

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第3996933号  
(P3996933)

(45) 発行日 平成19年10月24日(2007.10.24)

(24) 登録日 平成19年8月10日(2007.8.10)

(51) Int.C1.

F 1

<b>G06T</b>	<b>1/00</b>	<b>(2006.01)</b>	<b>G06T</b>	<b>1/00</b>	<b>315</b>
<b>G06T</b>	<b>17/40</b>	<b>(2006.01)</b>	<b>G06T</b>	<b>17/40</b>	<b>F</b>
<b>G09G</b>	<b>5/36</b>	<b>(2006.01)</b>	<b>G09G</b>	<b>5/36</b>	<b>510V</b>
<b>HO4N</b>	<b>13/00</b>	<b>(2006.01)</b>	<b>HO4N</b>	<b>13/00</b>	

請求項の数 13 外国語出願 (全 20 頁)

(21) 出願番号 特願2005-179123 (P2005-179123)  
 (22) 出願日 平成17年6月20日 (2005.6.20)  
 (65) 公開番号 特開2006-24206 (P2006-24206A)  
 (43) 公開日 平成18年1月26日 (2006.1.26)  
 審査請求日 平成18年11月10日 (2006.11.10)  
 (31) 優先権主張番号 10/879,235  
 (32) 優先日 平成16年6月28日 (2004.6.28)  
 (33) 優先権主張国 米国(US)

早期審査対象出願

(73) 特許権者 500046438  
 マイクロソフト コーポレーション  
 アメリカ合衆国 ワシントン州 98052  
 2-6399 レッドモンド ワン マイ  
 クロソフト ウェイ  
 (74) 代理人 100077481  
 弁理士 谷 義一  
 (74) 代理人 100088915  
 弁理士 阿部 和夫  
 (72) 発明者 チャールズ ジトニック ザ サード  
 アメリカ合衆国 98052 ワシントン  
 州 レッドモンド ワン マイクロソフト  
 ウェイ マイクロソフト コーポレーション内

最終頁に続く

(54) 【発明の名称】シーンの2レイヤ3D表現を生成するシステムおよび処理

## (57) 【特許請求の範囲】

## 【請求項1】

イメージおよび前記イメージの視差マップからデジタルイメージまたはデジタル化された前記イメージの2レイヤ表現を生成するコンピュータ実施される処理であって、

前記イメージの視差マップを使用して、前記イメージ内の深さ不連続の位置を識別する処理アクションと、

前記深さ不連続を囲む指定されたサイズの領域内で見つかる前記イメージのピクセルを識別し、これらのピクセルを、前記イメージの深さ不連続領域内にあるものとして指定する処理アクションと、

前記深さ不連続領域内のピクセルごとに、フォアグラウンドカラー、バックグラウンドカラー、およびアルファ値を推定する処理アクションと、

前記深さ不連続領域内のピクセルごとに、フォアグラウンド視差値およびバックグラウンド視差値を推定する処理アクションと、

前記深さ不連続領域のピクセルに対応する各めいめいの位置にあるピクセルを含む前記イメージの境界レイヤを確立する処理アクションであって、各境界レイヤピクセルは、前記深さ不連続領域の対応するピクセルに関連する前記フォアグラウンドカラー、前記フォアグラウンド視差、および前記アルファ値を割り当てる、処理アクションと、

前記イメージのメインレイヤを確立する処理アクションであって、前記メインレイヤは

、  
前記深さ不連続領域のピクセルに対応する各めいめいの位置のピクセルであって、該 20

ピクセルのそれぞれは、前記深さ不連続領域の対応するピクセルに関連する前記バックグラウンドカラーおよび前記バックグラウンド視差値を割り当てられる、ピクセルと、

前記深さ不連続領域のピクセルに対応しない各位置のピクセルであって、該ピクセルのそれぞれは、前記イメージの対応するピクセルに関連する前記カラーおよび前記視差値を割り当てられる、ピクセルと

を含む、処理アクションと

をコンピュータを使用して実行することを備えることを特徴とするコンピュータ実施される処理。

#### 【請求項 2】

前記イメージ内の深さ不連続の位置を識別する前記処理アクションは、視差レベルの指定された数を超える、隣接ピクセルの間の視差値の差を示す任意の位置を、深さ不連続として識別するアクションを含むことを特徴とする請求項 1 に記載の処理。 10

#### 【請求項 3】

視差レベルの前記指定された数が、4 であることを特徴とする請求項 2 に記載の処理。

#### 【請求項 4】

前記深さ不連続を囲む前記指定されたサイズの領域は、識別された深さ不連続位置から各方向に 3 ピクセル拡張した領域として定義されることを特徴とする請求項 1 に記載の処理。

#### 【請求項 5】

前記深さ不連続領域内のピクセルごとにフォアグラウンドカラー、バックグラウンドカラー、およびアルファ値を推定する前記処理アクションは、前記カラーおよびアルファ値を推定するのにマッティング技法を使用するアクションを含むことを特徴とする請求項 1 に記載の処理。 20

#### 【請求項 6】

前記深さ不連続領域内のピクセルごとにフォアグラウンドおよびバックグラウンド視差値を推定する前記処理アクションは、前記フォアグラウンド視差値および前記バックグラウンド視差値を推定するのに前記イメージのフォアグラウンド部分およびバックグラウンド部分の近傍視差のアルファ加重平均を使用するアクションを含むことを特徴とする請求項 1 に記載の処理。

#### 【請求項 7】

前記フォアグラウンド視差値および前記バックグラウンド視差値を推定するのに前記イメージのフォアグラウンド部分およびバックグラウンド部分の近傍視差のアルファ加重平均を使用する処理アクションは、 30

検討中のピクセルに隣接する前記イメージのフォアグラウンド部分内の指定されたサイズのウィンドウ内の各ピクセルにそのアルファ値をそれぞれ乗じ、結果の積の平均をとることによって、各深さ不連続領域内の各ピクセルの前記フォアグラウンド視差値を計算するアクションと、

検討中のピクセルに隣接する前記イメージのバックグラウンド部分内の指定されたサイズのウィンドウ内の各ピクセルに 1 からそのアルファ値を減じたものをそれぞれ乗じ、結果の積の平均をとることによって、各深さ不連続領域内の各ピクセルの前記バックグラウンド視差値を計算するアクションと 40

を含むことを特徴とする請求項 6 に記載の処理。

#### 【請求項 8】

前記境界レイヤピクセルに対応する領域を指定された量だけ膨張させ、各追加されたピクセルに前記メインレイヤ内の対応するピクセルと同一のカラーおよび視差値と、1 のアルファ値とを割り当てる処理アクションをさらに含むことを特徴とする請求項 1 に記載の処理。

