(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7469084号
(P7469084)

(45)発行日　令和6年4月16日(2024.4.16)

(24)登録日　令和6年4月8日(2024.4.8)

(51)国際特許分類
G06T 19/00 (2011.01)

FI
G06T 19/00 A

請求項の数 13 外国語出願 （全23頁）

(21)出願番号　特願2020-41913(P2020-41913)
(22)出願日　令和2年3月11日(2020.3.11)
(65)公開番号　特開2020-155121(P2020-155121A)
(43)公開日　令和2年9月24日(2020.9.24)
審査請求日　令和4年12月16日(2022.12.16)
(31)優先権主張番号　1903702.7
(32)優先日　平成31年3月19日(2019.3.19)
(33)優先権主張国・地域又は機関
英国(GB)

(73)特許権者　310021766
株式会社ソニー・インタラクティブエンタテインメント
東京都港区港南１丁目７番１号
(74)代理人　100105924
弁理士　森下 賢樹
(72)発明者　アンドリュー　ダミアン フォースフィールド
イギリス国 ダブリュ１エフ ７エルピー ロンドン グレート マールボロ ストリート １０、ソニー インタラクティブエンタテインメント ヨーロッパ リミテッド内
(72)発明者　ニコラ オール
イギリス国 ダブリュ１エフ ７エルピー

最終頁に続く

(54)【発明の名称】　画像を生成するための方法およびシステム

(57)【特許請求の範囲】

【請求項１】

仮想カメラの視点から被写体の画像を生成する方法であって、

シーン内の被写体の複数のソース画像を取得するステップであり、各ソース画像が異なるそれぞれの視点からキャプチャされる、ステップと、

ソース画像それぞれについて、前記被写体に対するカメラの姿勢を示すカメラ姿勢データを取得するステップと、

前記被写体に対する仮想カメラの姿勢を示す仮想カメラ姿勢データを取得するステップと、

前記ソース画像の少なくともいくつかについて、それぞれの歪み画像を生成するステップであり、各歪み画像が、前記ソース画像をキャプチャした前記カメラの前記姿勢と、前記仮想カメラの前記姿勢との差に基づいて生成される、ステップと、

前記歪み画像をキャプチャした前記カメラの前記姿勢と、前記仮想カメラの前記姿勢との類似性に基づいて各歪み画像に関連付けられた重みを決定し、前記決定された重みに従って前記歪み画像を合成して、前記仮想カメラの前記視点からの前記被写体の画像を生成するステップと、

を含み、

各歪み画像に関連付けられた重みを決定するステップが、少なくとも３つのカメラによって規定される平面と、前記仮想カメラを前記被写体と結ぶ線との交点を決定するステップを含む、方法。

【請求項２】
前記歪み画像を生成するステップが、
各ソース画像に関連付けられた深度データを取得するステップと、
ソース画像それぞれについて、前記ソース画像およびその画像に関連付けられた深度データに基づいて、前記被写体の多角形メッシュを生成するステップと、
多角形メッシュそれぞれについて、前記多角形メッシュの少なくとも一部を、前記仮想カメラが向いている方向に対して垂直になるように再投影するステップと、
を含む、請求項１に記載の方法。
【請求項３】
前記歪み画像を生成するステップが、
再投影された多角形メッシュそれぞれについて、その多角形メッシュに適用するためのテクスチャを生成するステップであって、前記テクスチャが、対応する前記ソース画像を前記多角形メッシュ上に視差マッピングすることによって生成される、ステップと、
をさらに含み、
前記歪み画像を生成するステップが、前記テクスチャが適用された前記それぞれの再投影された多角形メッシュの２Ｄ画像をレンダリングするステップを含む、
請求項２に記載の方法。
【請求項４】
各ソース画像が、緑色スクリーンの前でキャプチャされた前記被写体の画像に対応し、
前記歪み画像を生成するステップが、
ソース画像それぞれについて、前記被写体に対応するピクセルを、背景に対応するピクセルから分離するステップをさらに含む、
請求項１から３のいずれか一項に記載の方法。
【請求項５】
前記多角形メッシュを生成するステップが、
各ソース画像に二次元グリッドを適用するステップと、
それぞれのソース画像およびその画像に対して得られた前記深度データについて、前記二次元グリッドの各ユニットにわたって深度の変化を決定するステップと、
ソース画像それぞれについて、対応する多角形メッシュを生成するステップであって、前記多角形メッシュのテッセレーション密度が、前記ソース画像に適用された前記二次元グリッドの各ユニットにわたる深度の前記変化に基づいて決定される、ステップと、
を含む、請求項２、請求項３、請求項２を引用する請求項４のいずれか一項に記載の方法。
【請求項６】
各ソース画像が立体視カメラによってキャプチャされた立体視画像を含み、
少なくとも５つのソース画像が取得され、前記ソース画像に関連付けられた前記姿勢が、前記被写体に対して垂直方向および／または水平方向にｘ字形に配置された立体視カメラのアレイに対応する、
請求項１から５のいずれか一項に記載の方法。
【請求項７】
前記仮想カメラ姿勢データが、前記被写体の再構成を見るために使用されているＨＭＤの姿勢に対応し、
前記被写体の前記画像を生成するステップが、立体視画像を生成して、前記生成された立体視画像を前記ＨＭＤに出力するステップを含む、
請求項１から６のいずれか一項に記載の方法。
【請求項８】
請求項１から７のいずれか一項に記載の方法をコンピュータ・システムに実行させるように適合されたコンピュータ実行可能命令を有するコンピュータ可読媒体。
【請求項９】
オブジェクトの画像を視聴者に表示するように動作可能なディスプレイエレメントと、

前記視聴者の頭部の姿勢を検出するように動作可能な姿勢検出器と、
前記オブジェクトの複数のソース画像、および各ソース画像をキャプチャしたカメラの姿勢を示すカメラ姿勢データを取得するように動作可能な画像プロセッサであって、
前記姿勢検出器からの入力を受信し、それに応答して、前記ソース画像の少なくともいくつかを、前記ソース画像に関連付けられた前記カメラの前記姿勢と、前記オブジェクトの再構成に対する前記視聴者の頭部の現在の姿勢との差に基づいて歪ませるように構成されている、画像プロセッサと、
各歪み画像に関連付けられた重みを、その歪み画像に関連付けられた前記カメラの前記姿勢と、前記再構成に対する前記視聴者の頭部の前記姿勢との類似性に基づいて決定するように構成された画像合成器であって、
前記関連付けられた重みに従って前記歪み画像を一緒に合成して、前記ディスプレイエレメントに表示するための前記オブジェクトの画像を形成するように構成されている、画像合成器と、
を備え、
前記視聴者の頭部の前記姿勢が、前記オブジェクトを見る仮想カメラの前記姿勢に対応し、前記ソース画像をキャプチャした前記カメラの前記姿勢が、カメラアレイ平面に対応し、
前記画像合成器が、前記仮想カメラを前記オブジェクトと結ぶベクトルと、前記カメラアレイ平面との交点を決定するように構成され、
前記画像合成器が、前記ソース画像をキャプチャした前記カメラの前記姿勢および前記カメラアレイ平面上の前記交点に基づいて、各歪み画像の重みを決定するように構成されている、表示システム。
【請求項１０】
前記画像プロセッサが、各ソース画像に関連付けられた深度データを取得し、ソース画像それぞれについて、前記ソース画像およびその画像に関連付けられた深度データに基づいて前記オブジェクトの多角形メッシュを生成するようにさらに構成され、
前記画像プロセッサが、各ソース画像からの歪み画像を、そのソース画像に対して生成された前記多角形メッシュの少なくとも一部を、前記視聴者の頭部が向いている方向に対して垂直になるように再投影することによって生成するように構成されている、
請求項９に記載の表示システム。
【請求項１１】
前記画像プロセッサが、再投影された多角形メッシュそれぞれについて、その多角形メッシュに適用するためのテクスチャを生成するように構成され、
前記画像プロセッサが、前記ソース画像の部分を対応する前記再投影された多角形メッシュの部分に視差マッピングするように動作可能な視差シェーダを備える、
請求項１０に記載の表示システム。
【請求項１２】
ヘッドマウントディスプレイを備え、
前記ディスプレイエレメントが、前記ヘッドマウントディスプレイの前記ディスプレイエレメントに対応し、前記姿勢検出器が、前記ヘッドマウントディスプレイの位置および／または向きを検出するように動作可能である、
請求項９から１１のいずれか一項に記載の表示システム。
【請求項１３】
前記ヘッドマウントディスプレイと通信するゲームコンソールを備え、前記ゲームコンソールが、前記画像プロセッサを備え、前記合成された歪み画像を前記ヘッドマウントディスプレイに送信して前記ディスプレイエレメントに表示するように動作可能である、請求項１２に記載の表示システム。
【発明の詳細な説明】
【技術分野】
【０００１】

本開示は、被写体の画像を仮想カメラの視点から生成するための方法およびシステムに関する。本開示は、ヘッドマウントディスプレイ（ＨＭＤ）に表示するためのそのような画像を生成することにさらに関する。

【背景技術】

【０００２】

近年、ビデオコンテンツは、ますます没入型になっている。これは、一部には、視聴者がそのようなコンテンツを見ることができるＨＭＤが急増していることに起因する。ＨＭＤにより、視聴者は、自分のワールドビュー上に重畳されたコンテンツを、部分的に（拡張現実の場合）または完全に（仮想現実の場合）見ることができる。場合によっては、仮想コンテンツは、視聴者が位置する物理的な環境を考慮したやり方で視聴者に表示されることがある（いわゆる複合現実）。

