、などで）細分する。「最良」適合四分木ノードは、現在の四分木ノードになるように設定される。

50

【 0 0 8 7 】

現在の四分木ノードが現在の画像フラグメントのための「最良」適合四分木ノードであると決定されると、システムは、「最良」適合四分木ノードにおいて現在の画像フラグメントをスタンプまたは空間的に画定する。

【 0 0 8 8 】

ディスオクルージョンアトラスまたは現在の四分木ノードのレイアウトマスクに現在の画像フラグメントをスタンプするのに続いて、現在の画像フラグメントが（サイズでソートされた）画像フラグメントのリストから除去され、（除去された）現在の四分木ノードの後の次の四分木ノードが、（新しいまたは現在の（present））現在の四分木ノードとして設定される。

10

【 0 0 8 9 】

ブロック 4 1 0 において、システムは、ディスオクルージョンアトラスのレイアウトマスク全体を表すルートノードの下のどこかに、現在の画像フラグメントをホストするために利用可能な空の四分木ノードまたは候補四分木ノードがあるかどうかを決定する。ある場合、現在の画像フラグメントをホストするために、空の四分木ノードまたは候補四分木ノードが（例えば、2 つ以上のノードが使用される場合は集合的に、など）使用される。その後、プロセスフローはブロック 4 0 4 に進む。したがって、現在の画像フラグメントが現在の四分木ノードの下の子の任意の既存の（子）四分木ノードに（その全体が）適合しない場合、レイアウトマスク内の任意の場所にフラグメントを適合させる試みを行うことができる。多くの動作シナリオでは、四分木（複数可）は、アトラス構築をより高速にするように設計された加速データ構造にすぎないことに留意されたい。本明細書で説明される四分木は、レイアウト（またはレイアウトマスク）が決定されると、保存または必要とされない可能性がある。さらに、任意の画像フラグメントが配置され得る場所に対して、本明細書で説明される四分木によって課される（例えば、絶対的な、固有の、などの）制限は存在しない。画像フラグメントは、複数の四分木ノードに重複することができ、いくつかの動作シナリオではオーバーラップすることが多い。したがって、「最良適合」方法（例えば、現在の画像フラグメントなどのフラグメントのための単一の最良適合ノードを見つけるためなどの）が失敗した場合、レイアウトマスク全体にわたって、より網羅的（および高価な）探索を実行して、手近のフラグメントに適合させることができる。成功すると、このように配置されたフラグメントがオーバーラップするすべての四分木ノードが「占有」としてマークされ、処理が継続する。失敗すると、処理フローはブロック 4 1 2 に進み、四分木を成長させる。図 4 A に示されるような全体的なアルゴリズムが依然として効率的かつ効果的である理由は、多くの動作シナリオにおける大抵の場合、最良適合四分木探索が成功するからである。現在の画像フラグメントのための最良適合ノードが見つからない場合にのみ、より高価なまたは網羅的なフォールバック探索を行うか、または呼び出して、現在の画像フラグメントをホストするためにオーバーラップする可能性のある四分木ノードを見つける。これは、ディスオクルージョンアトラスのレイアウトマスク全体におけるすべての空の四分木ノードまたは候補四分木ノード（任意の画像フラグメントによってまだ占有されていない）まで（例えば、探索ループなどで）探索することを伴い得る。

20

30

40

【 0 0 9 0 】

レイアウトマスク内に残っている空の四分木ノードまたは候補四分木ノードのいずれも、現在の画像フラグメントをホストするのに十分な大きさではないと決定したことに応答して、処理フローはブロック 4 1 2 に進む。

【 0 0 9 1 】

そうでなければ、レイアウトマスク内の空の四分木ノードまたは候補四分木ノードが現在の画像フラグメントをホストするのに十分に大きいと決定したことに応答して、空の四分木ノードまたは候補四分木ノードが（新しい）現在の四分木ノードとして設定され、処理フローはブロック 4 0 8 に進む。

【 0 0 9 2 】

50

ブロック 4 1 2 において、システムは、ディスオクルージョンアトラスまたはディスオクルージョンアトラスのレイアウトマスクのサイズを、水平寸法および垂直寸法のそれぞれで 2 倍 (2 x) に拡張または増加させる。この拡張の前の既存の四分木 (または古い四分木) は、第 1 の四分木ノード (例えば、新たに拡張された四分木の左上象限など) にリンクまたは配置され得る。第 2 の四分木ノード (例えば、新たに拡張された四分木の右上象限など) は、(新たな) 現在の四分木ノードになるように設定される。処理フローはブロック 4 0 8 に進む。

【 0 0 9 3 】

画像フラグメントに属するものとしてディスオクルージョンアトラスのレイアウトマスクにおいて識別された各ピクセルのテクスチャ値および深度値は、ディスオクルージョンアトラスのレイアウトマスクとともにディスオクルージョンアトラスの一部として記憶、キャッシュ、またはバッファされ得る。

10

【 0 0 9 4 】

7 . 時間的に安定したグループレベルのレイアウトマスク

ビデオシーケンス内の連続するディスオクルージョンアトラスを安定化させるために、複数の連続する時点についての連続するディスオクルージョンアトラスのグループ (これは、ビデオシーケンス内のテクスチャ画像 GOP、深度画像 GOP などに対応し得る) 内のディスオクルージョンアトラスのレイアウトマスクは、連続するディスオクルージョンアトラスのグループのためのグループレベルのレイアウトマスクを形成するために、「or」演算によって分離的に結合され得る。

20

【 0 0 9 5 】

レイアウトマスク内の各レイアウトマスクは、等しいサイズであり得、レイアウトマスク内の任意のピクセルが、それぞれのディスオクルージョンアトラスにおいてホストされる画像フラグメントに属するかどうかを示すために、それぞれのインジケータまたはビットを用いて同じピクセルアレイを圧縮する。