#### 【請求項 9】

前記指定された膨張量は、1 ピクセルであることを特徴とする請求項 8 に記載の処理。

#### 【請求項 10】

イメージおよび該イメージのピクセル深さマップから、ディジタルイメージまたはディジタル化されたイメージの2レイヤ表現を生成するシステムであって、

汎用コンピューティングデバイスと、

前記コンピューティングデバイスによって実行可能なプログラムモジュールを含むコンピュータプログラムと

を備え、前記コンピューティングデバイスは、前記コンピュータプログラムの前記プログラムモジュールによって、

前記イメージの深さマップを使用して、前記イメージ内の深さ不連続の位置を識別し、

前記深さ不連続を囲む指定されたサイズの領域内で見つかる前記イメージのピクセルを識別し、これらのピクセルを、前記イメージの深さ不連続領域内にあるものとして指定し、

前記深さ不連続領域内のピクセルごとに、フォアグラウンドカラー、バックグラウンドカラー、およびアルファ値を推定し、

前記深さ不連続領域内のピクセルごとに、フォアグラウンド深さ値およびバックグラウンド深さ値を推定し、

前記深さ不連続領域の対応する位置のピクセルに関連する前記バックグラウンドカラーおよび前記バックグラウンド深さを示すピクセルと前記深さ不連続領域内にない前記イメージの対応する位置のピクセルに関連する前記カラーおよび前記深さを示すピクセルとを含む前記イメージのメインレイヤ、ならびに前記深さ不連続領域の対応する位置のピクセルに関連する前記フォアグラウンドカラー、前記フォアグラウンド深さ、および前記アルファ値を示すピクセルを含む前記イメージの境界レイヤを生成する

ように指示されることを特徴とするシステム。

#### 【請求項 1 1】

前記イメージの深さ不連続の前記位置を識別する前記プログラムモジュールは、指定された量を超える、隣接ピクセルの間の深さ値の差を示すすべての位置を、深さ不連続として識別するサブモジュールを含むことを特徴とする請求項 1 0 に記載のシステム。

#### 【請求項 1 2】

イメージおよび前記イメージのピクセル視差マップからディジタルイメージまたはディジタル化された前記イメージの2レイヤ表現を生成する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であって、前記コンピュータ実行可能命令は、

前記イメージの視差マップを使用して、前記イメージ内の深さ不連続の位置を識別することと、

前記識別された深さ不連続を囲む指定されたサイズの深さ不連続領域を確立することと、

前記深さ不連続領域内のピクセルごとに、フォアグラウンドカラー、バックグラウンドカラー、およびアルファ値を推定することと、

前記深さ不連続領域内のピクセルごとに、フォアグラウンド視差値およびバックグラウンド視差値を推定することと、

前記深さ不連続領域の対応する位置のピクセルに関連する前記バックグラウンドカラーおよび前記バックグラウンド視差を示すピクセルと前記深さ不連続領域内にない前記イメージの対応する位置のピクセルに関連する前記カラーおよび前記視差を示すピクセルとを含む前記イメージのメインレイヤ、ならびに前記深さ不連続領域の対応する位置のピクセルに関連する前記フォアグラウンドカラー、前記フォアグラウンド視差、および前記アルファ値を示すピクセルを含む前記イメージの境界レイヤを生成することと

を含むことを特徴とするコンピュータ可読媒体。

#### 【請求項 1 3】

ディジタルイメージまたはディジタル化されたイメージの2レイヤ表現を生成する装置であって、前記2レイヤ表現とは、メインレイヤと境界レイヤから構成され、

前記メインレイヤは、深さ不連続領域の対応する位置のピクセルに関連するバックグラ

10

20

30

40

50

ウンドカラーおよびバックグラウンド視差を示すピクセルと、前記深さ不連続領域内にない前記イメージの対応する位置のピクセルに関するカラーおよび視差を示すピクセルとを含み、

前記境界レイヤは、前記深さ不連続領域の対応する位置のピクセルに関するカラー、フォアグラウンドカラー、フォアグラウンド視差、およびアルファ値を示すピクセルを含み、

前記深さ不連続領域が、前記イメージの視差マップを使用して前記イメージ内で見つかる深さ不連続を囲む指定されたサイズの領域に対応するものである

ことを特徴とする装置。

#### 【発明の詳細な説明】

##### 【技術分野】

10

##### 【0001】

本発明は、デジタルイメージまたはデジタル化されたイメージのレイヤ表現に関し、具体的には、シーンの2レイヤ3D表現を生成するシステムおよび処理に関する。

##### 【背景技術】

##### 【0002】

最近数年間、TVコマーシャルおよび長編映画を観る人は、時間を止め、カメラ視点を移動するという幻想を作成するのに使用される「静止画」効果を見てきた。最も初期のコマーシャルは、時間の凍結されたスライスを通じて移動するという幻想を与るためにレールに沿って配列された異なる静止画カメラの間で素早くジャンプする、フィルムベースのシステムを使用して作成された。

20

##### 【0003】

はじめて現れた時に、その効果は、新鮮で華々しく見えたが、ほどなく、多数の作品でエミュレートされ、そのうちで最も有名なものが、多分、「マトリックス」という題名の映画で見られる「バレットタイム（bullet time）」効果である。残念ながら、この効果は、1回限りの事前に計画された仕事である。視点の軌跡が、事前に計画され、多数の人月（man hours）が、所望の補間されたビューを作ることに費やされる。より新しいシステムは、ビデオカメラアレイに基づくが、それでも、ソフトウェアビュー補間を避けるために多数のカメラを備えることを当てにする。

##### 【0004】

したがって、既存のシステムは、ユーザが、動的なイメージベースのシーンを見ながら、任意の所望の視点に対話式に変更することを受け入れない。過去のイメージベースレンダリング（IBR）に関する作業のほとんどが、静的なシーンのレンダリングを伴い、最もよく知られた技法の2つが、ライトフィールドレンダリング（Light Field Rendering）[11]およびルミグラフ（Lumigraph）[7]である。高品質レンダリングでのこれらの成功は、多数のサンプリングされたイメージの使用に由来し、この分野での多数の仕事を示唆した。この画期的な仕事の刺激的な潜在的拡張の1つに、ビデオを見ながら視点を対話式に制御することが含まれる。ユーザがビデオの視点を対話式に制御する能力は、見る経験をかなり高め、新しい視点のインスタント再生、ドラマでの視点の変更、および意のままの「静止画」視覚効果の作成などの別個の応用例が可能になる。

30

##### 【0005】

しかし、多数のカメラを同期させることならびにイメージの獲得および記憶の難しさ（およびコスト）のゆえに、動的シーンへのIBRの拡張は、自明ではない。複数の視点からの動的シーンのキャプチャ、表現、およびレンダリングでの克服すべき大きな障害があるだけではなく、これを対話式に行うことを可能にすることが、さらなる大きな複雑化をもたらす。現在まで、この目標を達成する試みは、非常に不満足なものである。

##### 【0006】

対話型視点ビデオシステムのビデオベースレンダリング態様について、動的シーンのキャプチャおよびレンダリングの最も初期の試みの1つが、カナデ（Kanade）他のVirtualized Realityシステム[10]であり、これには、5mのジオ

40

50

デシックドームの周囲に配置された 51 台のカメラが用いられた。各カメラの解像度は、 $512 \times 512$  であり、キャプチャレートは、30 f p s である。これらのカメラは、シーンフロー式 [17] に基づくボクセルカラーリング [14] を使用して、各タイムフレームにグローバルサーフェス表現を抽出する。残念ながら、その結果は、低い解像度、マッチングエラー、およびオブジェクト境界の不適切なハンドリングのゆえに、非現実的に見える。

#### 【0007】

カレンザ他 [3] は、3D の人間の動きをキャプチャするために、部屋の中心を見る、部屋の回りに分散した 7 台の同期化されたカメラを使用した。各カメラは、CIF 解像度 ( $320 \times 240$ ) であり、15 f p s でキャプチャする。彼らは、修道院長として 3D 人間モデルを使用して、各タイムフレームに 3D 形状を計算した。10