【０００３】

一部のアプリケーションでは、視聴者に現実のイベントの３Ｄ再構成が提示されることがある。例えば、現実のイベントは、演奏者の録音、例えば、音楽の演奏、コンテンツの提示、フィクション・シリーズでの演技などに対応することがある。そのようなコンテンツの一例は、ＰＳ４上で（ＰＳＶＲを介して）利用可能な「Ｊｏｓｈｕａ　Ｂｅｌｌ　ＶＲ　Ｅｘｐｅｒｉｅｎｃｅ」である。

【０００４】

仮想現実において演奏者を表現するための１つの知られている技法は、演奏者の立体視ビデオをキャプチャし、このビデオを環境の３Ｄ再構成内のメッシュにテクスチャとして投影することを含む。そのような技法の例が英国特許出願第２０１８／０５２６７１号明細書に記載されている。この技法は、視聴者の頭部の動きがほとんど予測されない場合に十分である傾向があるが、頭部の動きが大きい場合は、この技法の限界がより明らかになる可能性がある。

【０００５】

仮想現実（ＶＲ）において演奏者を表現するための別の知られている技法は、例えば、リグを使用して、演奏者の周りに複数のカメラを配置することを含む。カメラによってキャプチャされたビデオを使用して、演奏者の完全な幾何学的モデルを作成し、このモデルを、モデルに対する視聴者の向きに基づいて視聴者に適切に表示することができる。理解されるように、このようにしてＶＲコンテンツをキャプチャすることは、時間がかかり、高価な処理である。さらに、このようなキャプチャ処理から生じるＶＲコンテンツは、多くの場合、生気がないように見え、ソフトエッジ（髪の毛など）や複雑な表面素材が、仮にあったとしても、正確に表現されない場合がある。

【０００６】

ライトフィールド・レンダリングは、オブジェクトを３Ｄで表現するために使用することができる別の技法である。しかしながら、ライトフィールド・レンダリングは、計算コストが高くなる傾向があり、多くの場合、静的オブジェクトのみの表現に限定される。

【０００７】

一般に、当技術分野において、３Ｄでのオブジェクト（静的および／または動的）のレンダリングを、過度な量のデータまたは計算処理を必要とすることなく行うことに関して問題がある。理解されるように、このレンダリングがビデオゲームコンソールで実行される場合、利用可能な計算能力とリソースは、ある程度制限される可能性がある。さらに、画像がＨＭＤに表示される場合、視聴者の頭部の動きと新たな視点からのオブジェクトのレンダリングとの間のいかなる遅延も、視聴者にとって不快なものになる可能性がある。本発明は、これらの問題のいくつかに対処すること、または少なくとも軽減しようとするものである。

【発明の概要】

【０００８】

本明細書に開示される第１の態様によると、請求項１に記載のヘッドマウントディスプレイ（ＨＭＤ）に表示するための被写体の画像を生成する方法が提供される。

【０００９】
本明細書に開示される第２の態様によると、請求項１０に記載のオブジェクトの画像を生成するためのシステムが提供される。
【図面の簡単な説明】
【００１０】
本開示の理解を助け、実施形態がどのようにして実施され得るかを示すために、例として添付の図面が参照される。
【００１１】
【図１】ヘッドマウントディスプレイデバイスの例を概略的に示す図である。
【図２】オブジェクトの画像を仮想カメラの視点から再構成するための方法の例である。
【図３】被写体の例、および被写体のソース画像をキャプチャするための１つまたは複数のカメラの配置を示す図である。
【図４】被写体に対して異なる位置に配置された１つもしくは複数のカメラによってキャプチャされた立体視ソース画像の例を示す図である。
【図５】対応する立体視ソース画像それぞれについて生成された深度マップの例を示す図である。
【図６Ａ】所与のソース画像および対応する深度マップのメッシュを生成するためのテッセレーション処理の例を示す図である。
【図６Ｂ】所与のソース画像および対応する深度マップのメッシュを生成するためのテッセレーション処理の例を示す図である。
【図７】生成されたメッシュの一部をメッシュからトリミングすることができるトリミング・ステップの例を示す図である。
【図８Ａ】対応するソースカメラと仮想カメラの姿勢との間の姿勢の差に基づいて再投影されるメッシュの部分の例の概略図である。
【図８Ｂ】対応するソースカメラと仮想カメラの姿勢との間の姿勢の差に基づいて再投影されるメッシュの部分の例の概略図である。
【図９】ソース画像の領域を、対応する再投影されたメッシュの領域にマッピングするための視差マッピング技法の例である。
【図１０Ａ】各歪み画像に関連付けられた重みを決定するための方法の例の概略図である。
【図１０Ｂ】各歪み画像に関連付けられた重みを決定するための方法の例の概略図である。
【図１１】本開示に従って生成された歪み画像の例、およびこれらの画像を一緒に混合することによって得られた最終画像を示す図である。
【図１２】オブジェクトの画像を仮想カメラの視点から生成するためのシステムの例である。
【発明を実施するための形態】
【００１２】
ここで図１を参照すると、ユーザ１０は、（一般的な頭部装着可能な装置の例として）ＨＭＤ２０を着用している。ＨＭＤは、本例では、後部ストラップおよび頂部ストラップで形成されたフレーム４０、ならびに表示部５０を備える。
【００１３】
図１のＨＭＤは、周囲環境に対するユーザのビューを完全に（または少なくとも実質的に完全に）覆い隠す。ユーザが見ることができるのは、ＨＭＤ内に表示される画像のペアだけである。
【００１４】
ＨＭＤは、ユーザの左右の耳７０に適合する関連付けられたヘッドフォン・オーディオ・トランスデューサまたはイヤピース６０を有する。イヤピース６０は、外部ソースから提供されるオーディオ信号を再生し、外部ソースは、ユーザの目に表示用のビデオ信号を提供するビデオ信号源と同じであってもよい。
【００１５】
ユーザがＨＭＤによって表示されたもののみを見ることができるという事実、ならびに

イヤピースのノイズ遮断特性またはアクティブキャンセル特性、および関連付けられた電子機器の制限を受けて、イヤピースを介して提供されるもののみを聞くことができるという事実の組合せは、このＨＭＤをいわゆる「完全な没入型」ＨＭＤと見なすことができることを意味する。しかしながら、一部の実施形態では、ＨＭＤは、完全な没入型ＨＭＤではなく、ユーザの周囲を見るかつ／または聞くための少なくとも何らかの機能をユーザに提供することができることに留意されたい。この機能は、ディスプレイ装置にある程度の透明度または部分的な透明度を提供すること、および／またはＨＭＤのディスプレイを介して（カメラ、例えば、ＨＭＤに取り付けられたカメラを使用してキャプチャされた）外部のビューを投影すること、および／またはイヤピースを通過した周囲音の伝播を可能にすること、および／または周囲音に応じて（イヤピースへの伝播のための）入力音声信号を生成するマイクロフォンを提供することによるものであってもよい。

【００１６】

前方カメラ（図示せず）は、使用時に、ＨＭＤの前方に画像をキャプチャすることができる。Ｂｌｕｅｔｏｏｔｈ（登録商標）アンテナ（図示せず）は、通信機能を提供することができ、または近くのＢｌｕｅｔｏｏｔｈ送信機の方向の検出を可能にする指向性アンテナとして単に配置されてもよい。

【００１７】

動作において、ビデオ信号は、ＨＭＤによる表示のために提供される。このビデオ信号は、ビデオゲーム機またはデータ処理装置（パーソナルコンピュータなど）などの外部ビデオ信号源８０によって提供することができ、その場合、信号は、有線または無線接続によってＨＭＤに送信されてもよい。適切な無線接続の例には、Ｂｌｕｅｔｏｏｔｈ（登録商標）接続が挙げられる。イヤピース６０のためのオーディオ信号は、同じ接続によって運ばれてもよい。同様に、ＨＭＤからビデオ（オーディオ）信号源に渡される任意の制御信号は、同じ接続によって運ばれてもよい。

【００１８】

さらに、（１つもしくは複数のバッテリを含む、かつ／または主電源コンセントに接続可能な）電源をケーブルによってＨＭＤにつなぐことができる。電源およびビデオ信号源８０は、別個のユニットであってもよく、または同じ物理ユニットとして具現化されてもよいことに留意されたい。電源およびビデオ（実際にはオーディオ）信号供給用の別個のケーブルがあってもよく、またはこれらのケーブルは、単一のケーブルで運ぶために組み合わされてもよい（例えば、ＵＳＢケーブルのように別個の導体を使用して、またはデータが平衡信号として運ばれ、電力が直流として、同じ物理的なワイヤの集合体を介して運ばれる「パワー・オーバ・イーサネット（登録商標）」構成と同様の仕方で）。ビデオおよび／またはオーディオ信号は、例えば、光ファイバケーブルによって運ばれてもよい。他の実施形態では、ユーザに提示するための画像および／またはオーディオ信号の生成に関連付けられた機能の少なくとも一部は、ＨＭＤ自体の一部を形成する回路および／または処理によって実行されてもよい。電源は、ＨＭＤ自体の一部として提供されてもよい。

【００１９】

したがって、図１の構成は、観察者の頭部に取り付けられたフレームであって、使用時に、観察者のそれぞれの目の前に位置決めされる１つまたは２つの目の表示位置を規定するフレームと、目の表示位置のそれぞれに対して取り付けられたディスプレイエレメントであって、ビデオ信号源からのビデオ信号のビデオ表示の虚像を観察者のその目に提供するディスプレイエレメントと、を備えるヘッドマウントディスプレイシステムの例を提供する。

【００２０】

図１は、ＨＭＤの単なる一例を示す。他の形態が可能であり、例えば、ＨＭＤは、従来のメガネに関連付けられたものにより類似したフレーム、すなわち、表示部分からユーザの耳の上部後方に延び、場合によっては耳の後ろでカールする実質的に水平な脚部を使用することができる。他の（完全な没入型ではない）例では、外部環境のユーザのビューが実際には覆い隠されていない場合があり、表示される画像を、外部環境上に（ユーザの視