【 0 0 9 6 】

グループレベルのレイアウトマスクは、連続するディスオクルージョンアトラスのグループのための (個々の) レイアウトマスクと同じサイズであってもよく、(個々の) レイアウトマスクの場合と同じピクセルのアレイを含む。和集合演算または分離的な「OR」演算を通してグループレベルのレイアウトマスクを生成するために、特定のピクセルロケーションまたはインデックスにおけるピクセルのためのインジケータまたはビットは、連続するディスオクルージョンアトラスのグループのための (個々の) レイアウトマスク内の同じ特定のピクセルロケーションまたはインデックスにおける対応するピクセルのためのインジケータまたはビットのいずれかが真または 1 である場合、真または 1 に設定され得る。

30

【 0 0 9 7 】

グループレベルのレイアウトマスクまたはそのインスタンスは、連続するディスオクルージョンアトラスのグループにおいてカバーされる複数の連続する時点のそれぞれの時点についてのディスオクルージョンアトラスで表されるべき画像フラグメントをホストまたはレイアウトするために、連続するディスオクルージョンアトラスのグループ内の各ディスオクルージョンアトラスに対して繰り返し使用され得る。ある時点についてのディスオクルージョンテクスチャおよび深度情報を有さないピクセルは、その時点 (またはタイムスタンプ) についてのグループレベルのレイアウトマスクの対応するインスタンスにおいて除外され得る (例えば、未定義、非占有など)。図 3 C は、本明細書で説明される共通のグループレベルのレイアウトマスクを用いて生成された連続するディスオクルージョンアトラスの例示的なグループを示す。

40

【 0 0 9 8 】

いくつかの動作シナリオでは、同じグループレベルのレイアウトマスクの複数のインスタンスを使用して、別個のディスオクルージョンアトラスが生成され得る。連続するディスオクルージョンアトラスのグループ内の最初のディスオクルージョンアトラスを、組み

50

合わされたグループレベルのレイアウトマスクの最初のインスタンスとともに使用して、開始アトラスフレームが生成され得、その後、連続するディスオクルージョンアトラスのグループ内の他のディスオクルージョンアトラスから生成された他のアトラスフレームが続く。開始アトラスフレームおよび他のアトラスフレームは、開始アトラスフレームと、グループの終了前の次の開始アトラスフレームとによって区切られた連続するアトラスフレームのグループを形成し得る。時間的に安定しているグループレベルのレイアウトマスクを使用して、データ類似性を見つけるためにインター予測および/またはイントラ予測を適用するなど、データ圧縮動作を容易にし、受信者ビデオデコーダに送信されるべき連続するディスオクルージョンアトラスのグループ内の全体的なデータを低減することができる。いくつかの実装例では、(フレーム)時間間隔にわたってレイアウトマスク(またはビットマスク)の結合(union)を使用することで、ビデオ圧縮を2倍以上改善することができる。

10

【0099】

いくつかの動作シナリオでは、連続するディスオクルージョンアトラスのための個々のレイアウトマスクの各々において識別されたすべてのピクセルのためのテクスチャおよび深度データは、連続するディスオクルージョンアトラスのための組み合わされたグループレベルのレイアウトマスクを生成することなく、含まれるかまたは送信され得る。互いに時間的および空間的に異なる個々のレイアウトマスクを使用するデータ圧縮動作は、本明細書で説明されるグループレベルのレイアウトマスクを使用したデータ圧縮動作ほどデータ量を低減する際に効率的でないことがある。

20

【0100】

いくつかの動作シナリオでは、ディスオクルージョンアトラスのレイアウトマスク上に画像フラグメントをレイアウトするために、これらの画像フラグメントは、最初に回転されることなく、空の四分木ノードまたは候補四分木ノードなどの利用可能な空間領域に適合され得る。いくつかの動作シナリオでは、パッキング効率性を高めるために、画像フラグメントは、最初に回転させられてから、「最良」適合四分木ノードに配置されてもよい。その結果、回転前には画像フラグメントをホストすることができなかった可能性がある四分木ノードが、回転後には画像フラグメントをホストできるようになり得る。

【0101】

マルチビュー画像またはその中の任意の単一ビュー画像は360度画像であり得る。360度画像の画像データ(ディスオクルージョンデータを含む)は、矩形フレーム(例えば、「ベースボールカバー」ビューなど)などの画像フレームで表され得る。図2Aおよび図2Bに示されるように、そのような画像は、例えば「野球のカバー」ビューにおいて、矩形の画像フレームへと組み合わされる複数の画像セグメントを含む。しかしながら、複数の画像セグメントは、異なるビューの形状、例えば正方形のビューの形状に組み合わされてもよい。

30

【0102】

本明細書で説明されるディスオクルージョンアトラスは、画像フラグメントが、画像セグメントの境界に接するテクスチャおよび深度情報を含むことを示すために、レイアウトマスク内にマスクストライピングを含むか、または示し得る。マスクストライピングをレイアウトマスク内に配置する理由は、アトラスホストされた画像フラグメントが、1つまたは複数の継ぎ目で接された複数の画像セグメントを含み得る360度画像における継ぎ目に対応する C^0 (または0次)不連続部(discontinuity)を横切る場合を回避するためである。例えば、360度画像のベースボールカバー表現では、中央に1つの長い水平継ぎ目があり、そこでは、継ぎ目の異なる側の隣接ピクセルは、視覚シーンの(例えば、実際の)ビューの隣接部分に対応しない。マスクストライピングは、画像フラグメントがこの境界を横切らないことを保証するために、この継ぎ目に沿って入力マスク内の線をゼロ化することにより、ディスオクルージョンアトラス(複数可)において実装され得る。したがって、マスクストライピングを有する画像フラグメントは、画像フラグメントと同じ線の側の穴またはギャップを充填するために適用されるように制約され、正しく解釈