#### 【0008】

ヤン (Yang) 他 [18] は、動的シーンをキャプチャするためにカメラ（各  $320 \times 240$ ）の  $8 \times 8$  グリッドを設計した。データを記憶し、レンダリングするのではなく、これらは、所望の仮想ビューを構成するのに必要な光線だけを送信する。このシステムでは、カメラがゲンロック (genlock) されず、6 台の PC にまたがる内部クロックに依存する。カメラキャプチャレートは、15 f p s であり、対話型ビューアイinguレートは、18 f p s である。

#### 【0009】

前述のシステムに共通しているのが、シーンジオメトリが、未知であるか、おおまかにのみ既知であるのいずれかであることに部分的に起因して、現実的なレンダリングに大量のイメージが必要であることである。ジオメトリが正確に既知の場合には、イメージの要件を大幅に減らすことが可能である [7]。シーンジオメトリを抽出する 1 つの実用的な形が、立体視を介するものであり、多数の立体視アルゴリズムが、静的シーンについて提案してきた [13]。しかし、動的シーンに立体視技法を使用する試みはほとんど行われてこなかった。Virtualized Reality 作業 [10] の一部として、ベデュラ (Vedula) 他 [17] は、2D オプティカルフローおよび 3D シーン形状を使用して 3D 動きを抽出する（たとえば、経時的なシーン形状の間の対応）アルゴリズムを提案した。この手法では、ボクセルカラーリング [14] に似た投票方式 (voting scheme) が使用され、そこで使用される指標 (measure) は、仮定されたボクセル位置が 3D フロー式にどれほどよくあてはまるかである。2030

#### 【0010】

ザングおよびカムバメツ (Zhang and Kambhamettu) [19] は、3 シーンフローおよび構造をフレームワークに統合した。彼らの 3D アフィン動きモデルは、空間正則化 (spatial regularization) と共にローカルに使用され、不連続性が、カラーセグメント化を使用して保存される。タオ (Tao) 他 [16] は、シーンが区分的に平面であると仮定した。彼らは、動的深さマップ推定を制約するために、平面パッチごとに一定の速度を仮定した。

#### 【0011】

より野心的な努力で、カーセロニおよびクツラコス (Carceroni and Kutulakos) [2] は、既知の照明位置を伴う非剛体運動の下で区分的に連続なジオメトリおよび反射率 (フォンモデル) を正常な状態に戻した。彼らは、空間をサーフェス要素（「サーフェル (surface)」）に離散化し、位置、方位、および反射率パラメータに対する検索を実行して、観察されたイメージとの一致を最大にする。40

#### 【0012】

普通のローカルウィンドウマッチングに対する興味深い別の手法で、ザング他 [20] は、空間と時間にまたがるマッチングウィンドウを使用した。この方法の利益は、経時的な明るさの定常性に対する依存が少ないことである。

#### 【0013】

アクティブレンジファインディング (active range finding) 技法50

も、移動するシーンに適用されてきた。ホールホルトおよびラシンキウイツ ( Hall - Holt and Russinkiewicz ) [ 8 ] は、経時的に変化する射影された境界コーディングされたストライプパターンを使用した。市場に、イスラエルの 3 D V

S y s t e m s 社が製造する Z C a m ( 商標 ) と称する市販システムもあり、これは、放送ビデオカメラと一緒に使用されるレンジセンシングビデオカメラアドオンである。しかし、これは、高価なシステムであり、単一の視点深さだけを提供し、これによって、複数視点ビデオに適しなくなっている。

#### 【 0 0 1 4 】

しかし、立体視およびイメージベースレンダリングのすべての進展にかかわらず、動的シーンの高品質高解像度ビューをレンダリングすることは、まだ非常に困難である。ライ 10  
トフィールドレンダリング論文 [ 1 1 ] で提案されているように、1つの手法は、入力カメラおよび仮想カメラの相対位置だけに基づいて単純に光線を再サンプリングすることである。ルミグラフ [ 7 ] および後続の研究で実証されているように、シーンジオメトリに 3 D インポスター ( i m p o s t o r ) またはプロキシを使用することによって、補間されたビューの品質を大きく高めることができる。もう1つの手法は、単純なテクスチャマッピングされた 3 D モデル [ 1 0 ] を作成することであるが、これは、一般に、複数の基準ビューの使用より劣る結果を作る。もう1つの手法が、3 D プロキシを必要とする、ジオメトリによって補助されたイメージベースレンダリング手法を使用することである。1つの可能性は、 L u m i g r a p h a n d U n s t r u c t u r e d L u m i g r a p h 論文 [ 1 ] のように単一のグローバル多面体モデルを使用することである。もう1つの可能性は、 L a y e r e d D e p t h I m a g e s [ 1 5 ] のようにピクセルごとの深さ、オフセット深さマップ F a c a d e [ 5 ] のように、または深さ付きスプライト [ 1 5 ] を使用することである。一般に、基準ビュー [ 1 2 、 6 、 9 ] ごとに異なるローカル幾何プロキシを使用することによって、より高い品質の結果が作られる。

#### 【 0 0 1 5 】

しかし、複数の深さマップさえも、新しいビューを生成する時にレンダリングアーティファクトすなわち、フォアグラウンドからバックグラウンドへの推移の急激な性質に起因するエイリアス ( ジャギー ) と、新規の背景またはオブジェクトの上に合成する時に可視になる、混合されたピクセルに起因する汚染されたカラーを示す。

#### 【 0 0 1 6 】

この問題は、本発明で、入力イメージの独自の 2 レイヤ 3 D 表現を介して対処される。この 2 レイヤ 3 D 表現は、対話型視点ビデオシステムで新規のビューをレンダリングすることに関する前述のエイリアス問題を解決するのに使用できるだけではなく、他のコンテキストでも有利に使用できることに留意されたい。一般に、デジタルイメージまたはデジタル化されたイメージは、この 2 レイヤ 3 D 表現を使用して表すことができる。

#### 【 0 0 1 7 】

前の段落および本明細書の残りの説明で、大括弧の対に挟まれた数値指定子によって識別されるさまざまな個々の刊行物に言及することに留意されたい。たとえば、そのような参考文献は、「参考文献 [ 1 ]」または単に「 [ 1 ]」と引用することによって識別することができる。複数の参照は、複数の指定子を含む大括弧の対、たとえば [ 2 、 3 ] によって識別される。各指定子に対応する刊行物を含む参照のリストを、詳細な説明のセクションの末尾に示す。

#### 【 発明の開示 】

#### 【 課題を解決するための手段 】

#### 【 0 0 1 8 】

本発明は、デジタルイメージまたはデジタル化されたイメージの 2 レイヤ表現を生成するシステムおよび方法を対象とする。一般に、2つのレイヤに、イメージの深さ不連続領域 ( d e p t h d i s c o n t i n u i t y a r e a s ) の対応する位置のピクセルに関連するバックグラウンドカラーおよびバックグラウンド視差を示すピクセル、な

10

20

20

30

40

50

らびにこれらの深さ不連続領域で見つからないイメージの対応する位置のピクセルに関するカラーおよび視差を示すピクセルを有するメインレイヤが含まれる。もう一方のレイヤは、深さ不連続領域の対応する位置のピクセルに関するフォアグラウンドカラー、フォアグラウンド視差、およびアルファ値を示すピクセルから構成された境界レイヤである。深さ不連続領域は、イメージ内で見つかる深さ不連続を囲む指定されたサイズの領域に対応する。