点から）重ね合わせるように配置することができる。
【００２１】
図１に示すＨＭＤを使用して、没入型ビデオを視聴者に表示することができる。没入型ビデオという用語は、本明細書では、仮想現実、拡張現実、および／または混合現実での視聴に適したビデオコンテンツを指すために使用される。一般に、ビデオは、少なくとも部分的に、１つまたは複数の現実世界のオブジェクトを含む現実シーンのビデオに対応する。没入型ビデオは、現実世界のオブジェクトの画像が、視聴者にそれらのオブジェクトを３Ｄで表すために使用されているという点で、ゲームコンテンツとは異なることがある。しかしながら、それでもなお、没入型ビデオにより、視聴者は、仮想現実環境、拡張現実環境、または混合現実環境で表されるオブジェクトと対話することが可能である場合がある。
【００２２】
前述したように、没入型ビデオは、演奏のビデオに対応する場合ある。没入型ビデオコンテンツの例には、例えば、ＰＳ４で利用可能な「Ｊｏｓｈｕａ　Ｂｅｌｌ　ＶＲ　Ｅｘｐｅｒｉｅｎｃｅ」、「Ｔｏｍ　Ｇｒｅｎｎａｎ　ＶＲ」などが挙げられる。文脈上の目的で、本明細書で説明される実施形態は、主に、ＶＲでの没入型ビデオコンテンツの表示に関連して説明される。
【００２３】
前述したように、没入型ビデオコンテンツを生成するための多くの技法が存在する。これらには、例えば、シーンをボリュメトリック・ビデオとしてキャプチャすること、シーンの立体視ビデオをキャプチャし、これをシーンの３Ｄ再構成内のメッシュ上に投影すること、ライトフィールド・レンダリングなどが含まれる。しかしながら、典型的には、これらの技法は、ビデオをレンダリングするために必要とされるリソースの点で、時間がかかり、高価であり、計算集約的である。
【００２４】
ビデオゲームコンソールが没入型ビデオのレンダリングに使用される場合、利用可能な計算能力は、ある程度制限される可能性がある。さらに、ビデオのレンダリングのいかなる遅延も、視聴者の頭部の動きと同期してＨＭＤでのビデオの表示が更新されない場合、視聴者に不快感または吐き気さえ感じさせることがある。したがって、当技術分野では、上述の技法に関連付けられた障害の影響をあまり受けない、没入型ビデオコンテンツを生成することができるレンダリング処理が必要とされている。
【００２５】
ここで、本開示による、ＨＭＤに表示するための被写体の画像を生成するための方法を、図２に関連して説明する。ここでは、「被写体」は、没入型ビデオの一部として表示される被写体に対応する。
【００２６】
第１のステップＳ２０１において、シーン内の被写体の複数のソース画像が取得される。被写体は、没入型ビデオの一部として表される静的オブジェクトまたは動的オブジェクトに対応してもよい。各ソース画像は、異なるそれぞれの視点からキャプチャされる。すなわち、ソース画像それぞれについての被写体に対するカメラの姿勢は、他のソース画像に対するカメラの姿勢とは異なる。その結果、各ソース画像は、被写体の異なるビューに対応する。姿勢という用語は、本明細書では、被写体に対するカメラの位置および／または向きを指すために使用される。
【００２７】
各ソース画像は、立体視カメラによってキャプチャされた立体視画像に対応することができる。一部の例では、複数の立体視カメラ（例えば、ソース画像それぞれについて１つの立体視カメラ）を使用して、立体視画像をキャプチャすることができる。すなわち、立体視カメラのアレイが被写体の周りに配置されてもよく、ソース画像がアレイ状の立体視カメラによってキャプチャされた立体視画像に対応することができる。追加のまたは代替の例では、単一の立体視カメラを使用して、この立体視カメラを被写体の周りの異なる位

置に移動させて、異なるそれぞれの視点から被写体の複数のソース画像を取得してもよい。立体視画像は、カラー画像、例えば、ＲＧＢまたはＹＵＶ画像に対応してもよい。一部の例では、ソース画像は、本明細書で説明される処理の前に圧縮されてもよい。

【００２８】

一部の例では、各立体視カメラは、２つの画像センサを備え、各画像センサは、対応するセンサ上に光を集束するための１つまたは複数のレンズに関連付けられている。画像センサによってキャプチャされた画像は、左画像および右画像にそれぞれ対応することができる。左の画像および右の画像は、（以下に説明するように、処理の後に）ＨＭＤの左および右のディスプレイエレメントにそれぞれ表示されてもよい。

【００２９】

追加のまたは代替の例では、カメラの少なくともいくつかは、単眼視カメラ（すなわち、１つの画像センサ、およびそのセンサ上に光を集束させるための１つまたは複数のレンズを備える）を含むことができる。これらは、後述するように、飛行時間構造化光センサなどの追加の深度センサと組み合わせて使用されることがある。

【００３０】

一部の例では、各ソース画像は、緑色スクリーンの前でキャプチャされた被写体の画像に対応してもよい。緑色スクリーンに対応するソース画像の部分を削除して、被写体をその環境から分離することができる。緑色スクリーンを使用することにより、３Ｄで表現することができる被写体の高品質のソフトエッジをキャプチャすることができる。

【００３１】

図３は、被写体の例、および被写体のソース画像をキャプチャするための立体視カメラの位置を示す。図３では、被写体は、緑色スクリーン上におよびその前に配置されたぬいぐるみライオンの玩具３００に対応する。これは単なる被写体の例であり、演奏者などの他の被写体も同じやり方でキャプチャされ得ることが理解されよう。図３では、被写体およびカメラは、上面図で示されている。

【００３２】

図３では、被写体に対する立体視カメラの姿勢が、ｘ字形に配置された立体視カメラのアレイに対応していることがわかる。すなわち、一対の立体視カメラ３０２Ａ、３０２Ｂが、被写体の左右にそれぞれ配置され、別の立体視カメラ３０２Ｃが、第１の対の立体視カメラの３０２Ａ、３０２Ｂの後方に、被写体の前方のほぼ中央に配置され、第２の対の立体視カメラ３０２Ｄ、３０２Ｅが、中央に配置された立体視カメラの後方に、ここでもまた被写体の左右にそれぞれ配置されている。一部の例では、このｘ字形は、単一の立体視カメラを、ｘ字形を構成する個々の位置のそれぞれに移動させることによってエミュレートされてもよい。

【００３３】

追加のまたは代替の例では、立体視カメラは、異なる垂直高さに配置されてもよい。例えば、ｘ字形は、（または代替として）カメラの垂直配置に対応することもできる。これは、第１の対の立体視カメラが、下方の対の立体視カメラの上方に配置され、２つの対の立体視カメラ間の（垂直方向の）中央の位置を立体視カメラが占有してもよいことを意味する。このようなカメラの配置は、（ｚが深度に対応する）ｘ－ｙ平面内で同一平面上にあるということができる。このように配置された立体視カメラアレイの例については、図１０に関連して後述する。

【００３４】

理解されるように、他の例では、カメラは、同一平面上にある必要はなく、カメラのうちの１つまたは複数は、ｚ方向にオフセットされていてもよい。場合によっては、カメラの非同一平面配置により、深度マップの計算に使用される写真測量の精度を向上させることができる。

【００３５】

立体視カメラの相対的な分離は、３Ｄで被写体を再構成するために利用可能なソース画像データの量を決定することが理解されるであろう。遮られた領域が見えるようになるよ

うに、カメラをさらに離すことによって、被写体のより良好なカバレッジを得ることができる。カバレッジが良好なほど、３Ｄで表現することができる被写体が多くなり、したがって、再構成を見るときに視聴者が有する動きの自由度が大きくなる。しかしながら、カメラの間隔が離れすぎた場合は、ソース画像間で共有される画像データが少なくなり、ソース画像を一緒に混合しようと試みる際に困難が生じる可能性がある。当然のことながら、被写体を３Ｄで表現することができる精度は、異なるそれぞれの視点からより多くの画像をキャプチャすることによって改善され得る。しかしながら、より多くの画像をキャプチャすると、３Ｄ再構成を生成するために必要なメモリおよび処理能力の点で、例えば、ボリュメトリック・ビデオについては高いことが知られているように、より大きなコストを招く。したがって、視点の数は、計算リソースおよび３Ｄ再構成の品質との間の経験的なトレードオフである場合がある。

【００３６】

本発明者らは、図３に示すｘ配置が、被写体に対する視聴者の動きが相対的に制限されると予想される没入型ビデオ目的には十分であることを見出した。この動きは、コンテンツが典型的に消費される性質の結果として、例えば、（着座などの）静止位置から、またはコンテンツが消費されている物理的環境から、制限されることがある（例えば、視聴者がＶＲで表現される被写体の後方を歩くことを試みるのに十分な物理的空間がない場合がある）。したがって、没入型ビデオでは、被写体は、制限された範囲のビューに対してのみ表現されればよい。加えて、図３に示す配置を介してキャプチャされたソース画像には、重複領域が含まれ、結果としてソース画像を一緒に合成したときに「孔（ｈｏｌｅ）」が少なくなる。

【００３７】

立体視カメラをｘ字形に配置することは、被写体に対して立体視カメラを配置することができるやり方の単なる一例であることが理解されよう。他の例では、代替のカメラキャプチャレイアウトを使用することができる。最終的に、使用されるレイアウト、およびキャプチャされる画像の数は、キャプチャされる被写体およびそれらの画像を処理するために使用されるハードウェアに依存する場合がある。より一般的には、図３に見られるｘ字形は、視聴者にとって所定の動きの範囲内の視点の均一な分布の一例として考えられてもよい。