40

50

され得る。

【0103】

8. 例示的なプロセスフロー

図4Bは、本発明の例示的な実施形態による例示的なプロセスフローを示す。いくつかの例示的な実施形態では、1つまたは複数のコンピューティングデバイスまたは構成要素が、このプロセスフローを実行し得る。ブロック422において、上流デバイスは、1つまたは複数の基準ビューからの視覚シーンを描写する1つまたは複数の基準画像ではオクルードされ、1つまたは複数の基準ビューに隣接する非基準ビューでは少なくとも部分的にディスオクルードされるようになる画像フラグメントをサイズでソートする。画像フラグメントは、画像フラグメント内の任意の他の画像フラグメント以上のサイズである第1の画像フラグメントを含む。

10

【0104】

ブロック424において、上流デバイスは、画像フラグメントを記憶するために使用されるディスオクルージョンアトラスのためのレイアウトマスクを生成する。レイアウトマスクは、第1の画像フラグメントに対して特にサイズ決定された第1の最良適合ノードを含む四分木でカバーされる。第1の適合ノードは、第1の画像セグメントを（例えば、完全に）カバーするようにサイズ決定される。

【0105】

ブロック426において、上流デバイスは、レイアウトマスク内で識別された最良適合ノードにソートされた画像フラグメントを降順で記憶する。ソートされた画像フラグメント内の各画像フラグメントは、最良適合ノード内のそれぞれの最良適合ノードに記憶される。最良適合ノードは、レイアウトマスクをカバーする四分木内の少なくとも1つのノードを反復的に分割することによって取得される少なくとも1つの最良適合ノードを含む。最良適合ノードの各々は、それぞれの画像フラグメントの各々を完全にカバーするための最小サイズの四分木ノードとして識別され得る。

20

【0106】

ブロック428において、上流デバイスは、1つまたは複数の基準画像で符号化されたボリュメトリックビデオ信号を生成する。ボリュメトリックビデオ信号は、ディスオクルージョンアトラス内の画像フラグメントを用いてさらに符号化される。1つまたは複数の基準画像は、画像ディスプレイ上にレンダリングするための非表現ビューにおける表示画像を合成するためにボリュメトリックビデオ信号の受信者デバイスによって使用される。ディスオクルージョンアトラス内の画像フラグメントは、表示画像におけるディスオクルードされた空間領域内のディスオクルードされた画像データを充填するために受信者デバイスによって使用される。

30

【0107】

一実施形態では、1つまたは複数の基準画像の各々は、360度画像、180度画像、ビューポート画像、規則的な空間形状画像フレーム内の画像、または不規則な空間形状画像フレーム内の画像のうちの1つを表す。

【0108】

一実施形態では、画像フラグメントの各々は、1つまたは複数の基準ビューにおいてオクルードされた連続するピクセルによって形成される空間領域について、テクスチャ画像値および深度画像値を含む。

40

【0109】

一実施形態では、1つまたは複数の顕著性ビデオストリームのセットは、第1の顕著性ランクが割り当てられた第1の顕著性ビデオストリームと、第1の顕著性ランクよりも低い第2の顕著性ランクが割り当てられた第2の顕著性ビデオストリームとを含み、第2のビデオストリームは、利用可能なデータレートが低減されたと決定したことに応答して、後の時間にビデオストリーミングクライアントに送信されるべき1つまたは複数の顕著性ビデオストリームのセットから除去される。

【0110】

50

一実施形態では、1つまたは複数の基準画像は、複数の連続する時点についての複数の連続するマルチビュー画像を含むマルチビュー画像グループ内のマルチビュー画像に含まれ、ディスオクルージョンアトラスは、複数の連続する時点についての複数のディスオクルージョンアトラスを含むディスオクルージョンアトラスグループに含まれる。

【0111】

一実施形態では、レイアウトマスクは、複数のディスオクルージョンアトラスのために生成された複数の個別のレイアウトマスクに含まれ、グループレベルのレイアウトマスクは、和集合演算により複数の個別のレイアウトマスクから生成され、ボリュメトリックビデオ信号に符号化されたディスオクルージョンアトラスは、グループレベルのレイアウトマスクで表される。

10

【0112】

一実施形態では、ディスオクルージョンアトラスグループは、アトラスフレームのグループとしてボリュメトリックビデオ信号に符号化され、アトラスフレームのグループは、アトラスIフレームから開始し、異なるアトラスIフレームの前に終了する。

【0113】

一実施形態では、ディスオクルージョンアトラスは、マスクストライピングを含み、マスクストライピングは、ディスオクルージョンアトラスに記憶された画像フラグメントが画像セグメントの1つまたは複数の境界で接していることを示す。

【0114】

一実施形態では、レイアウトマスクは、拡張前のサイズのレイアウトマスク内に最良適合ノードが見つからないと決定したことに応答して拡張される。

20

【0115】

一実施形態では、画像フラグメントについての空間領域は、ビットマスクで識別され、画像フラグメントは、画像フラグメントについての空間領域のサイズを使用してソートされる。

【0116】

一実施形態では、ディスオクルージョンアトラスに記憶された画像フラグメントは、視覚シーンから識別される1つまたは複数の顕著性領域に位置する。顕著性領域は、視覚シーンのより興味深いまたはより重要な部分または関心領域であり得る。

【0117】

図4Cは、本発明の例示的な実施形態による例示的なプロセスフローを示す。いくつかの例示的な実施形態では、1つまたは複数のコンピューティングデバイスまたは構成要素が、このプロセスフローを実行し得る。