#### 【0019】

この2レイヤ表現は、まず、イメージの視差マップを使用して、検討中のイメージ内の深さ不連続の位置を識別することによって生成される。深さ不連続は、隣接ピクセル間の視差値の相違が、指定されたレベルの数を超える位置で発生する。次に、深さ不連続を囲む指定されたサイズの領域内で見つかるイメージのピクセルが、識別される。次に、フォアグラウンドアルファ値と共に、フォアグラウンドカラーおよびバックグラウンドカラーが、マッティング技法を使用して、これらの深さ不連続領域内のピクセルごとに推定される。さらに、フォアグラウンド視差値およびバックグラウンド視差値が、イメージのフォアグラウンド部分とバックグラウンド部分の近傍視差のアルファ加重平均を使用して、深さ不連続領域内のピクセルごとに推定される。次に、深さ不連続領域のピクセルに対応する各位置のピクセルを含むイメージの境界レイヤを確立する。深さ不連続領域の対応するピクセルに関するフォアグラウンドカラー、フォアグラウンド視差、およびアルファ値が、各境界レイヤピクセルに割り当てられる。さらに、イメージのメインレイヤが確立される。このメインレイヤには、深さ不連続領域のピクセルに対応する各位置のピクセル、ならびに深さ不連続領域のピクセルに対応しないイメージの各ピクセル位置のピクセルが含まれる。深さ不連続領域の対応するピクセルに関するバックグラウンドカラーおよびバックグラウンド視差値は、メインレイヤの対応する位置のピクセルのそれぞれに割り当てられ、深さ不連続領域内にないイメージのピクセルに関するカラーおよび視差値は、メインレイヤの対応する位置のピクセルのそれぞれに割り当てられる。フォアグラウンドカラー、フォアグラウンド視差、およびアルファ値が、深さ不連続領域の各ピクセルについて確立されたならば、レイヤからのイメージのレンダリング中に割れ目(crack)が見えないようにするために、普通の膨張(dilation)技法を使用して、所定の量だけこれらの領域のサイズを増やすことができることに留意されたい。

#### 【0020】

上で説明した利益のほかの、本発明の他の利益は、添付図面と共に解釈される時の下の詳細な説明から明白になる。

#### 【0021】

本発明の特定の特徴、態様、および長所は、次の説明、請求項、および添付図面に関してよりよく理解されよう。

#### 【0022】

図4C～4Eのイメージが、反転して示され、その結果、透明／空のピクセルが白く見えていることに留意されたい。

#### 【発明を実施するための最良の形態】

#### 【0023】

本発明の好ましい実施形態の次の説明では、添付図面を参照するが、この添付図面は、本明細書の一部を形成し、本発明を実施できる特定の実施形態を図によって示すものである。本発明の範囲から逸脱することなく、他の実施形態を使用することができ、構造的変更を行うことができることを理解されたい。

#### 【0024】

##### 1.0 コンピューティング環境

本発明の好ましい実施形態の説明を提供する前に、本発明を実施できる適切なコンピューティング環境の簡単な全般的な説明を説明する。図1に、適切なコンピューティングシステム環境100の例を示す。コンピューティングシステム環境100は、適切なコンピューティング環境の1例にすぎず、本発明の使用または機能の範囲に関するいかなる限定

10

20

30

40

50

を示唆することを意図するものではない。コンピューティング環境 100 を、例示的なオペレーティング環境 100 に図示されたコンポーネントのいずれかまたはその組合せに対する依存性または要件を有するものと解釈してもならない。

#### 【 0025 】

本発明は、多数の汎用のまたは専用のコンピューティングシステム環境またはコンピューティングシステム構成と共に動作する。本発明と共に使用するのに適し得る周知のコンピューティングシステム、コンピューティング環境、および / またはコンピューティング構成の例には、パーソナルコンピュータ、サーバコンピュータ、ハンドヘルドデバイス、ラップトップデバイス、マルチプロセッサシステム、マイクロプロセッサベースのシステム、セットトップボックス、プログラマブルな家庭用電化製品、ネットワーク PC、ミニコンピュータ、メインフレームコンピュータ、上記のシステムまたはデバイスのいずれかを含む分散コンピューティング環境などが含まれるが、これに限定するものではない。 10

#### 【 0026 】

本発明を、プログラムモジュールなど、コンピュータによって実行される、コンピュータ実行可能命令の一般的なコンテキストで説明することができる。一般に、プログラムモジュールには、特定のタスクを実行するか特定の抽象データ型を実装する、ルーチン、プログラム、オブジェクト、コンポーネント、データ構造などが含まれる。本発明は、通信ネットワークを介してリンクされたリモート処理デバイスによってタスクが実行される分散コンピューティング環境でも実施することができる。分散コンピューティング環境では、プログラムモジュールを、メモリ記憶装置を含む、ローカルとリモートの両方のコンピュータ記憶媒体に配置することができる。 20

#### 【 0027 】

図 1 を参照すると、本発明を実施する例示的システムに、コンピュータ 110 の形の汎用コンピューティングデバイスが含まれる。コンピュータ 110 のコンポーネントに、処理装置 120、システムメモリ 130、およびシステムメモリを含むさまざまなシステムコンポーネントを処理装置 120 に結合するシステムバス 121 が含まれる。システムバス 121 は、メモリバス、メモリコントローラ、周辺バス、およびさまざまなバスアーキテクチャのいずれかを使用するローカルバスを含む複数のタイプのバス構造のいずれかとすることができます。限定ではなく例として、そのようなアーキテクチャに、Industrial Standard Architecture (ISA) バス、マイクロチャネルアーキテクチャ (MCA) バス、Enhanced ISA (EISA) バス、Video Electronics Standards Association (VESA) ローカルバス、およびメザニンバスとも称する Peripheral Component Interconnect (PCI) バスが含まれる。 30

#### 【 0028 】

コンピュータ 110 には、通常は、さまざまなコンピュータ可読媒体が含まれる。コンピュータ可読媒体は、コンピュータ 110 によってアクセスでき、揮発性媒体および不揮発性媒体、取外し可能媒体および取り外し不能媒体の両方を含む使用可能な媒体のいずれかとすることができます。限定ではなく例として、コンピュータ可読媒体に、コンピュータ記憶媒体および通信媒体を含めることができます。コンピュータ記憶媒体に、コンピュータ可読命令、データ構造、プログラムモジュール、または他のデータなどの情報を記憶する任意の方法または技術で実施された、揮発性および不揮発性、取外し可能および取り外し不能の両方の媒体が含まれる。コンピュータ記憶媒体に、RAM、ROM、EEPROM、フラッシュメモリ、および他のメモリ技術、CD-ROM、ディジタル多用途ディスク (DVD)、または他の光学ディスク記憶装置、磁気力セット、磁気テープ、磁気ディスク記憶装置、または他の磁気記憶装置、あるいは所望の情報の格納に使用でき、コンピュータ 110 によってアクセスできる他のすべての媒体が含まれるが、これに制限はされない。通信媒体は、通常は、搬送波または他のトランスポート機構などの変調されたデータ信号中にコンピュータ可読命令、データ構造、プログラムモジュール、または他のデータを具現化するものであり、任意の情報伝達媒体が含まれる。用語「変調されたデータ信号 40