【００３８】

図４は、各立体視カメラのそれぞれによってキャプチャされた立体視画像４０２Ａ、４０２ｂ、４０２、４０２Ｄ、４０２Ｅの例を示す。図４では、立体視画像の位置決めは、被写体に対する立体視カメラの位置決めに対応する。図４に示す立体視画像は、上述のｘ字形に配置された立体視カメラアレイによってキャプチャされている。

【００３９】

図２に戻ると、ステップＳ２０１において、ソース画像それぞれについてシーン内の被写体に対するカメラの姿勢を示すカメラ姿勢データが取得される。カメラ姿勢データは、ソース画像それぞれについて、被写体に対するカメラの位置および／または向きを規定する。カメラ姿勢データは、ソース画像に関連付けて記憶されたメタデータとして規定されてもよい。

【００４０】

カメラ姿勢データは、ソース画像のキャプチャ中に取得されてもよい。例えば、被写体のカメラからの距離は、例えば、飛行時間センサを使用して測定され、カメラの向きの変化は、例えば、カメラに関連付けられた加速度計またはジャイロスコープを使用して追跡されてもよい。一部の例では、カメラの相対的な姿勢は、（例えば、複数のカメラのアレイが使用される場合）キャプチャ処理中に変化しない場合があり、被写体に対するカメラの姿勢は、知られている写真測量技法を使用して決定されてもよい。例えば、チェス盤などの試験被写体を使用して、カメラの外的要因を決定する。

【００４１】

第２のステップＳ２０２において、被写体に対する仮想カメラの姿勢を示す仮想カメラ

姿勢データが取得される。仮想カメラのデータは、被写体の再構成が視認可能なＨＭＤの姿勢に対応してもよい。ＨＭＤの姿勢は、被写体の再構成に対して視聴者がどのように位置し、向いているか決定し、したがって、被写体を視認可能にする仮想カメラに対応する。被写体に対する仮想カメラの姿勢を決定するために、被写体の再構成をＨＭＤに実際に表示する必要がないことが理解されよう。仮想カメラの姿勢は、レンダリング時に決定することができる。

【００４２】

仮想カメラ姿勢データは、ＨＭＤに組み込まれた姿勢検出器から取得されてもよい。例えば、ＨＭＤは、視聴者の頭部の姿勢の変化を追跡するための加速度計および／またはジャイロスコープを備えてもよい。代替として、または加えて、１つまたは複数の外部カメラを使用して、視聴者の頭部の姿勢を、例えば、コンスタレーション追跡を介して、追跡することができる。再構成されたシーン内の被写体の位置がわかっている場合があり、ＨＭＤの姿勢を使用して、再構成されたシーンにおいて被写体に対する視聴者の頭部の（仮想）姿勢を決定することができる。

【００４３】

図２に戻ると、第３のステップＳ２０３において、ソース画像の少なくともいくつかを歪ませる。ソース画像をキャプチャしたカメラの姿勢と、被写体の再構成に対する仮想カメラの姿勢との差に基づいて各ソース画像を歪ませる。言い換えれば、所与のソース画像の歪みの程度は、仮想カメラのビューが、ソース画像をキャプチャしたカメラのビューにどれだけ厳密に一致しているかに依存する。

【００４４】

［歪み画像の生成］

ここで、歪み画像を生成するための例示的な方法を、図４～図１１に関連して説明する。

【００４５】

一部の例では、歪み画像を生成することは、各ソース画像に関連付けられた深度データを取得することを含む。前述したように、各ソース画像は、被写体の立体視画像に対応することができ、したがって、ソース画像を構成する左右の画像に対してステレオマッチングを実行することによって、所与のソース画像について深度データを取得することができる。所与のソース画像について取得された深度データを使用して、その画像内の被写体の深度マップ（すなわち、深度画像）を生成することができる。被写体が緑色スクリーンの前に置かれている場合、それらの画像から深度マップを生成する前に、背景に対応するソース画像のピクセルを除去することができる。代替として、ある特定の深度しきい値を超える深度マップの領域を背景領域として排除することができる。各深度マップは、対応するソース画像をキャプチャしたカメラと同じ視点からキャプチャされた深度画像に対応することができる。一部の例では、深度マップは、その深度マップからメッシュを生成する前に平滑化される。

【００４６】

図５は、図３および図４に従ってキャプチャされた立体視画像について生成された深度マップ５０２Ａ、５０２Ｂ、５０２Ｃ、５０２Ｄ、５０２Ｅの例を示す。図５では、各深度マップは、深度マップの生成に使用される立体視画像がキャプチャされた立体視カメラの位置に対応する位置に示されている。図５では、深度マップとソース画像との間に１対１の対応関係がある。

【００４７】

代替の例では、各ソース画像の深度データは、ソース画像のキャプチャとは別個にキャプチャされてもよい。例えば、ソース画像それぞれについて、被写体のＬＩＤＡＲスキャンが、対応する視点から実行されてもよい。そのような例では、ソース画像は、立体視的である必要はない。対応する視点の深度データを取得するための他の知られている技法が使用されてもよい。

【００４８】

ソース画像は、ＲＧＢピクセル、および所望により各ピクセルの透明度を表すアルファ

値に関して規定されてもよい。ソース画像それぞれについて、緑色スクリーンに対応する画像内の領域および対応する深度マップを使用して透明マスクが生成されてもよい（すなわち、均一な深度の大きな領域は、背景領域に属するものとして識別されてもよい）。透明マスクのエッジからの距離は、必要に応じて、透明マスクのエッジを簡単に拡大または縮小することができるように、符号付き距離フィールド（ｓｉｇｎｅｄ　ｄｉｓｔａｎｃｅ　ｆｉｅｌｄ）（ＳＤＦ）に記憶することができる。ソース画像が立体視画像を含む例では、各立体視画像を構成する左右の画像に対して透明マスクを生成することができる。

【００４９】

一部の例では、被写体の深度マップをわずかに拡大して、アルファ輪郭を超えて、任意の孔（すなわち、深度データが欠落している領域）を塗りつぶすことができる。所与のソース画像の深度マップは、複数のミップマップとして記憶することができ、連続する各ミップマップは、より低い解像度の深度マップに対応する。ミップマップレベルは、画像品質に基づいて事前に決定されてもよく、すなわち、ミップマップレベルは、データがどの程度ノイズが多いかに基づいて調整することができるパラメータであってもよい。深度ミップマップは、上記のミップマップレイヤの例えば４個のピクセルの最も近い（カメラに最も近い）深度を記憶することによって生成されてもよい。

【００５０】

歪み画像を生成することは、ソース画像それぞれについて、ソース画像およびそのソース画像について取得された深度データに基づいて、被写体の多角形メッシュ（例えば、三角形メッシュ）を生成することをさらに含むことができる。各多角形メッシュは、それぞれのソース画像に二次元グリッドを適用し、そのソース画像について、二次元グリッドの各ユニットにわたる深度の変化を決定することによって生成することができる。メッシュのテッセレーション密度は、対応するソース画像に適用される二次元グリッドの各ユニットにわたる深度の変化に基づいて決定される。このテッセレーション処理の一例を図６Ａおよび図６Ｂに示す。

【００５１】

図６Ａでは、三角形タイルを有する二次元グリッド６０４が、所与のソース画像に関連付けられた深度マップ６０２Ａと位置合わせされているものとして示されている。図６Ｂでは、同じグリッド６０４が、対応するソース画像６０２Ｂと位置合わせされて示されており、このソース画像は、この場合、立体視画像を構成するカラー画像の１つに対応する。図６Ａおよび図６Ｂでは、ソース画像と、対応する深度マップの部分とが正確に位置合わせされている。タイルのｘ座標とｙ座標は、スクリーンで見たときに等間隔になるように設定され、ｚ座標は、対応する深度マップの適切なミップマップレベルをサンプリングすることによって決定される。テッセレーション密度は、三角形タイルごとに、タイルの頂点にわたる深度の変化量によって決定される。これにより、確実に、可能な限り多くの表面の形状が幾何学形状に取り込まれる。

【００５２】

ソース画像の多角形メッシュの生成に続いて、メッシュの少なくとも一部分がトリミングされてもよい。このトリミングは、メッシュの少なくとも一部が被写体の輪郭に対応するしきい値境界の外側にあるという判定に基づいて実行することができる。

【００５３】

このトリミングの例が図７に示されており、最上部の画像７０２は、被写体を表す符号付き距離フィールドに対応し、しきい値境界７０４が重畳されている。しきい値境界７０４は、ＳＤＦに基づいて決定されてもよく、すなわち、被写体の形状の「内側」と「外側」との間の境界がＳＤＦによって規定されてもよい。図７では、最下部の画像７０６は、図６に示すソース画像に対応し、三角形メッシュが重畳されている。三角形は、境界７０４からのしきい値距離を規定することによって、メッシュから除去されてもよく、このしきい値距離を超えると、メッシュから三角形が除去される。一例では、境界７０４の外側の４ピクセルを超える位置にある任意の三角形が除去されることがあってもよい。最下部の画像７０６では、領域７０８は、内部に含まれているメッシュの任意の三角形が除去さ

れる領域に対応する。

【００５４】

ソース画像それぞれについて多角形メッシュを生成した後、歪み画像は、それぞれのメッシュの少なくとも一部を、仮想カメラが向いている方向に対して垂直になるように再投影することによって生成されてもよい。すなわち、ソース画像それぞれについて、歪み画像は、対応するメッシュの少なくとも一部を、仮想カメラに面するように再投影することによって生成されてもよい。この技法の例を図８Ａおよび図８Ｂに示す。

