

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4904264号  
(P4904264)

(45) 発行日 平成24年3月28日 (2012.3.28)

(24) 登録日 平成24年1月13日 (2012.1.13)

(51) Int. Cl. F I  
**G06T 13/40 (2011.01)** G O 6 T 15/70 B  
**G06T 1/00 (2006.01)** G O 6 T 1/00 3 1 5

請求項の数 15 (全 39 頁)

(21) 出願番号	特願2007-523240 (P2007-523240)	(73) 特許権者	507031974
(86) (22) 出願日	平成17年7月31日 (2005.7.31)		エクストリーム リアリティー エルティ ーディー.
(65) 公表番号	特表2008-508590 (P2008-508590A)		イスラエル国 ヘルツェリア ベタック
(43) 公表日	平成20年3月21日 (2008.3.21)		4 6 7 2 5 アッパ イーヴン プールバ ード 1 6
(86) 国際出願番号	PCT/IL2005/000813	(74) 代理人	100078880
(87) 国際公開番号	W02006/011153		弁理士 松岡 修平
(87) 国際公開日	平成18年2月2日 (2006.2.2)	(74) 代理人	100169856
審査請求日	平成20年7月29日 (2008.7.29)		弁理士 尾山 栄啓
(31) 優先権主張番号	60/592,136	(74) 代理人	100148895
(32) 優先日	平成16年7月30日 (2004.7.30)		弁理士 荒木 佳幸
(33) 優先権主張国	米国 (US)	(72) 発明者	ギボン, ドール
前置審査			イスラエル リション-レジオン 7 5 2 4 1, ヴィニクストリート 1 1 / 1 0 最終頁に続く

(54) 【発明の名称】 3D空間次元に基づく画像処理のためのシステムおよび方法

(57) 【特許請求の範囲】

【請求項1】

互いの相対位置が3次元骨格モデルに実質的に従う構造を有する任意の非剛体の本体の1又はそれ以上の部分の3次元座標を推定するための装置であって、前記装置が：

前記非剛体の本体の2次元画像を取得するための手段と、

前記2次元画像に基づき、前記非剛体の本体の一連の構造を計算で特定する構造特定器と、

骨格に基づく位置推定部と、

を具備しており、前記位置推定部が：

(1) 予め記憶された3次元骨格モデルを用い、特定された構造を、前記非剛体の本体の2又はそれ以上の構造間の空間的関係の制約を規定する3次元骨格モデルの対応する要素に適合させ、

(2) 前記2次元画像における、前記特定された構造の2次元の空間的関係に基づいて、前記3次元骨格モデル要素の移動関係を定義することによって、前記特定された構造の3次元位置情報を推定するよう構成されていることを特徴とする装置。

【請求項2】

前記関係の制約が、前記非剛体の本体の2又はそれ以上の構造間の角度の比及び距離の比を有することを特徴とする請求項1に記載の装置。

【請求項3】

さらに、前記推定部が、前記2次元画像の中の前記特定された構造の2次元の空間的関

係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することによって、他の構造の 3 次元位置情報を推定するよう構成されていることを特徴とする請求項 1 に記載の装置。

【請求項 4】

前記 2 次元画像の中の前記特定された構造の 2 次元の空間的關係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することが、対応する 2 次元投影が前記 2 次元画像と少なくとも部分的に一致するモデルに関する許容される 3 次元方向を判断することを含んでいることを特徴とする請求項 1 に記載の装置。

【請求項 5】

前記少なくとも部分的に一致することが、前記画像の特定された構造を骨格投影の対応する 3 次元骨格要素に一致させることを含んでいることを特徴とする請求項 4 に記載の装置。

10

【請求項 6】

互いの相対位置が 3 次元骨格モデルに実質的に従う構造を有する任意の非剛体の本体の 1 又はそれ以上の部分の 3 次元座標を推定するためのシステムであって、

前記推定が、前記非剛体の本体の 2 次元画像に基づいており、前記システムが：

前記 2 次元画像に基づき前記非剛体の本体の一連の構造を特定し、予め記憶された 3 次元骨格モデルを用いて、特定された構造を前記非剛体の本体の 2 又はそれ以上の構造間の空間的制約を規定する 3 次元骨格モデルの対応する要素に適合させるよう構成された画像処理ロジックを具えており、

20

前記ロジックが、さらに、前記 2 次元画像の中の前記特定された構造の 2 次元の空間的關係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することによって、前記特定された構造の 3 次元位置情報を推定するよう構成されていることを特徴とするシステム。

【請求項 7】

前記関係の制約が、前記非剛体の本体の 2 又はそれ以上の構造間の角度の比及び距離の比を有することを特徴とする請求項 6 に記載のシステム。

【請求項 8】

さらに、前記 2 次元画像の中の前記特定された構造の 2 次元の空間的關係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することによって、他の構造の 3 次元位置情報を推定することを含むことを特徴とする請求項 6 に記載のシステム。