30

【0118】

ブロック460において、下流デコーダ（例えば、受信者デバイスまたはデコーダ）が、ボリュメトリックビデオ信号を受信する。ボリュメトリックビデオ信号は、例えば図4Bを参照して上述した実施形態のいずれかで符号化／生成され得る。ボリュメトリックビデオ信号は、1つまたは複数の基準画像およびディスオクルージョンアトラス内の画像フラグメントで符号化される。ディスオクルージョンアトラスは、画像フラグメントを記憶するために使用される。1つまたは複数の基準ビューからの視覚シーンを描写する1つまたは複数の基準画像ではオクルードされ、1つまたは複数の基準ビューに隣接する非基準ビューでは少なくとも部分的にディスオクルードされるようになる画像フラグメントは、図4Bを参照して説明されたように、サイズでソートされる（ブロック422）。

40

【0119】

ブロック462において、下流デバイスは、ボリュメトリックビデオ信号から、1つまたは複数の基準画像を復号する。

【0120】

ブロック464において、下流デバイスは、ボリュメトリックビデオ信号から、ディスオクルージョンアトラス内の画像フラグメントを復号する。

【0121】

50

ブロック 4 6 6 において、下流デバイスは、1 つまたは複数の基準画像に基づいて、非表現ビューにおける表示画像を合成する。

【 0 1 2 2 】

ブロック 4 6 8 において、下流デバイスは、ディスオクルージョンアトラス内の画像フラグメントを使用して、表示画像におけるディスオクルードされた空間領域内のディスオクルードされた画像データを充填する。

【 0 1 2 3 】

ブロック 4 7 0 において、上流デバイスは、画像ディスプレイ上に表示画像をレンダリングする。

【 0 1 2 4 】

一実施形態では、画像フラグメントの各々は、1 つまたは複数の基準ビューにおいてオクルードされた連続するピクセルによって形成される空間領域について、テクスチャ画像値および深度画像値を含む。

【 0 1 2 5 】

一実施形態では、ブロック 4 6 6 において、表示画像を合成することは、1 つまたは複数の基準ビューに利用可能なテクスチャ画像値および深度画像値を使用することを含む。

【 0 1 2 6 】

一実施形態では、合成された表示画像内のディスオクルードされた空間領域は、1 つまたは複数の基準ビューに利用可能なテクスチャ画像値および深度画像値が、1 つまたは複数の基準ビューに隣接する非基準ビューについて取得可能でないと決定することによって識別される。

【 0 1 2 7 】

一実施形態では、ディスオクルージョンアトラスに記憶された画像フラグメントは、視覚シーンから識別される 1 つまたは複数の顕著性領域に位置し、ディスオクルージョンアトラスは、合成された表示画像において 1 つまたは複数の顕著性領域が識別されるように、1 つまたは複数の顕著性領域から離れた空間領域をカバーするためのテクスチャ画像値または深度画像値を含まない。

【 0 1 2 8 】

一実施形態では、ポリュメトリックビデオ信号は、単射関数を指定する画像メタデータを含み、単射関数は、画像フラグメント内の各ピクセルを、画像フレーム内のピクセルのピクセルロケーションから、視覚シーンが表される 3 次元座標系内の対応するロケーションにマッピングする。

【 0 1 2 9 】

様々な例示的な実施形態では、装置、システム、装置、または 1 つもしくは複数の他のコンピューティングデバイスが、説明された前述の方法のいずれかまたは一部を実行する。一実施形態では、非一時的コンピュータ可読記憶媒体は、1 つまたは複数のプロセッサによって実行されると、本明細書で説明される方法を実行させる、ソフトウェア命令を記憶する。

【 0 1 3 0 】

別個の実施形態が本明細書で述べられているが、本明細書で述べられている実施形態および/または部分的な実施形態の任意の組合せが、さらなる実施形態を形成するために組み合わせられることに留意されたい。

【 0 1 3 1 】

9 . 実装機構 - ハードウェアの概要

一実施形態によれば、本明細書で説明される技法は、1 つまたは複数の専用コンピューティングデバイスによって実装される。専用コンピューティングデバイスは、本技法を実行するためにハードワイヤードされ得るか、または本技法を実行するように永続的にプログラムされる 1 つもしくは複数の特定用途向け集積回路 (ASIC) もしくはフィールドプログラマブルゲートアレイ (FPGA) などのデジタル電子デバイスを含み得るか、またはファームウェア、メモリ、他の記憶装置、もしくは組合せにおけるプログラム命令に

10

20

30

40

50

したがって本技法を実行するようにプログラムされた１つもしくは複数の汎用ハードウェアプロセッサを含み得る。そのような専用コンピューティングデバイスはまた、本技法を達成するために、カスタムハードワイヤードロジック、ＡＳＩＣ、またはＦＰＧＡをカスタムプログラミングと組み合わせ得る。専用コンピューティングデバイスは、デスクトップコンピュータシステム、ポータブルコンピュータシステム、ハンドヘルドデバイス、ネットワークデバイス、または本技法を実装するためにハードワイヤードおよび／またはプログラムロジックを組み込む任意の他のデバイスであり得る。

【０１３２】

例えば、図５は、本発明の例示的な実施形態が実装され得るコンピュータシステム５００を示すブロック図である。コンピュータシステム５００は、情報を通信するためのバス５０２または他の通信機構と、情報を処理するために、バス５０２に結合されたハードウェアプロセッサ５０４とを含む。ハードウェアプロセッサ５０４は、例えば、汎用マイクロプロセッサであり得る。

10

【０１３３】

コンピュータシステム５００はまた、情報およびプロセッサ５０４によって実行されるべき命令を記憶するためにバス５０２に結合された、ランダムアクセスメモリ（ＲＡＭ）または他の動的記憶デバイスなどのメインメモリ５０６を含む。メインメモリ５０６はまた、プロセッサ５０４によって実行されるべき命令の実行中、テンポラリ変数または他の中間情報を記憶するために使用され得る。そのような命令は、プロセッサ５０４にアクセス可能な非一時的記憶媒体に記憶されると、コンピュータシステム５００を、命令で指定された動作を実行するようにカスタマイズされた専用マシンにする。