【００５５】

図８Ａでは、カメラ８０２、および被写体を表すメッシュ８０４Ａのテッセレーションされた表面の上面図が示されている。被写体に対するカメラ８０２の位置および向きは、ソース画像のうちの１つについての被写体に対するカメラの位置および向きに対応する。図８Ｂでは、メッシュの頂点は、レンダーカメラの位置に対応する仮想カメラの位置（位置「ｘ」で示す）に従って再投影されたものとして示されている。図８Ｂでは、メッシュの再投影部分が再投影部分８０４Ｂとして示されている。図８Ｂからわかるように、これらの部分の頂点は、ビュー平面に対して平行（および仮想カメラのビュー方向に対して垂直）になるように再投影されている。この再投影は、例えば、スクリーン空間において、三角形のｘ値およびｙ値は、保持されるが、ｚ値が３つの頂点の視点に最も近い値に設定されるように、または代替として３つの頂点の中央値または平均値に設定されるように、多角形メッシュ内の三角形の少なくとも一部を再位置合わせすることを含むことができる。仮想カメラに面するようにメッシュの部分を再投影するこの処理は、ソース画像それぞれについてのカメラの姿勢および仮想カメラの姿勢に基づいて、メッシュそれぞれについて繰り返されてもよい。

【００５６】

各メッシュの少なくとも一部を再投影すると、それらのメッシュに適用するためのテクスチャを生成することができる。すなわち、ソース画像から歪み画像を生成することは、再投影されたメッシュそれぞれについて、そのメッシュに適用するためのテクスチャを生成することを含むことができる。所与のメッシュのテクスチャは、対応するソース画像をメッシュに視差マッピングすることによって生成されてもよい。このマッピングを実行するために、視差シェーダを使用することができる。視差シェーダは、仮想カメラのビューにマッピングする際に、ソース画像の視点の変化を補正する。（所与のソース画像についての）各歪み画像は、テクスチャが適用された対応する再投影されたメッシュのレンダリングに対応することができる。

【００５７】

図９は、ソース画像から取得された色情報が、（上述したように）少なくとも一部が再投影された対応するメッシュにテクスチャとして適用され得る技法の例を示す。ここで、メッシュは、ソース画像を介して生成されたという意味で「対応」している。すなわち、各メッシュは、異なるソース画像を使用して生成されてもよい。

【００５８】

図９の左側には、深度マップ９０２の一部、およびソース画像の対応する部分（歪みのない）９０４Ａが示されている。この部分内に、領域９０６Ａが、視差マッピングが適用されるテクスチャ領域に対応して示されている。図９に示す領域９０６は、対応するメッシュの多角形（例えば、三角形）上にマッピングされるソース画像の一部に対応することができる。同じテクスチャ領域９０６Ｂが、視差マッピングに続いて、図９の右側にソース画像９０４Ｂの歪んだ部分内に示されている。

【００５９】

図９の中央領域には、被写体９０８の表面が、対応するソース画像および被写体の深度マップが取得されたソースカメラ９１０に対して示されている。図９に示す被写体９０８の表面は、深度マップから生成された多角形のプロキシジオメトリ表面として表されている。

【００６０】

ソースカメラの姿勢は、問題となっているソース画像がキャプチャされたカメラの姿勢に対応する。図９では、（位置「ｘ」で示す）仮想カメラの姿勢も、被写体に対して示されている。仮想カメラの姿勢は、被写体に対する視聴者の姿勢に対応する。

【００６１】

図９では、第１の光線９１２が、（メッシュの再投影された多角形に対応する）ＶＲ画像に平行な多角形エッジから開始して示されており、光線マーチングを使用して深度マップの表面エッジに当たるまで投射される。次いで、光線を追跡して、ソーステクスチャ上のその位置を見出す。これは、（第２の光線９１４を介して示されるように）ＶＲ画像に平行な多角形エッジに沿って複数の光線について繰り返される。これから、ソース画像の所与の部分（すなわち、ソーステクスチャ）をＶＲ視点の画像空間内にどのように湾曲（ｗａｒｐ）させるべきかを決定することができる。ソーステクスチャの所与の部分について適切な湾曲が決定されると、ソーステクスチャのその部分をメッシュの対応する多角形に当てはめることができる。

【００６２】

［歪み画像の合成］

図２に戻ると、第４のステップＳ２０４において、歪ませたソース画像それぞれについて重みが決定される。所与のソース画像の重みは、歪み画像をキャプチャしたカメラの姿勢と、仮想カメラの姿勢との類似性（すなわち、実際のカメラと仮想カメラのビューとの一致度）に基づいて決定される。これは、ソースカメラと、仮想カメラの姿勢との類似性が大きいほど、最終的な混合された画像における歪み画像の寄与が大きくなることを意味する。一部の例では、所与の歪み画像の重みは、対応するソースカメラの、仮想カメラおよび被写体への近接度に基づいて決定することができる。

【００６３】

図１０Ａは、歪み画像に関連付けられた重みを決定することができる方法の一例を示す。図１０Ａでは、仮想カメラ１００２は、被写体１００４に対して所定の姿勢で示されている。仮想カメラを被写体と結ぶ線は、破線の矢印１００６を介して示されており、この矢印は、ｚ方向の被写体の距離の表示を提供する。被写体に対するソースカメラの姿勢は、姿勢１００８で示されている。図１０Ａおよび図１０Ｂでは、ソースカメラは、共平面配置で配置されている。

【００６４】

図１０Ａは、被写体、ソースカメラ、および仮想カメラの主に上から見下ろした図に対応する。図１０Ｂは、同じソースカメラを横断面図で示す。すなわち、図１０Ｂでは、ソースカメラの相対的な水平位置および垂直位置が示されている（しかし、被写体からのそれらの相対的な距離は示されていない）。

【００６５】

歪み画像に関連付けられた重みは、仮想カメラを被写体と結ぶ線と、カメラアレイ平面との交点を計算することによって決定することができる。図１０Ａおよび図１０Ｂでは、この交点は右端の３つのカメラについて「ｘ」としてマークされている。カメラアレイ平面は、少なくとも３つの隣接するカメラを結ぶ平面として規定されてもよい。例えば、図１０Ａおよび図１０Ｂでは、ｘ－ｙ平面は、中央のカメラを中心とした４つのスライスに分割されて示されており、各スライスは、異なるサブセットの３つのカメラによって規定されるカメラアレイ平面に対応する。

【００６６】

所与の歪み画像に関連付けられた重みは、所与のカメラアレイ平面内の対応するソースカメラと、（「ｘ」とマークされた）交点との間の距離に基づいて決定することができる。例えば、右端の３つのカメラについては、「ｘ」とマークされた点と、そのカメラの位置との間の距離が小さいほど、各カメラに関連付けられた重みが大きくなり得る。所望により、ますます遠く離れたカメラからの寄与も使用されてもよいことが理解されよう。

【００６７】

一部の例では、各歪み画像に関連付けられた重みを決定することは、観察者および被写

体からの光線と、カメラアレイ平面を構成する三角形のそれぞれとの交点の符号付き重心座標を決定することを含むことができる。典型的な非縮退の使用事例では、三角形のうちの１つについては、重心座標のうちの２つまたは３つが正であり、合致したビューに対する重みとして使用することができる。

【００６８】

一般に、所与の歪み画像の重みは、対応するソースカメラの、仮想カメラおよび被写体への近接度に基づいて決定することができる。図１０Ａに示す例では、例えば、右上のカメラによってキャプチャされたソース画像と比較して、左上のカメラによってキャプチャされたソース画像により低い重みが割り当てられてもよい。これは、左上のカメラのビューが仮想カメラのビューとあまり一致しないためである。

【００６９】

一部の例では、各ソースカメラが向いている方向がわかっている場合があり、この情報を、仮想カメラを被写体と結ぶベクトルと比較して、各ソースカメラが仮想カメラとどのくらい厳密に位置合わせされているかを判定することができる。ソースカメラおよび仮想カメラの位置を使用して、ソースカメラが仮想カメラとどのくらい厳密に位置合わせされているかも判定することができる。繰り返すが、所与のソースカメラが仮想カメラに近くなるほど、そのソースカメラを介して生成された歪み画像に割り当てられる重みが大きくなる。

【００７０】

図１０Ｂに示すカメラは、共平面配置にあるが、他の配置が使用されてもよいこと、および異なる（実際には複数の）カメラアレイ平面が存在してもよく、そのカメラアレイ平面について、その平面またはサブ平面を規定するカメラに対して交点が決定されて、それらのカメラから取得された歪み画像に関連付けられた重みを決定することができることが理解されよう。

【００７１】

上述した技法は、中央のカメラと、中央のカメラの周りに対称または非対称に配置されたＮ個の遠位カメラと、が存在する任意のカメラ配置に拡張することができる。例えば、Ｎ個ものカメラアレイ平面が存在してもよく（ここで、Ｎ＋１はカメラの総数に等しく、「１」は中央のカメラからのものである）、仮想カメラを被写体と結ぶ光線と、これらの平面のうちの１つまたは複数との交点を決定することができる。次いで、この交点を使用して、上述した技法のいずれかを使用して、各カメラから取得した画像に関連付けられた重みを決定することができる。

【００７２】

より一般的には、任意のカメラ配置を使用することができ、ｘ－ｙ平面上に投影されたそれらのカメラの原点を使用して、その平面の三角形の不規則ネットワークカバレッジを決定することができる。次いで、仮想カメラとの交点「ｘ」を含む三角形を形成する少なくとも３つのカメラからの重みの寄与を使用して、最終的な混合画像を生成することができる。

【００７３】

図２に戻ると、第５のステップＳ２０５において、決定された重みに従って、歪み画像が合成され、それにより被写体の画像が仮想カメラの視点から生成される。生成された画像は、ＨＭＤに表示するために出力される立体視画像に対応することができる。所望により、各歪み画像の完全に透明なピクセルは、歪み画像を混合する前に破棄されてもよい。