50

」は、信号内で情報をエンコードする形でその特性の1つまたは複数を設定または変更された信号を意味する。限定ではなく例として、通信媒体に、有線ネットワークまたは直接配線接続などの有線媒体と、音響、RF、赤外線、および他の無線媒体などの無線媒体が含まれる。上記のいずれかの組合せも、コンピュータ可読媒体の範囲に含めるべきである。

#### 【0029】

システムメモリ130に、読み出し専用メモリ(ROM)131およびランダムアクセスメモリ(RAM)132などの揮発性メモリおよび/または不揮発性メモリの形のコンピュータ記憶媒体が含まれる。起動中などにコンピュータ110内の要素の間での情報の転送を助ける基本ルーチンを含む基本入出力システム133(BIOS)が、通常はROM131に格納される。RAM132には、通常は、処理装置120から即座にアクセス可能、かつ/または処理装置120によって現在操作中のデータおよび/またはプログラムモジュールが含まれる。限定ではなく例として、図1に、オペレーティングシステム134、アプリケーションプログラム135、他のプログラムモジュール136、およびプログラムデータ137を示す。

#### 【0030】

コンピュータ110に、他の取外し可能/取外し不能、揮発性/不揮発性のコンピュータ記憶媒体も含めることができる。例としてのみ、図1に、取外し不能不揮発性磁気媒体から読み取るかこれに書き込むハードディスクドライブ141、取外し可能不揮発性磁気ディスク152から読み取るかこれに書き込む磁気ディスクドライブ151、CD-ROMまたは他の光媒体などの取外し可能不揮発性光ディスク156から読み取るかこれに書き込む光ディスクドライブ155を示す。例示的なオペレーティング環境で使用できる他の取外し可能/取外し不能、揮発性/不揮発性のコンピュータ記憶媒体に、磁気テープカセット、フラッシュメモリカード、デジタル多用途ディスク、デジタルビデオテープ、ソリッドステートRAM、ソリッドステートROMなどが含まれるが、これに限定するものではない。ハードディスクドライブ141は、通常は、インターフェース140などの取外し不能メモリインターフェースを介してシステムバス121に接続され、磁気ディスクドライブ151および光ディスクドライブ155は、通常、インターフェース150などの取外し可能メモリインターフェースによってシステムバス121に接続される。

#### 【0031】

上で説明し、図1に示したドライブおよびそれに関連するコンピュータ記憶媒体は、コンピュータ110のコンピュータ可読命令、データ構造、プログラムモジュール、および他のデータの記憶を提供する。図1では、たとえば、ハードドライブ141が、オペレーティングシステム144、アプリケーションプログラム145、他のプログラムモジュール146、およびプログラムデータ147を格納するものとして図示されている。これらのコンポーネントを、オペレーティングシステム134、アプリケーションプログラム135、他のプログラムモジュール136、およびプログラムデータ137と同一のまたは異なるもののいずれかとすることができるに留意されたい。オペレーティングシステム144、アプリケーションプログラム145、他のプログラムモジュール146、およびプログラムデータ147は、最低限でも異なるコピーであることを示すために、異なる符号を与えられている。ユーザは、キーボード162および、一般にマウス、トラックボール、またはタッチパッドと称するポインティングデバイス161などの入力デバイスを介してコンピュータ110にコマンドおよび情報を入力することができる。他の入力デバイス(図示せず)に、マイクロホン、ジョイスティック、ゲームパッド、衛星パラボラアンテナ、スキャナなどを含めることができる。上記および他の入力デバイスは、しばしば、システムバス121に結合されたユーザ入力インターフェース160を介して処理装置120に接続されるが、パラレルポート、ゲームポート、またはuniversal serial bus(USB)などの他のインターフェースおよびバス構造によって接続することができる。モニタ191または他の種類のディスプレイデバイスも、ビデオインターフェース190などのインターフェースを介してシステムバス121に接続される。

10

20

30

40

50

モニタのほかに、コンピュータに、スピーカ 197 およびプリンタ 196 など、出力周辺装置インターフェース 195 を介して接続できる他の周辺出力デバイスも含めることができる。イメージ 193 のシーケンスをキャプチャできるカメラ 192（ディジタル／電子静止画カメラまたはビデオカメラ、あるいはフィルム／写真スキャナ）も、入力デバイスとしてパーソナルコンピュータ 110 に含めることができる。さらに、1つのカメラだけが図示されているが、複数のカメラを、入力デバイスとしてパーソナルコンピュータ 110 に含めることができる。1つまたは複数のカメラからのイメージ 193 が、適当なカメラインターフェース 194 を介してパーソナルコンピュータ 110 に入力される。このインターフェース 194 は、システムバス 121 に接続され、これによって、RAM 132 またはコンピュータ 110 に関連する他のデータ記憶装置にイメージをルーティングし、格納することができるようになる。しかし、イメージデータを、カメラ 192 の使用を必要とせずに、前述のコンピュータ可読媒体のいずれかを使用してコンピュータ 110 に入力することもできることに留意されたい。10

#### 【0032】

コンピュータ 110 は、リモートコンピュータ 180 などの1つまたは複数のリモートコンピュータへの論理接続を使用して、ネットワーク化された環境で動作することができる。リモートコンピュータ 180 は、パーソナルコンピュータ、サーバ、ルータ、ネットワーク PC、ピアデバイス、または他の一般的なネットワークノードとすることができ、通常は、上でコンピュータ 110 に関して説明した要素の多くまたはすべてが含まれるが、図 1 には、メモリ記憶装置 181 だけを示した。図 1 に示された論理接続に、ローカルエリアネットワーク（LAN）171 および広域ネットワーク（WAN）173 が含まれるが、他のネットワークも含めることができる。そのようなネットワーキング環境は、オフィス、会社全体のコンピュータネットワーク、インターネット、およびインターネットでごく一般的なものである。20

#### 【0033】

LAN ネットワーキング環境で使用される時に、コンピュータ 110 は、ネットワークインターフェースまたはネットワークアダプタ 170 を介して LAN 171 に接続される。WAN ネットワーキング環境で使用される時に、コンピュータ 110 に、通常は、インターネットなどの WAN 173 を介する通信を確立する、モデム 172 または他の手段が含まれる。モデム 172 は、内蔵または外付けとすることができますが、ユーザ入力インターフェース 160 または他の適当な手段を介してシステムバス 121 に接続することができます。ネットワーク化された環境では、コンピュータ 110 に関して図示されたプログラムモジュールまたはその一部を、リモートメモリ記憶装置に保管することができます。限定ではなく例として、図 1 に、メモリデバイス 181 に常駐するものとしてリモートアプリケーションプログラム 185 を示す。図示のネットワーク接続が例示的であり、コンピュータの間の通信リンクを確立する他の手段を使用できることを理解されたい。30

#### 【0034】

##### 2.0.2 レイヤイメージ表現

例示的なオペレーティング環境を述べたので、この説明セクションの残りは、本発明を実施するプログラムモジュールの説明に捧げられる。一般に、本発明に、イメージの圧縮、転送、および記憶を容易にする、独自の 2 レイヤ 3D 表現の生成が含まれる。この表現は、イメージが動的シーンのビデオのフレームであり、フレームデータがリアルタイムレンダリングのためにエンコードされる時に、特に有用である。この表現には、ピクセル視差情報または深さ情報も含まれ、したがって、表現の 3D 様態が提供される。イメージまたはフレームは、ディジタルカメラによってキャプチャされたか、そうでない場合にさらなる処理を実行する前にディジタル化されたのいずれかであるディジタルイメージである。ディジタルイメージデータを、カメラを使用してシーンの実際のイメージをキャプチャすることによって入手されたイメージベースのものまたは合成されたイメージデータとすることができることにも留意されたい。40