20

【０１３４】

コンピュータシステム５００は、プロセッサ５０４のための静的情報および命令を記憶するために、バス５０２に結合された読取り専用メモリ（ＲＯＭ）５０８または他の静的記憶デバイスをさらに含む。

【０１３５】

磁気ディスクまたは光ディスク、ソリッドステートＲＡＭなどの記憶デバイス５１０が提供され、情報および命令を記憶するためにバス５０２に結合される。

【０１３６】

コンピュータシステム５００は、情報をコンピュータユーザに表示するために、バス５０２を介して、液晶ディスプレイなどのディスプレイ５１２に結合され得る。英数字および他のキーを含む入力デバイス５１４は、情報およびコマンド選択をプロセッサ５０４に通信するためにバス５０２に結合される。別のタイプのユーザ入力デバイスは、方向情報およびコマンド選択をプロセッサ５０４に通信し、ディスプレイ５１２上でのカーソル移動を制御するためのマウス、トラックボール、またはカーソル方向キーなどのカーソル制御５１６である。この入力デバイスは、典型的には、第１の軸（例えば、ｘ）および第２の軸（例えば、ｙ）という２つの軸において２つの自由度を有し、これにより、デバイスが平面内の位置を指定することができる。

30

【０１３７】

コンピュータシステム５００は、カスタマイズされたハードワイヤードロジック、１つもしくは複数のＡＳＩＣまたはＦＰＧＡ、ファームウェア、および／またはコンピュータシステムと組み合わせてコンピュータシステム５００を専用マシンにするかまたは専用マシンになるようにプログラムするプログラムロジックを使用して、本明細書で説明される技法を実装し得る。一実施形態によれば、本明細書の技法は、プロセッサ５０４がメインメモリ５０６に含まれる１つまたは複数の命令の１つまたは複数のシーケンスを実行することに応答して、コンピュータシステム５００によって実行される。そのような命令は、記憶デバイス５１０などの別の記憶媒体からメインメモリ５０６に読み込まれ得る。メインメモリ５０６に含まれる命令のシーケンスの実行は、プロセッサ５０４に、本明細書で説明されるプロセスステップを実行させる。代替の実施形態では、ハードワイヤード回路が、ソフトウェア命令の代わりに、またはそれと組み合わせて使用されてもよい。

40

50

【 0 1 3 8 】

本明細書で使用される場合、「記憶媒体 (storage media)」という用語は、機械を特定の方式で動作させるデータおよび/または命令を記憶する任意の非一時的媒体を指す。そのような記憶媒体は、不揮発性媒体および/または揮発性媒体を含み得る。不揮発性媒体は、例えば、記憶デバイス 510 のような光ディスクまたは磁気ディスクを含む。揮発性媒体は、メインメモリ 506 などの動的メモリを含む。記憶媒体の一般的な形態は、例えば、フロッピー (登録商標) ディスク、フレキシブルディスク、ハードディスク、ソリッドステートドライブ、磁気テープもしくは任意の他の磁気データ記憶媒体、CD-ROM、任意の他の光学データ記憶媒体、穴のパターンを有する任意の物理媒体、RAM、PROM、および EPROM、FLASH (登録商標) - EPROM、NVRAM、任意の他のメモリチップまたはカートリッジを含む。

10

【 0 1 3 9 】

記憶媒体は、伝送媒体とは別個であるが、伝送媒体とともに使用され得る。伝送媒体は、記憶媒体間の情報の転送に関与する。例えば、伝送媒体は、バス 502 を構成するワイヤを含む、同軸ケーブル、銅線、および光ファイバを含む。伝送媒体はまた、電波および赤外線データ通信中に生成されるものなどの音波または光波の形態をとることもできる。

【 0 1 4 0 】

様々な形態の媒体が、実行のためにプロセッサ 504 に 1 つまたは複数の命令の 1 つまたは複数のシーケンスを搬送することに関与し得る。例えば、命令は、最初に、リモートコンピュータの磁気ディスクまたはソリッドステートドライブ上で搬送され得る。リモートコンピュータは、命令をその動的メモリにロードし、モデムを使用して電話回線上で命令を送信することができる。コンピュータシステム 500 にローカルなモデムは、電話回線上でデータを受信し、赤外線送信機を使用してデータを赤外線信号に変換することができる。赤外線検出器は、赤外線信号で搬送されるデータを受信することができ、適切な回路が、データをバス 502 上に配置することができる。バス 502 はデータをメインメモリ 506 に搬送し、プロセッサ 504 は、そこから命令を取り出して実行する。メインメモリ 506 によって受信された命令は、プロセッサ 504 による実行の前または後のいずれかに記憶デバイス 510 に記憶され得る。

20

【 0 1 4 1 】

コンピュータシステム 500 はまた、バス 502 に結合された通信インターフェース 518 を含む。通信インターフェース 518 は、ローカルネットワーク 522 に接続されたネットワークリンク 520 への双方向データ通信結合を提供する。例えば、通信インターフェース 518 は、統合サービスデジタル網 (ISDN) カード、ケーブルモデム、衛星モデム、または対応するタイプの電話回線へのデータ通信接続を提供するモデムであり得る。別の例として、通信インターフェース 518 は、互換性のあるローカルエリアネットワーク (LAN) へのデータ通信接続を提供する LAN カードであってもよい。ワイヤレスリンクを実装することもできる。任意のそのような実装形態では、通信インターフェース 518 は、様々なタイプの情報を表すデジタルデータストリームを搬送する電気信号、電磁信号、または光信号を送受信する。