【００７４】

一部の例では、立体視カメラによってキャプチャされた左右の画像の各セットに対して、一組の歪み画像を生成することができる。左の画像用に生成された歪み画像は、ＨＭＤの左目のディスプレイエレメントに表示するための画像を形成するために、一緒に混合されてもよい。同様に、右の画像用に生成された歪み画像は、ＨＭＤの右目のディスプレイエレメントに表示するための画像を形成するために、一緒に混合されてもよい。

【００７５】

図１１は、上述の方法に従って生成された歪み画像１１０２Ａ、１１０２Ｂ、１１０２Ｃ、１１０２Ｄ、１１０２Ｅの例を示す。図１１の歪み画像の相対的な配置は、対応するソース画像をキャプチャした立体視カメラの相対的な配置に対応する。図１１では、歪み画像はすべて、前述した立体視画像の左の（または右の）画像から生成された一組の歪み画像に対応する。

【００７６】

図１１からわかるように、歪み画像には様々なレベルの孔（すなわち、欠落した色データ）がある。所与の歪み画像に関連付けられた孔のレベルは、一般に、関連付けられたソース画像が、被写体が表現される仮想カメラの視点にどれだけ厳密に一致しているかに依存する。図１１では、仮想カメラの視点は、被写体の左側に向けられているものに対応し、したがって、対応する位置のソースカメラから取得した歪み画像（例えば、画像１１０２Ａ、１１０２Ｃ、１１０２Ｄ）は、孔が少ない。どんな場合でも、歪み画像が合成されると、合成画像の孔がより少なくなる。

【００７７】

決定された重みに従って、歪み画像を混合することによって生成された画像の例が図１１の最終画像１１０４として示されている。図１１からわかるように、最終画像１１０４には、歪み画像のいずれもが塗りつぶすことのできなかった被写体の領域に対応して、依然としていくつかの間隙がある。これらの領域は、例えば、単純なしきい値ベースのフィルタを使用して塗りつぶすことができる。したがって、実際に表示される最終画像は、少なくともいくつかの孔塗りつぶしが実行された混合歪み画像に対応することができる。

【００７８】

一部の例では、テクスチャ空間が制限されており、（立体ペアそれぞれについて）立体ペアからの１つの画像のみが最終的な混合において使用されることがあってもよい。これらの画像が再投影されるとき、特に、この妥協が外側のカメラ（例えば、すべての左側またはすべての右側のカメラ）にのみ使用される場合、最終画像の品質は、一般に十分である。これは、例えば、最終画像が立体視画像としてではなく、単一画像として表示される場合に適切である場合がある。代替としてまたは加えて、このようにして生成された単一画像から立体視画像が引き続き生成されることがあってもよい。

【００７９】

最終画像を生成する際に立体視画像からの一組の画像のみが使用される場合、鏡面ハイライトおよび反射が依然として正確に表現されることが多いことが本発明者らによって見出された。これは、反射の深さが反射面ではなく深度マップに記憶されているためである。しかしながら、結果は、一般に、各立体視ペアの両方の画像が最終画像の生成に使用される場合に改善される。

【００８０】

被写体の最終画像が生成されると、この画像が没入型ビデオの一部として表示されてもよい。被写体が表示される環境は、被写体自体とは別に生成される場合がある。例えば、被写体は、コンピュータで生成された環境の前に重畳されてもよい。場合によっては、例えば、スタジオまたはセットなどの環境のＬＩＤＡＲスキャンが実行され、これを使用してその環境の再構成を生成することがあってもよい。これは、例えば、没入型ビデオが現実の環境で行われるライブ演奏またはゲームに対応している場合に適切な場合がある。次いで、被写体は、被写体の表現が上述の方法に従って調整される、この再構成された環境内に適切に配置されてもよい。

【００８１】

上述の実施形態は、ＨＭＤに表示するための画像の生成に関連して説明されたが、任意のタイプの表示スクリーンに表示するために、同じ技法を等しく使用して、オブジェクトを３Ｄでレンダリングすることができることが理解されよう。

【００８２】

コンピュータ・システムに前述した方法を実行させるように適合されたコンピュータ実行可能命令を有するコンピュータ可読媒体も提供することができる。一部の例では、レン

ダリング方法は、例えば、ゲームコンソールで実行されてもよく、最終画像がゲームコンソールと通信しているＨＭＤのディスプレイに表示されてもよい。代替としてまたは加えて、レンダリングの少なくとも一部は、例えば、サーバで実行されてもよく、最終画像が、例えばインターネット接続を介して、またはサーバと通信しているゲームコンソールを介して、サーバと通信しているＨＭＤのディスプレイに提供されてもよい。代替としてまたは加えて、ＨＭＤは、適切な処理手段を備えている場合、レンダリング方法自体の少なくとも一部を実行することができる。

【００８３】

一部の例では、レンダリングが、サーバ、ゲームコンソール、またはＨＭＤのうちの２つ以上で共有され、最終画像が、必要に応じてゲームコンソールからＨＭＤに中継されることがあってもよい。最終的に、サーバ、ゲームコンソール、およびＨＭＤの任意の組合せを使用して、それぞれのハードウェア能力およびレンダリングされる被写体の性質に応じて、レンダリングを実行することができる。

【００８４】

ここで、上述した方法を実施するための表示システムを図１２に関連して説明する。

【００８５】

表示システム１２００は、オブジェクトの画像を視聴者に表示するように動作可能なディスプレイエレメント１２０２を備える。オブジェクトは、静的（例えば、ぬいぐるみの玩具）または動的（例えば、演奏者）なオブジェクトである。オブジェクトは、前述した被写体に対応する。ディスプレイエレメントは、図１に関連して前述したように、ＨＭＤのディスプレイに対応することができる。

【００８６】

本システムは、視聴者の頭部の姿勢を検出するように動作可能な姿勢検出器１２０４をさらに備える。姿勢検出器１２０４は、例えば、ＨＭＤに組み込まれた加速度計および／またはジャイロスコープを備えることができる。代替としてまたは加えて、姿勢検出器１２０４は、１つまたは複数のカメラを備えてもよく、視聴者の頭部の姿勢は、１つまたは複数のカメラによってキャプチャされた画像内の視聴者の頭部の姿勢に基づいて検出されてもよい。一部の例では、ＨＭＤには複数の光源が取り付けられており、視聴者の頭部の姿勢は、（例えば、コンスタレーション追跡を使用して）カメラによってキャプチャされた画像にこれらの光源がどのように現れるかに基づいて決定されてもよい。

【００８７】

前述したように、視聴者の頭部の姿勢は、異なるそれぞれの視点からオブジェクトを見るために使用される仮想カメラの姿勢に対応することができる。したがって、姿勢検出器１２０４により、仮想カメラの姿勢を決定することが可能になる。視聴者が没入型ビデオを見ているとき、仮想カメラ（および視聴者の頭部）の姿勢は、オブジェクトのソース画像がキャプチャされたソースカメラの姿勢とは異なることが一般に予想される。

【００８８】

本システムは、オブジェクトの複数のソース画像と、各ソース画像をキャプチャしたカメラの姿勢を示すカメラ姿勢データと、を取得するように動作可能な画像プロセッサ１２０６をさらに備える。前述したように、各ソース画像は、複数のカメラのアレイ（例えば、立体視カメラ）によって、または単一のカメラをオブジェクトに対して異なる位置に移動させることによってキャプチャすることができる。各ソース画像は、カメラ姿勢データが対応する、対応するソースカメラに関連付けられている。ソース画像は、図３～図１１に関連して前述したやり方のいずれかでキャプチャされたものであってもよい。

【００８９】

画像プロセッサ１２０６は、姿勢検出器１２０４からの入力を受信し、それに応答して、ソース画像に関連付けられた１つまたは複数のカメラの姿勢と、オブジェクトに対する視聴者の頭部の現在の姿勢との差に基づいてソース画像の少なくともいくつかを歪ませるように構成されている。オブジェクトに対する視聴者の頭部の姿勢は、オブジェクトの再構成に対する仮想カメラの姿勢に対応する。