#### 【0035】

2 レイヤ表現は、まず、検討中のイメージまたはフレームの視差マップの深さ不連続を突き止めることによって生成される。これらの深さ不連続は、指定された数の視差レベル（たとえば、本発明のテストされた実施形態では 4 レベル）を超えるジャンプと定義される。検討中のイメージの視差マップは、普通の形で入手することができる。しかし、本発明のテストされた実施形態では、共通の譲受人に譲渡された、本出願の発明者による「Color Segmentation-Based Stereo Reconstruction System And Process」という名称の同時係属出願の主題である新しい手法を使用する。

#### 【 0 0 3 6 】

次に、検討中のイメージで見つかった深さ不連続の近傍で、小さい領域を識別する。この小さい領域は、深さ不連続の位置から 3 ピクセル以内のすべてのピクセルを含むものとして定義される。これを、図 2 のグラフに示すが、図 2 では、ピクセルの行の視差値が、ピクセル位置に対してプロットされている。視差値の急激なステップは、それが指定された視差レベルの数より大きいと仮定して、深さ不連続を表す。前述の小さい領域は、深さ不連続領域と称するが、不連続性の周囲で確立される。図 2 のグラフで表されたプロファイルビューでは、この領域が、イメージ内の深さ視差の位置にまたがる、ある幅を有する。通常そうであるように、深さ不連続がイメージ内の物体の輪郭に追随する（flow）場合に、深さ不連続領域は、この輪郭に追随するストリップを形成するためにマージされる。

#### 【 0 0 3 7 】

物体の境界に沿った一部のピクセルは、バックグラウンド領域とフォアグラウンド領域の両方からの寄与を受け取る。しかし、オリジナルの混合されたピクセルカラーがレンダリング中に使用される場合に、可視のアーティファクトが生じる可能性がある。したがって、この 2 つの寄与を分離することが重要である。したがって、次に、深さ不連続領域内の各ピクセルのフォアグラウンドカラーとバックグラウンドカラーが、確立され、ピクセルの不透明度（アルファ値の形）も確立される。これは、マッティング技法を使用して達成される。一般に、マッティング（matting）技法には、各ピクセルのフォアグラウンド要素とバックグラウンド要素に由来するカラーおよび不透明度を推定することによって、イメージのフォアグラウンド要素とバックグラウンド要素を抽出することが含まれる。ピクセルの不透明度は、0 と 1 の間の範囲を有するアルファ値によって定義される。事実上、アルファ値は、ピクセルのカラーのうちでフォアグラウンド要素に寄与するパーセンテージを定義する。通常のマッティング動作の正味の結果は、検討されるピクセルごとにフォアグラウンドカラー、バックグラウンドカラー、およびアルファ値を識別することである。本発明と共に、どのマッティング手順でも使用することができるが、テストされた実施形態では、参考文献 [4] で説明したベイジアンイメージマッティング技法を使用した。参考文献 [4] に、多数の他の既存のマッティング技法の説明も含まれることに留意されたい。

#### 【 0 0 3 8 】

前に説明したように、本発明の主な応用例に、シーンの 2 つの実際のイメージに関連する視点の中間の視点からのシーンの仮想イメージをレンダリングすることが含まれる。これを達成する方法の詳細は、本発明の範囲に含まれないが、この処理に、通常、各ピクセルに関連する深さを知ることが含まれることに留意されたい。このために、本発明によるイメージ表現に、ピクセル深さ（または、簡単に深さ値に変換できる視差値）が含まれる。これらの値は、前述の視差マップからほとんどのピクセルについて入手可能であるが、深さ不連続に隣接して見つかる混合されたピクセルに関連する 2 つの深さすなわち、フォアグラウンド要素に関連する深さおよびバックグラウンド要素に関連する深さがあることに留意されたい。したがって、これらのフォアグラウンド深さおよびバックグラウンド深さ（または視差）値が、次に、深さ不連続領域内のピクセルごとに推定される。一般に、これは、それぞれイメージのフォアグラウンド部分およびバックグラウンド部分の近傍深さのアルファ加重平均を使用することによって達成される。具体的に言うと、フォアグラウンド視差は、深さ不連続領域のオリジナルフォアグラウンド領域内のピクセルからの視

10

20

30

40

50

差値のウィンドウ（たとえば、 $7 \times 7$  ピクセル）内のアルファ加重平均を使用することによって見つかる。図2では、オリジナルフォアグラウンド領域が、視差不連続性の左側のピクセルを指す。バックグラウンド視差は、深さ不連続領域のオリジナルバックグラウンド領域内の視差値のウィンドウ内の加重平均によって見つかる。バックグラウンド視差の平均をとるのに使用される重みは、1からアルファ値を引いたものと等しくなるように設定される。

#### 【0039】

深さ不連続領域ピクセルのカラー、視差、および不透明度が確立されたならば、イメージ表現の2つのレイヤが形成される。具体的に言うと、深さ不連続領域内で見つかった各めいめいのピクセルに関連する前に計算されたフォアグラウンドカラー、フォアグラウンド視差、およびアルファ値が、イメージの境界レイヤの対応する位置のピクセルに割り当てられる。同様に、深さ不連続領域内の各めいめいのピクセルに関連する前に計算されたバックグラウンドカラーおよびバックグラウンド視差値が、深さ不連続領域内にないイメージ内のすべてのピクセルのカラー値および視差値（イメージおよびその視差マップからとられる）と共に、イメージのメインレイヤの対応する位置のピクセルに割り当てられる。

10

#### 【0040】

したがって、結果の2レイヤイメージ表現に、そのレイヤの各ピクセルのフォアグラウンドカラー、フォアグラウンド視差、およびアルファ値を識別する境界レイヤが含まれる。さらに、この表現に、深さ不連続領域に関連する、そのレイヤ内のピクセルのバックグラウンドカラーおよびバックグラウンド深さと、そのレイヤの他のピクセルのそれぞれのカラーおよび視差値を識別するメインレイヤが含まれる。これを与えられれば、メインレイヤに、イメージ内のすべてのピクセルに関するピクセル情報が含まれるが、このデータは、シーンの1つのイメージと次のイメージの間で、比較的類似するであろう。これは、イメージがビデオフレームである場合に時間的にそうであり、同一シーンの複数のイメージが送信または記憶される場合に空間的にそうである。これによって、メインレイヤが、標準的なイメージ圧縮方法を使用してかなり圧縮可能になる。さらに、通常、イメージのピクセルの総数のうちの小さい比率だけが、境界レイヤに含まれるので、境界レイヤには、比較的少ないデータが含まれる。したがって、このデータは、動的シーンのあるイメージから次のイメージへ、経時的に大きく変化する可能性が高いが、送信または記憶する必要があるデータは、多くない。このゆえに、圧縮しなくとも、このデータをすぐに送信または記憶することができる。適当な圧縮方式を使用すると、データの量およびこれを転送できる速度を、さらに減らすことができる。この形で、前述の2レイヤイメージ表現は、データ転送および記憶の望ましい容易化を提供し、リアルタイムレンダリングを実現可能にする。