30

【 0 1 4 2 】

ネットワークリンク 520 は、典型的には、1 つまたは複数のネットワークを通して他のデータデバイスにデータ通信を提供する。例えば、ネットワークリンク 520 は、ローカルネットワーク 522 を通して、ホストコンピュータ 524 へのまたはインターネットサービスプロバイダ (ISP) 526 によって動作されるデータ機器への接続を提供し得る。次に、ISP 526 は、現在一般に「インターネット」528 と呼ばれる世界規模の packets データ通信ネットワークを通してデータ通信サービスを提供する。ローカルネットワーク 522 およびインターネット 528 は両方とも、デジタルデータストリームを搬送する電気信号、電磁信号、または光信号を使用する。コンピュータシステム 500 との間でデジタルデータを搬送する、様々なネットワークを通る信号、およびネットワークリンク 520 上にあり通信インターフェース 518 を通る信号は、伝送媒体の例示的な形態

40

50

である。

【 0 1 4 3 】

コンピュータシステム 5 0 0 は、ネットワーク（複数可）、ネットワークリンク 5 2 0、および通信インターフェース 5 1 8 を通して、メッセージを送信し、プログラムコードを含むデータを受信することができる。インターネットの例では、サーバ 5 3 0 は、インターネット 5 2 8、ISP 5 2 6、ローカルネットワーク 5 2 2、および通信インターフェース 5 1 8 を通して、アプリケーションプログラムのための要求されたコードを送信することができる。

【 0 1 4 4 】

受信されたコードは、それが受信されるとプロセッサ 5 0 4 によって実行され、および / または後の実行のために記憶デバイス 5 1 0 または他の不揮発性記憶装置に記憶され得る。

10

【 0 1 4 5 】

1 0 . 同等物、拡張物、代替物、およびその他

前述の明細書では、本発明の例示的な実施形態が、実装形態ごとに異なり得る多数の具体的な詳細を参照して説明されてきた。したがって、何か本発明であり、および本出願人によって本発明であると意図されているかの唯一かつ排他的なインジケータは、本出願から発行される請求項のセットであり、任意のその後の補正を含め、当該請求項が発行される特定の形態である。このような請求項に含まれる用語について本明細書に明示的に記載される任意の定義は、請求項で使用されるかかる用語の意味に適用されるものとする。したがって、請求項に明示的に記載されていない限定、要素、特性、特徴、利点または属性は、当該請求項の範囲を決して限定すべきではない。したがって、本明細書および図面は、限定的な意味ではなく例示的な意味で考慮されるべきである。

20

【 0 1 4 6 】

いくつかの実施形態の態様は、以下の列举された例示的な実施形態（ E E E ）を含む。

【 0 1 4 7 】

E E E 1 . 方法であって、

1 つまたは複数の基準ビューからの視覚シーンを描写する 1 つまたは複数の基準画像ではオクルードされ、1 つまたは複数の基準ビューに隣接する非基準ビューでは少なくとも部分的にディスオクルードされるようになる画像フラグメントをサイズでソートするステップであって、画像フラグメントは、画像フラグメント内の任意の他の画像フラグメント以上のサイズである第 1 の画像フラグメントを含む、ステップと、

30

画像フラグメントを記憶するために使用されるディスオクルージョンアトラスのためのレイアウトマスクを生成するステップであって、レイアウトマスクは、第 1 の画像フラグメントをカバーするように特にサイズ決定された第 1 の最良適合ノードを含む四分木でカバーされ、ディスオクルージョンアトラスは、複数のオーバーラップしていない画像フラグメントを含む最小総面積の結合画像である、ステップと、

レイアウトマスク内で識別された最良適合ノードにソートされた画像フラグメントを降順で記憶するステップであって、ソートされた画像フラグメント内の各画像フラグメントは、最良適合ノード内のそれぞれの最良適合ノードに記憶され、最良適合ノードは、レイアウトマスクをカバーする四分木内の少なくとも 1 つのノードを反復的に分割することによって取得される少なくとも 1 つの最良適合ノードを含む、ステップと、

40

1 つまたは複数の基準画像で符号化されたボリュメトリックビデオ信号を生成するステップであって、ボリュメトリックビデオ信号は、ディスオクルージョンアトラス内の画像フラグメントを用いてさらに符号化され、1 つまたは複数の基準画像は、画像ディスプレイ上にレンダリングするための非表現ビューにおける表示画像を合成するためにボリュメトリックビデオ信号の受信者デバイスによって使用され、ディスオクルージョンアトラス内の画像フラグメントは、表示画像におけるディスオクルードされた空間領域内のディスオクルードされた画像データを充填するために受信者デバイスによって使用される、ステップと

50

を含む方法。

【 0 1 4 8 】

E E E 2 . 1 つまたは複数の基準画像の各々は、360度画像、180度画像、ビューポート画像、規則的な空間形状画像フレーム内の画像、または不規則な空間形状画像フレーム内の画像のうちの1つを表す、E E E 1 に記載の方法。

【 0 1 4 9 】

E E E 3 . 画像フラグメントの各々は、1 つまたは複数の基準ビューにおいてオクルードされた連続するピクセルによって形成される空間領域について、テクスチャ画像値および深度画像値を含む、E E E 1 または E E E 2 に記載の方法。

【 0 1 5 0 】

E E E 4 . 1 つまたは複数の基準画像は、複数の連続する時点についての複数の連続するマルチビュー画像を含むマルチビュー画像グループ内のマルチビュー画像に含まれ、ディスオクルージョンアトラスは、複数の連続する時点についての複数のディスオクルージョンアトラスを含むディスオクルージョンアトラスグループに含まれる、E E E 1 から 3 のいずれかに記載の方法。

【 0 1 5 1 】

E E E 5 . レイアウトマスクは、複数のディスオクルージョンアトラスのために生成された複数の個別のレイアウトマスクに含まれ、グループレベルのレイアウトマスクは、和集合演算により複数の個別のレイアウトマスクから生成され、ポリュメトリックビデオ信号に符号化されたディスオクルージョンアトラスは、グループレベルのレイアウトマスクで表される、E E E 4 に記載の方法。