【００９０】
具体的には、画像プロセッサ１２０６は、その画像に関連付けられたソースカメラの姿勢と、オブジェクトが視認可能な仮想カメラの姿勢との差に基づいて各ソース画像を歪ませるように構成されている。ソース画像の歪みは、図３～図１１に関連して上述した方法のいずれかに従って実行することができる。
【００９１】
本システムは、その画像に関連付けられたカメラの姿勢と、オブジェクト（またはその再構成）に対する視聴者の頭部の姿勢との類似性に基づいて、各歪み画像に関連付けられた重みを決定するように構成された画像合成器１２０８をさらに備える。画像合成器１２０８は、関連付けられた重みに従って歪み画像を一緒に合成して、ディスプレイエレメントに表示するための被写体の画像を形成するようにさらに構成されている。すなわち、画像合成器１２０８は、歪み画像を一緒に混合して、視聴者の視点（これは、ひいては、オブジェクトを視認可能にする仮想カメラの視点に対応する）からのオブジェクトの最終画像を生成するように構成されてもよい。最終的な混合画像は、視聴者に視認可能な現実の環境または仮想環境に対応し、これと重畳されてもよい。
【００９２】
一部の例では、画像プロセッサ１２０６は、各ソース画像に関連付けられた深度データを取得し、ソース画像それぞれについて、ソース画像およびその画像に関連付けられた深度データに基づいて被写体の多角形メッシュを生成するようにさらに構成されている。深度データは、前述したやり方のいずれかで取得することができる。画像プロセッサ１２０６は、ソース画像に対して生成されたメッシュの少なくとも一部を、視聴者の頭部が向いている方向（視聴者の視線方向に対応する）に対して垂直になるように再投影することによって、各ソース画像から歪み画像を生成するように構成されている。メッシュのこの再投影は、図８Ａおよび図８Ｂに関連して前述した。
【００９３】
画像プロセッサ１２０６は、再投影されたメッシュそれぞれについて、そのメッシュに適用するためのテクスチャを生成するように構成されてもよい。画像プロセッサ１２０６は、ソース画像の部分を、そのソース画像から生成された再投影されたメッシュの対応する部分に視差マッピングするように動作可能な視差シェーダを備えることができる。前述したように、この視差マッピング処理は、図９に示されている。
【００９４】
視聴者の頭部の姿勢は、オブジェクトを見るための仮想カメラの姿勢に対応し、ソース画像をキャプチャした１つまたは複数のカメラの姿勢は、カメラアレイ平面に対応する。画像合成器１２０８は、仮想カメラをオブジェクトと結ぶベクトルと、カメラアレイ平面との交点を決定することによって、歪み画像それぞれについての重みを決定するように構成されてもよい。画像合成器１２０８は、歪み画像それぞれについて、その画像に関連付けられたカメラの距離およびカメラアレイ平面上の交点に基づいて、重みを決定するように構成されてもよい。この処理は、前述した図１０Ａおよび図１０Ｂに示されている。
【００９５】
一部の例では、画像プロセッサ１２０６は、ディスプレイエレメント１２０２を備えるヘッドマウントディスプレイ（ＨＭＤ）と通信するビデオゲームコンソールに実装されてもよい。したがって、ビデオゲームコンソールは、歪み画像を生成して混合し、その結果をＨＭＤに送信して表示するように構成されてもよい。一部の例では、歪み画像が、例えばゲームコンソールで生成され、合成がＨＭＤで実行されることがあってもよい。例えば、ＰＳ５（商標）などの次世代コンソールは、上述した方法に従ってソース画像のそれぞれを歪ませるのに適している可能性がある。
【００９６】
代替としてまたは加えて、画像プロセッサ１２０６は、上述したようなビデオゲームコンソールと同様のやり方で、ＨＭＤ自体に実装されてもよい。
【００９７】

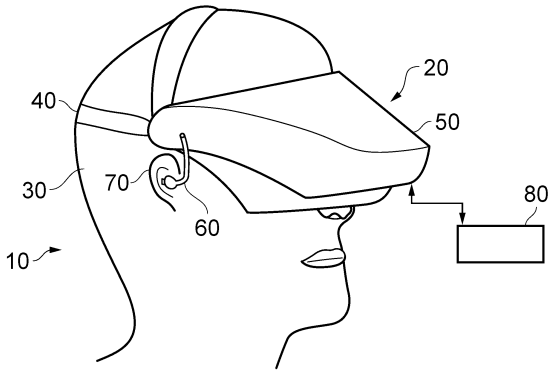
代替としてまたは加えて、画像プロセッサ１２０６は、例えば、ビデオゲームコンソールおよび／またはＨＭＤが通信しているサーバに実装されてもよい。したがって、歪み画像は、サーバで生成され、サーバから（場合によっては、ゲームコンソールを介して）ＨＭＤに送信されてもよい。理解されるように、視聴者は、自分の頭部のいくつかの小さな動きを行うことがあるため、歪み画像がサーバではなくゲームコンソールで生成されることが好ましい場合がある。これにより、オブジェクトが対応する視点からレンダリングされ得る前に、頭部姿勢データをサーバに送信しなければならないことによってもたらされる遅延が制限される。

【００９８】

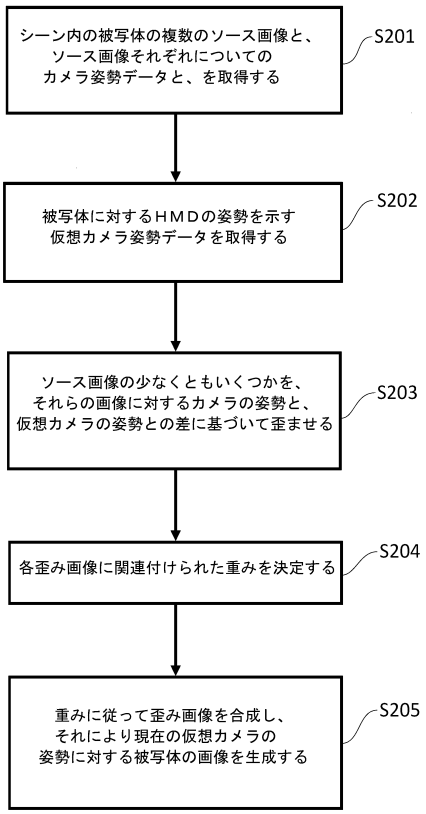
本明細書に記載の方法は、ソフトウェア命令によって、または専用ハードウェアの包含または置換によって適用可能なように適切に適合された従来のハードウェア上で実行することができることが理解されよう。したがって、従来の等価な機器の既存の部分への必要な適合は、フロッピー（登録商標）ディスク、光ディスク、ハードディスク、ＰＲＯＭ、ＲＡＭ、フラッシュメモリ、またはこれらもしくは他の記憶媒体の任意の組合せなどの非一時的な機械可読媒体に記憶されたプロセッサ実施可能命令を含むコンピュータ・プログラム製品の形態で実施されてもよく、あるいはＡＳＩＣ（特定用途向け集積回路）またはＦＰＧＡ（フィールド・プログラマブル・ゲート・アレイ）または従来の等価な機器に適合する際に使用するのに適した他の構成可能な回路としてハードウェアで実現されてもよい。別個に、そのようなコンピュータ・プログラムは、イーサネット（登録商標）、ワイヤレス・ネットワーク、インターネット、またはこれらもしくは他のネットワークの任意の組合せなどのネットワーク上のデータ信号を介して送信されてもよい。

【図面】

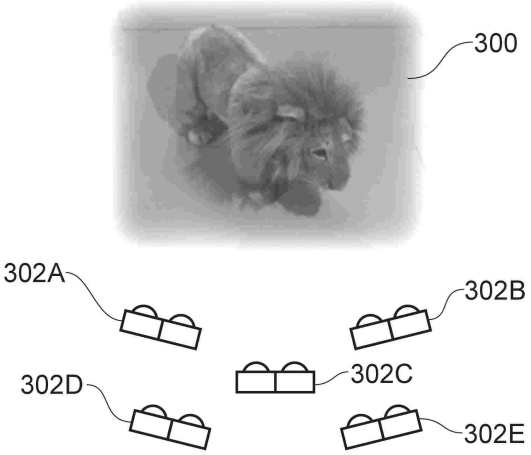
【図１】


20
40
50
30
70
80
60
10

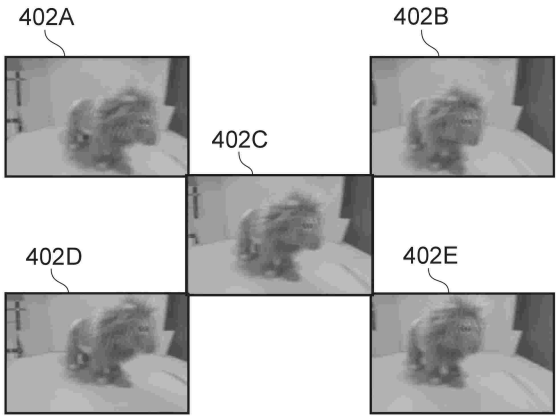
【図２】


シーン内の被写体の複数のソース画像と、
ソース画像それぞれについての
カメラ姿勢データと、を取得する
S201
被写体に対するＨＭＤの姿勢を示す
仮想カメラ姿勢データを取得する
S202
ソース画像の少なくともいくつかを、
それらの画像に対するカメラの姿勢と、
仮想カメラの姿勢との差に基づいて歪ませる
S203
各歪み画像に関連付けられた重みを決定する
S204
重みに従って歪み画像を合成し、
それにより現在の仮想カメラの
姿勢に対する被写体の画像を生成する
S205