20

#### 【0041】

前述の2レイヤイメージ表現生成処理の概要を、図3A～Bに示された流れ図を参照して示す。まず、表現されるデジタルイメージまたはデジタル化されたイメージとその視差マップを入力する（処理アクション300）。次に、視差マップを使用して、イメージ内で深さ不連続の位置を識別する（処理アクション302）。次に、深さ不連続を囲む指定されたサイズの領域内で見つかるピクセルを、識別し、イメージの深さ不連続領域内にあるものとして指定する（処理アクション304）。マッティング技法を使用して、深さ不連続領域内の各ピクセルのフォアグラウンドカラーおよびバックグラウンドカラーを、そのアルファ値と共に確立する（処理アクション306）。さらに、処理アクション308で、深さ不連続領域内のピクセルのそれについて、フォアグラウンド視差値およびバックグラウンド視差値を確立する。次に、深さ不連続領域のピクセルに対応するピクセルを有する境界レイヤを確立する（処理アクション310）。深さ不連続領域内で見つかる各めいめいのピクセルに関連するフォアグラウンドカラー、フォアグラウンド視差、およびアルファ値を、境界レイヤの対応するピクセルに割り当てる（処理アクション312）。イメージのすべてのピクセルに対応するピクセルを有するメインレイヤも確立する

30

40

50

(処理アクション 314)。この場合に、深さ不連続領域のピクセルに関連するバックグラウンドカラーおよびバックグラウンド視差を、メインレイヤの対応するピクセルに割り当て、深さ不連続領域にないイメージピクセルに関連するカラーおよび視差を、メインレイヤの対応するピクセルに割り当てる(処理アクション 316)。

#### 【0042】

本発明による2レイヤイメージ表現を生成する前述の処理に組み込むことができる改善の1つに、イメージ表現の境界レイヤおよびメインレイヤを形成する前の深さ不連続領域のフォアグラウンドマットの膨張が含まれる。これは、この処理の不正確さに起因して発生する可能性がある、レイヤからのイメージのレンダリング中に割れ目が見えることをなくすという長所を有する。具体的に言うと、フォアグラウンドカラー、フォアグラウンド視差、およびアルファ値が、深さ不連続領域のピクセルごとに確立されたならば、この領域のサイズを、指定された量(たとえば1ピクセル)だけ増やす。膨張中に境界レイヤに追加される新しいピクセルは、メインレイヤの対応するピクセルと同一のカラー値および視差値と1のアルファ値を割り当てる。

10

#### 【0043】

図4A～4Eに、ブレークダンサー(breakdancers)のグループのイメージに本発明の2レイヤイメージ表現生成処理を適用した結果を示す。図4Aは、メインレイヤカラー推定値を示し、図4Bは、メインレイヤ視差推定値を表す。同様に、図4Cは、境界レイヤカラー推定値を示し、図4Dは、境界レイヤ視差推定値を示す。最後に、図4Eは、境界レイヤアルファ推定値を示す。図4C～4Eのイメージが、反転して示され、その結果、透明／空のピクセルが白く見えていることに留意されたい。ソフトオブジェクト境界を考慮に入れるために送信する必要がある情報の量がどれほど少ないかということ、境界レイヤの不透明度および両方のレイヤのカラーがどれほどきれいに回復されているかということに留意されたい。

20

#### 【0044】

前述の説明を介して、イメージピクセルの視差値が使用されたことに留意されたい。しかし、生成処理のある点で、これらの値を、標準的な方法を使用して深さ値に変換することができる。その場合に、メインレイヤおよび境界レイヤに、それに含まれるピクセルの、視差値ではなく深さ値が含まれる。

#### 【0045】

30

#### 3.0 参考文献

#### 【0046】

## 【表1-1】

- [1] Buehler, C., Bosse, M., McMillan, L., Gortler, S. J., and Cohen, M. F. 2001. Unstructured Lumigraph rendering. *Proceedings of SIGGRAPH 2001* (August), 425–432.
- [2] Carceroni, R. L., and Kutulakos, K. N. 2001. Multi-view scene capture by surfel sampling: From video streams to non-rigid 3D motion, shape and reflectance. In *Eighth International Conference on Computer Vision (ICCV 2001)*, vol. II, 60–67. 10
- [3] Carranza, J., Theobalt, C., Magnor, M. A., and Seidel, H.-P. 2003. Free-viewpoint video of human actors. *ACM Transactions on Graphics* 22, 3 (July), 569–577.
- [4] Chuang, Y.-Y., et al. 2001. A Bayesian approach to digital matting. In *Conference on Computer Vision and Pattern Recognition (CVPR'2001)*, vol. II, 264–271. 20
- [5] Debevec, P. E., Taylor, C. J., and Malik, J. 1996. Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. *Computer Graphics (SIGGRAPH'96)* (August), 11–20. 30
- [6] Debevec, P. E., Yu, Y., and Borshukov, G. D. 1998. Efficient view-dependent image-based rendering with projective texture-mapping. *Eurographics Rendering Workshop 1998*, 105–116.

【0047】

## 【表1 - 2】

[7] Gortler, S. J., Grzeszczuk, R., Szeliski, R., and Cohen, M. F. 1996. The Lumigraph. In *Computer Graphics (SIGGRAPH'96) Proceedings*, ACM SIGGRAPH, 43–54.

[8] Hall-Holt, O., and Rusinkiewicz, S. 2001. Stripe boundary codes for real-time structured-light range scanning of moving objects. In *Eighth International Conference on Computer Vision (ICCV 2001)*, vol. II, 359–366. 10

[9] Heigl, B., et al. 1999. Plenoptic modeling and rendering from image sequences taken by hand-held camera. In *DAGM'99*, 94–101.

[10] Kanade, T., Rander, P. W., and Narayanan, P. J. 1997. Virtualized reality: constructing virtual worlds from real scenes. *IEEE MultiMedia Magazine* 1, 1 (Jan-March), 34–47. 20

[11] Levoy, M., and Hanrahan, P. 1996. Light field rendering. In *Computer Graphics (SIGGRAPH'96) Proceedings*, ACM SIG-GRAPH, 31–42.

[12] Pulli, K., et al. 1997. View-based rendering: Visualizing real objects from scanned range and color data. In *Proceedings of the 8th Eurographics Workshop on Rendering*. 30

[13] Scharstein, D., and Szeliski, R. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47, 1 (May), 7–42.