【 0 1 5 2 】

E E E 6 . ディスオクルージョンアトラスグループは、アトラスフレームのグループとしてポリュメトリックビデオ信号に符号化され、アトラスフレームのグループは、アトラスIフレームから開始し、異なるアトラスIフレームの前に終了する、E E E 4 または E E E 5 に記載の方法。

【 0 1 5 3 】

E E E 7 . ディスオクルージョンアトラスは、マスクストライピングを含み、マスクストライピングは、ディスオクルージョンアトラスに記憶された画像フラグメントが画像セグメントの1つまたは複数の境界で接していることを示す、E E E 1 から 6 のいずれかに記載の方法。

【 0 1 5 4 】

E E E 8 . レイアウトマスクは、拡張前のサイズのレイアウトマスク内に最良適合ノードが見つからないと決定したことに応答して拡張される、E E E 1 から 7 のいずれかに記載の方法。

【 0 1 5 5 】

E E E 9 . 画像フラグメントについての空間領域は、ビットマスクで識別され、画像フラグメントは、画像フラグメントについての空間領域のサイズを使用してソートされる、E E E 1 から 8 のいずれかに記載の方法。

【 0 1 5 6 】

E E E 1 0 . ディスオクルージョンアトラスに記憶された画像フラグメントは、視覚シーンから識別される1つまたは複数の顕著性領域に位置する、E E E 1 から 9 のいずれかに記載の方法。

【 0 1 5 7 】

E E E 1 1 . 1 つまたは複数の顕著性ビデオストリームは、第1の顕著性ランクが割り当てられた第1の顕著性ビデオストリームと、第1の顕著性ランクよりも低い第2の顕著性ランクが割り当てられた第2の顕著性ビデオストリームとを含む、E E E 1 0 に記載の方法。

【 0 1 5 8 】

E E E 1 2 . 方法であって、

10

20

30

40

50

ポリュメトリックビデオ信号から、１つまたは複数の基準画像を復号するステップと、
ポリュメトリックビデオ信号から、ディスオクルージョンアトラス内の画像フラグメントを復号するステップと、

１つまたは複数の基準画像から、非表現ビューにおける表示画像を合成するステップと、
ディスオクルージョンアトラス内の画像フラグメントを使用して、表示画像におけるディスオクルードされた空間領域内のディスオクルードされた画像データを充填するステップと、

画像ディスプレイ上に表示画像をレンダリングするステップと
を含む方法。

【０１５９】

10

ＥＥＥ１３．ポリュメトリックビデオ信号は、単射関数を指定する画像メタデータを含み、単射関数は、画像フラグメント内の各ピクセルを、画像フレーム内のピクセルのピクセルロケーションから、視覚シーンが表される３次元座標系内の対応するロケーションにマッピングする、ＥＥＥ１から１２のいずれかに記載の方法。

【０１６０】

ＥＥＥ１４．ソフトウェア命令を記憶する非一時的コンピュータ可読記憶媒体であって、ソフトウェア命令は、１つまたは複数のプロセッサによって実行されると、ＥＥＥ１から１３のいずれかに記載の方法を実行させる、非一時的コンピュータ可読記憶媒体。

【０１６１】

ＥＥＥ１５．１つまたは複数のプロセッサと、命令のセットを記憶する１つまたは複数の記憶媒体とを備えるコンピューティングデバイスであって、命令は、１つまたは複数のプロセッサによって実行されると、ＥＥＥ１から１３のいずれかに記載の方法を実行させる、コンピューティングデバイス。