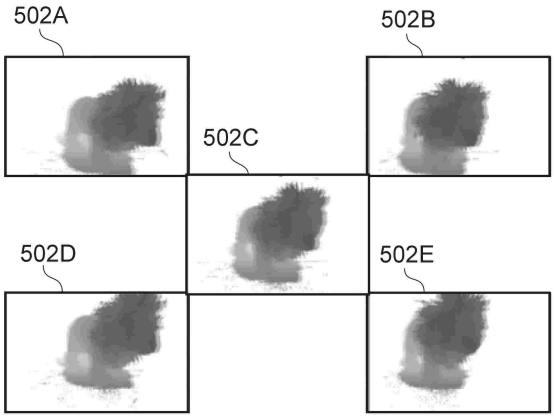
【図３】


300
302A
302B
302C
302D
302E

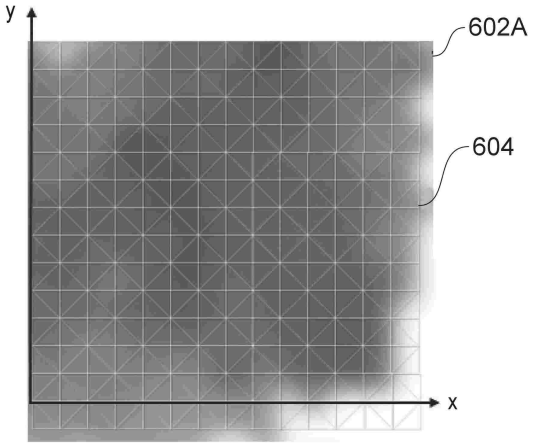
【図４】


402A
402B
402C
402D
402E

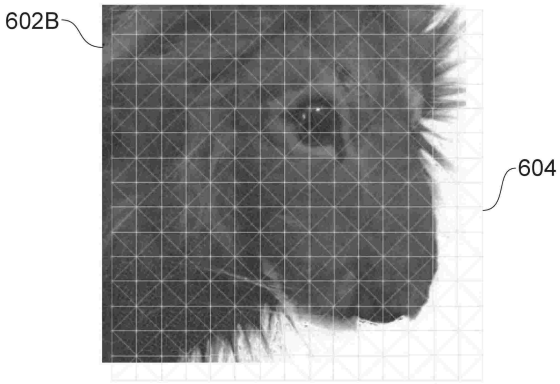
【図５】


502A
502B
502C
502D
502E

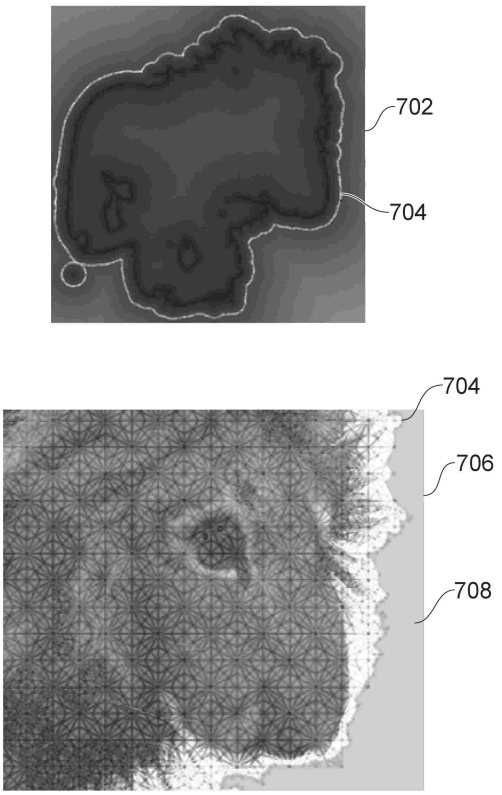
【図６Ａ】


y
602A
604
x

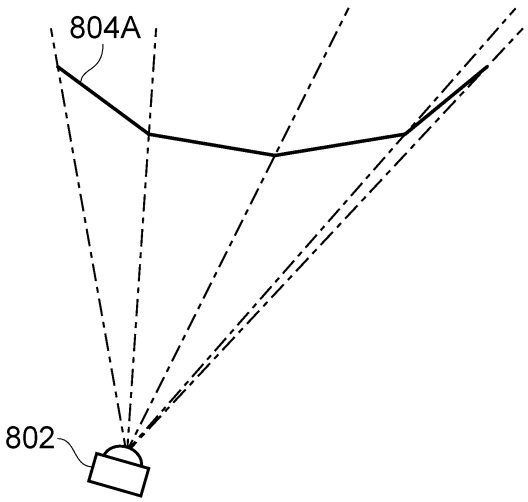
【図６Ｂ】


602B
604

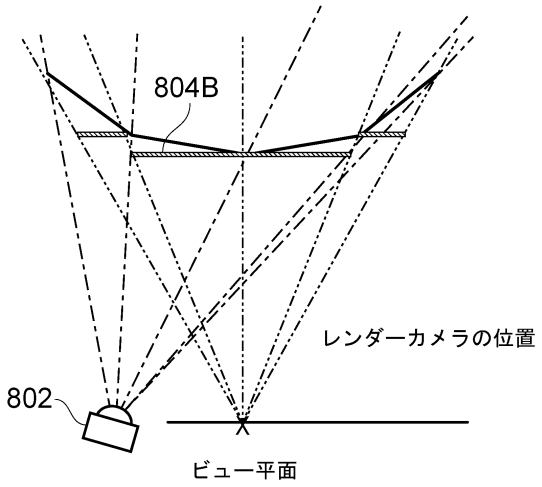
【図７】


702
704
704
706
708

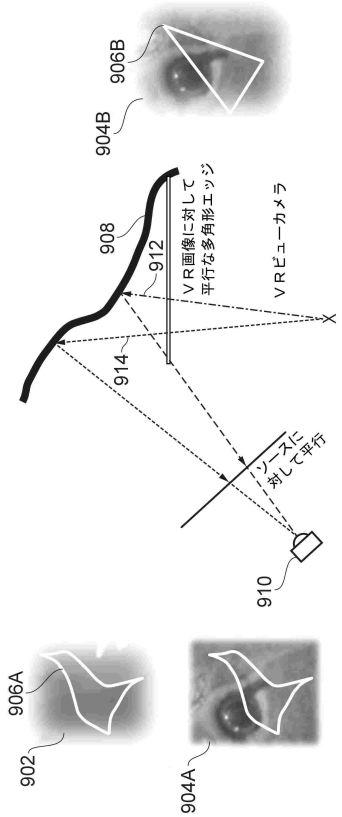
【図８Ａ】



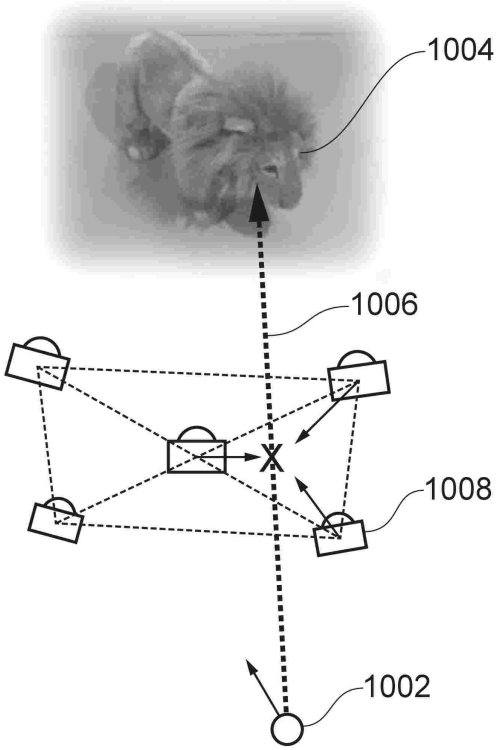
【図８Ｂ】



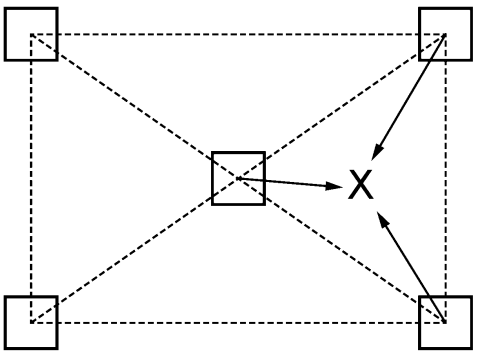
【図９】



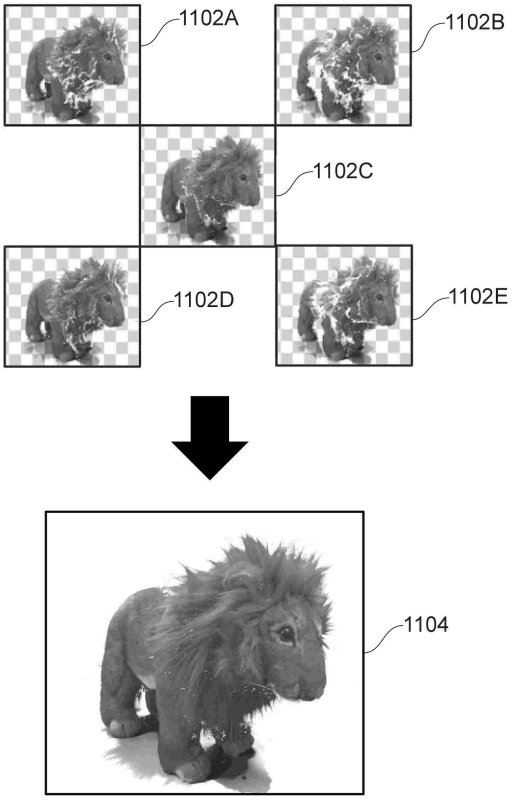
【図１０Ａ】



【図１０Ｂ】



【図１１】



【図１２】

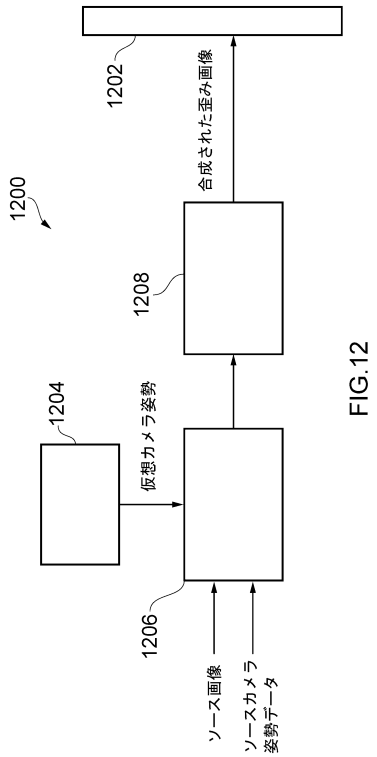


FIG.12

フロントページの続き

ロンドン グレート マールボロ ストリート １０、ソニー インタラクティブ エンタテインメント ヨーロッパ リミテッド内
(72)発明者 イアン ヘンリー ビッカースタッフ
イギリス国 ダブリュ１エフ ７エルピー ロンドン グレート マールボロ ストリート １０、ソニー インタラクティブ エンタテインメント ヨーロッパ リミテッド内

審査官 中田 剛史

(56)参考文献
特開２００９－２１１３３５（ＪＰ，Ａ）
特開２０１５－１８７７９７（ＪＰ，Ａ）
特開２０１５－２３２７８３（ＪＰ，Ａ）
国際公開第２０１２／０４２９９８（WO，Ａ１）
特表２０２０－５２４８５１（ＪＰ，Ａ）
特開２０１４－０５６４６６（ＪＰ，Ａ）

(58)調査した分野 (Int.Cl.，ＤＢ名)
Ｇ０６Ｔ １９／００