[14] Seitz, S. M., and Dyer, C. M. 1997. Photorealistic scene reconstruction by voxel coloring. In *Conference on Computer Vision and Pattern Recognition (CVPR'97)*, 1067–1073. 40

## 【0048】

## 【表 1 - 3】

- [15] Shade, J., Gortler, S., He, L.-W., and Szeliski, R. 1998. Layered depth images. In *Computer Graphics (SIGGRAPH'98) Proceedings*, ACM SIGGRAPH, Orlando, 231–242.
- [16] Tao, H., Sawhney, H., and Kumar, R. 2001. A global matching framework for stereo computation. In *Eighth International Conference on Computer Vision (ICCV 2001)*, vol. I, 532–539. 10
- [17] Vedula, S., Baker, S., Seitz, S., and Kanade, T. 2000. Shape and motion carving in 6D. In *Conference on Computer Vision and Pattern Recognition (CVPR'2000)*, vol. II, 592–598.
- [18] Yang, J. C., Everett, M., Buehler, C., and McMillan, L. 2002. A real-time distributed light field camera. In *Eurographics Workshop on Rendering*, P. Debevec and S. Gibson, Eds., 77–85. 20
- [19] Zhang, Y., and Kambhamettu, C. 2001. On 3D scene flow and structure estimation. In *Conference on Computer Vision and Pattern Recognition (CVPR'2001)*, vol. II, 778–785.
- [20] Zhang, L., Curless, B., and Seitz, S. M. 2003. Spacetime stereo: Shape recovery for dynamic scenes. In *Conference on Computer Vision and Pattern Recognition*, 367–374. 30
- [21] Zhang, Z. 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22, 11, 1330–1334.
- 【図面の簡単な説明】
- 【0049】
- 【図1】本発明を実施する例示的なシステムを構成する汎用コンピューティングデバイスを示す図である。 40
- 【図2】視差値の急激なステップが深さ不連続性を表す、ピクセル位置に対するピクセルの行の視差値をプロットしたグラフである。
- 【図3A】本発明によるディジタルイメージまたはディジタル化されたイメージの2レイヤ表現を生成する処理を示す流れ図である。
- 【図3B】本発明によるディジタルイメージまたはディジタル化されたイメージの2レイヤ表現を生成する処理を示す流れ図である。
- 【図4A】ブレークダンサのグループのイメージに図3A～Bの2レイヤイメージ表現生成処理を適用した結果を示すイメージであり、メインレイヤカラー推定値を示す図である 50

。

【図4B】ブレークダンサのグループのイメージに図3A～Bの2レイヤイメージ表現生成処理を適用した結果を示すイメージであり、メインレイヤ視差推定値を表す図である。

【図4C】ブレークダンサのグループのイメージに図3A～Bの2レイヤイメージ表現生成処理を適用した結果を示すイメージであり、境界レイヤカラー推定値を示す図である。

【図4D】ブレークダンサのグループのイメージに図3A～Bの2レイヤイメージ表現生成処理を適用した結果を示すイメージであり、境界レイヤ視差推定値を示す図である。

【図4E】ブレークダンサのグループのイメージに図3A～Bの2レイヤイメージ表現生成処理を適用した結果を示すイメージであり、境界レイヤアルファ推定値を示す図である

。

10

#### 【符号の説明】

##### 【0050】

130 システムメモリ

131 ROM

133 BIOS

132 RAM

134 オペレーティングシステム

135 アプリケーションプログラム

136 他のプログラムモジュール

137 プログラムデータ

20

120 処理装置

194 カメラインターフェース

190 ビデオインターフェース

195 出力周辺装置インターフェース

192 カメラ

191 モニタ

196 プリンタ

197 スピーカ

121 システムバス

140 取り外し不能不揮発性メモリインターフェース

30

150 取り外し可能不揮発性メモリインターフェース

160 ユーザ入力インターフェース

170 ネットワークインターフェース

171 ローカルエリアネットワーク

144 オペレーティングシステム

145 アプリケーションプログラム

146 他のプログラムモジュール

147 プログラムデータ

161 マウス

162 キーボード

40

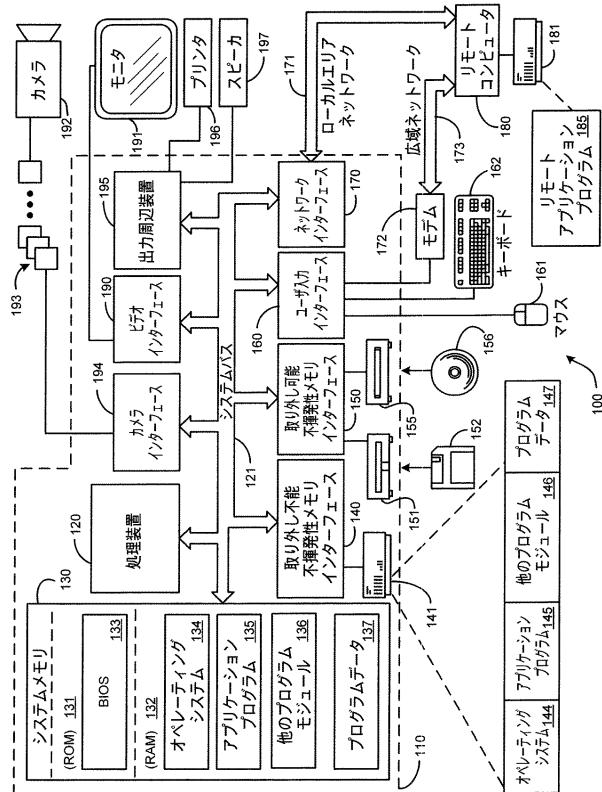
172 モデム

173 広域ネットワーク

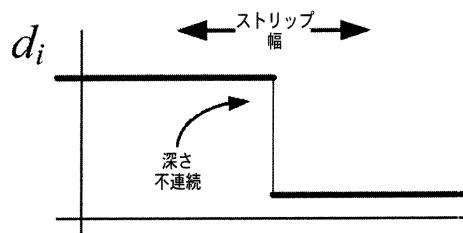
180 リモートコンピュータ

185 リモートアプリケーションプログラム

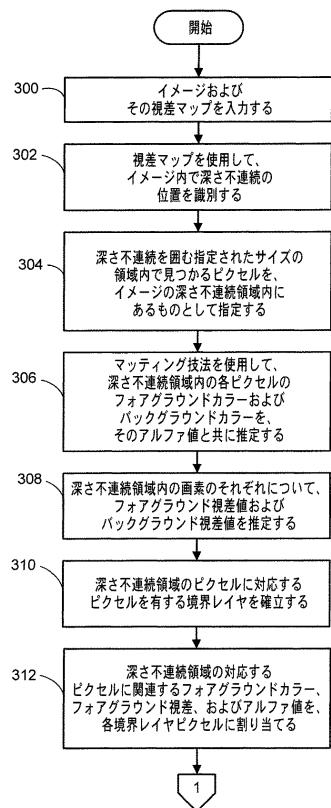
【図1】



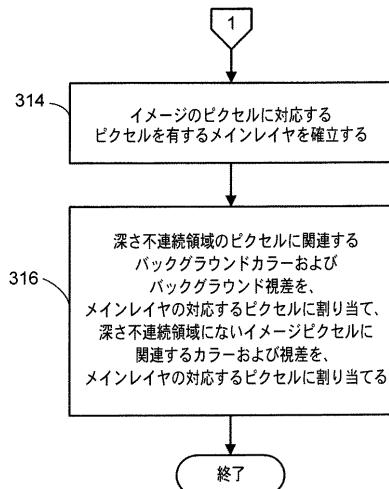
【図2】



【図3A】



【図3B】



【図4A】



【図4B】



【図4C】



【図4D】



【図4E】



---

フロントページの続き

(72)発明者 マシュー ウィッテンダール  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション内

(72)発明者 リチャード スゼリスキ  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション内

(72)発明者 シモン ウィンダー<sup>1</sup>  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション内

(72)発明者 シング ビン カン  
アメリカ合衆国 98052 ワシントン州 レッドモンド ワン マイクロソフト ウェイ マ  
イクロソフト コーポレーション内

審査官 岡本 俊威

(56)参考文献 特開2003-331317(JP,A)  
特開平08-227464(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 1/00, 17/00 - 17/40  
H04N 13/00