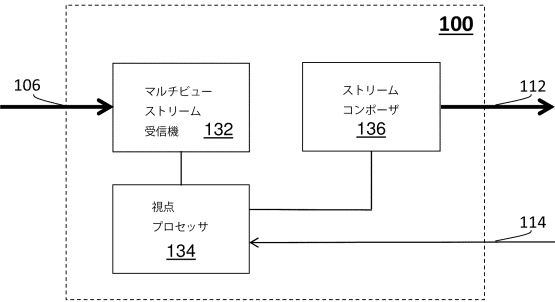
20

30

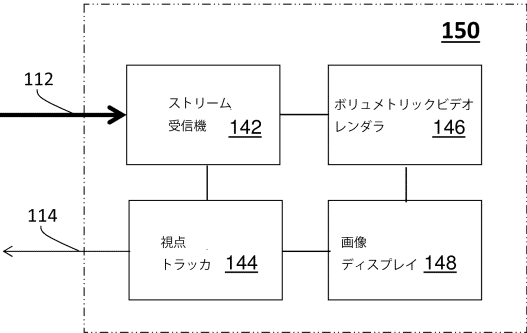
40

50

【図面】
【図 1 A】



【図 1 B】



【図 2 A】

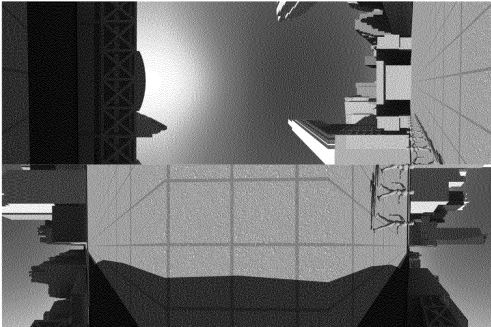


FIG. 2A

【図 2 B】

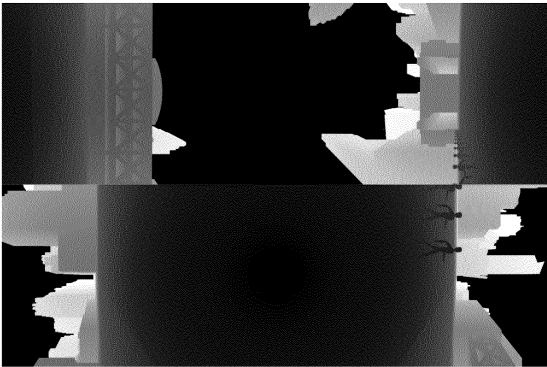


FIG. 2B

10

20

30

40

50

【図 2 C】

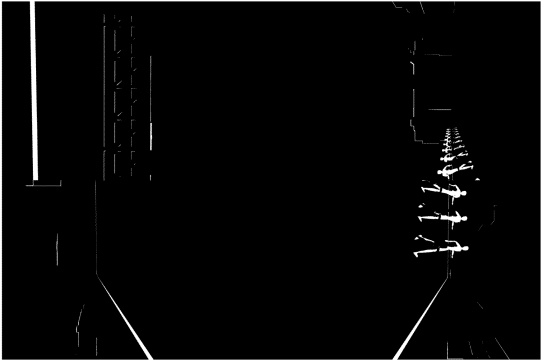


FIG. 2C

【図 3 A】

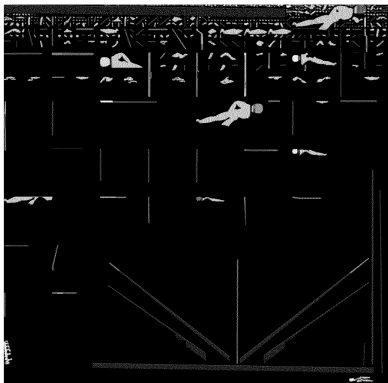


FIG. 3A

【図 3 B】



FIG. 3B

【図 3 C】



FIG. 3C

10

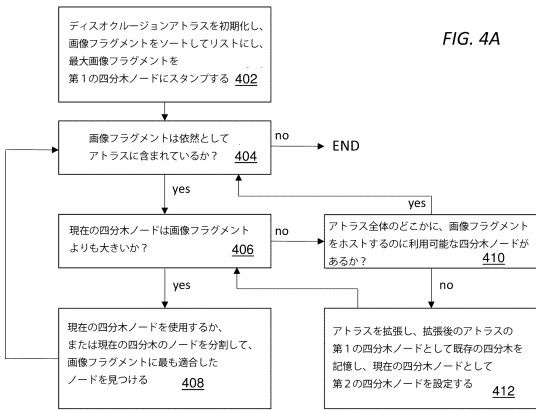
20

30

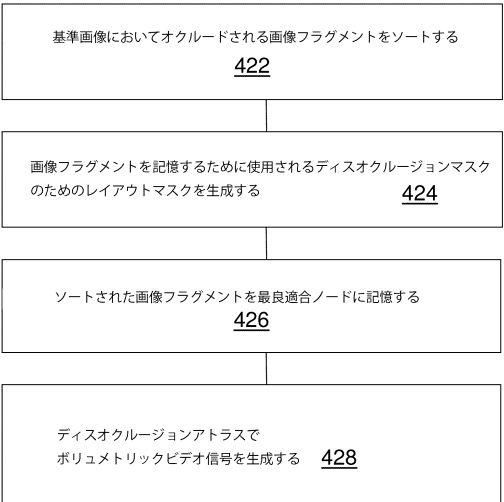
40

50

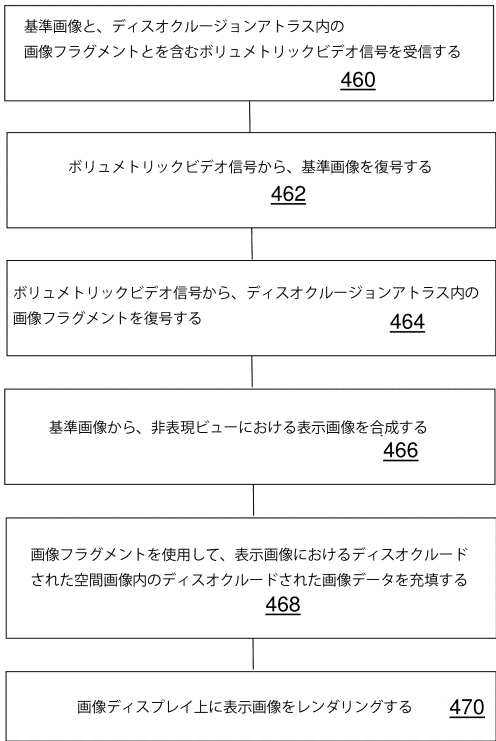
【図 4 A】



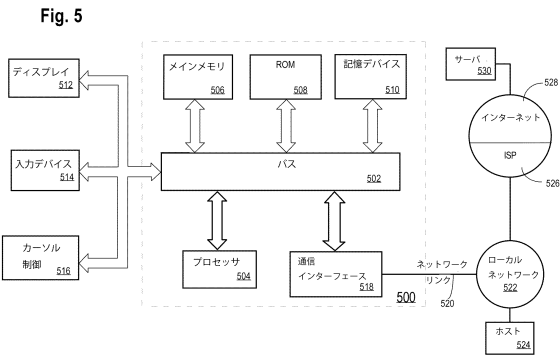
【図 4 B】



【図 4 C】



【図 5】



10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

ストリート 1 2 7 5 ドルビー ラボラトリーズ インコーポレイテッド 内

審査官 岩井 健二

(56)参考文献

国際公開第 2 0 1 7 / 0 8 0 4 2 0 (W O , A 1)

国際公開第 2 0 1 3 / 1 6 8 0 9 1 (W O , A 1)

国際公開第 2 0 0 9 / 0 0 1 2 5 5 (W O , A 1)

米国特許出願公開第 2 0 1 0 / 0 2 1 5 2 5 1 (U S , A 1)

(58)調査した分野 (Int.Cl., D B 名)

G 0 6 T 9 / 0 0 - 9 / 4 0

H 0 4 N 1 9 / 0 0 - 1 9 / 9 8