30

【請求項 9】

前記 2 次元画像の中の前記特定された構造の 2 次元の空間的關係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することが、対応する 2 次元投影が前記画像と少なくとも部分的に一致するモデルに関する許容される 3 次元方向を判断することを含んでいることを特徴とする請求項 6 に記載のシステム。

【請求項 10】

前記少なくとも部分的に一致することが、前記画像の特定された構造を骨格投影の対応する 3 次元の骨格要素に一致させることを含んでいることを特徴とする請求項 9 に記載のシステム。

40

【請求項 11】

互いの相対位置が 3 次元骨格モデルに実質的に従う構造を有する任意の非剛体の本体の 1 又はそれ以上の部分の 3 次元座標を推定するための画像トラッキングシステムであって、前記推定が、前記非剛体の本体の 2 次元画像に基づいており、前記システムが：

前記非剛体の本体を含む 2 次元画像に関するデータを受け取るよう構成された画像取得回路と；

前記 2 次元画像に基づき前記非剛体の本体の一連の構造を特定し、予め記憶された 3 次元骨格モデルを用いて、特定された構造を、3 次元骨格モデルの対応する要素に適合させるよう構成された画像処理ロジックと；

を具えており、

50

前記ロジックが、さらに、前記 2 次元画像の中の前記特定された構造の 2 次元の空間的關係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することによって、前記特定された構造の 3 次元位置情報を推定することを特徴とするシステム。

【請求項 1 2】

前記關係の制約が、前記非剛体の本体の 2 又はそれ以上の構造間の角度の比及び距離の比を有することを特徴とする請求項 1 1 に記載のシステム。

【請求項 1 3】

さらに、前記 2 次元画像の中の前記特定された構造の 2 次元の空間的關係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することによって、他の構造の 3 次元位置情報を推定することを含んでいることを特徴とする請求項 1 1 に記載のシステム。

10

【請求項 1 4】

前記 2 次元画像の中の前記特定された構造の 2 次元の空間的關係に基づいて、前記 3 次元骨格モデル要素の移動関係を定義することが、対応する 2 次元投影が前記画像と少なくとも部分的に一致するモデルに関する許容される 3 次元方向を判断することを含んでいることを特徴とする請求項 1 1 に記載のシステム。

【請求項 1 5】

前記少なくとも部分的に一致することが、前記画像の特定された構造を骨格投影の対応する3 次元の骨格要素に一致させることを含んでいることを特徴とする請求項 1 4 に記載のシステム。

【発明の詳細な説明】

20

【技術分野】

【0 0 0 1】

本発明は、撮影術、画像処理およびアニメーションに関し、より特定的には、3 次元 (3 D) 撮影術、3 次元画像処理および 3 次元アニメーションに関するが、これらに限定されるものではない。

【背景技術】

【0 0 0 2】

現在技術による 3 次元撮影術は、時間次元に基づく。

【0 0 0 3】

本発明は、例えば、立体画像、球体撮影システム、3 D コンピュータアニメーション、3 D 撮影術、および 3 D 画像処理アルゴリズム等、3 D 映像および画像処理の世界に属するいくつかの異なる分野に関する。

30

【0 0 0 4】

従来の 3 D 立体撮影術は、位置合わせされたレンズ間に平行な光軸を有し、それらレンズ間の距離が固定されたツインカメラを採用している。これらのツインカメラは、立体表示および立体視に関する公知技術のいずれかによって表示することができる一対の画像を作成する。これらの技術は、一般に、右側のレンズにより撮影された画像が観察者の右目に対して表示され、左側のレンズにより撮影された画像が観察者の左目に対して表示されるという原理に基づく。

【0 0 0 5】

40

例えば、「3 次元表示用途のためのデジタルフォーマッタ (Digital formatter for 3-dimensional display applications)」と題されるテキサス・インスツルメンツ社に付与された特許文献 1 は、4 重メモリバッファを用いて、右目および左目の両方に対する表示用の処理済ビデオデータを記憶したり読み出したりする 3 D デジタル投影ディスプレイを開示している。このフォーマッタを用いて、ビデオデータが 48 フレーム毎秒の速度で処理され、2 度 (繰り返して) 読み出されることにより、表示のちらつきに関する閾値を越える 96 (最大で 120) フレーム毎秒のフラッシュ速度を達成する。このデータは、次いで、右目および左目のフレームが正確に位相ずれを起こした状態でヘッドセットまたはゴーグルと同期して、知覚できる 3 D 画像を生成する。

50

## 【0006】

球体またはパノラマ写真撮影術では、伝統的に、「魚眼」レンズ等の非常に広角なレンズによって、または重なり合って隣接する画像の「継ぎ合わせ」によって、最大で完全な球体となるような広範な視野をカバーしている。このような技術を用いて得られるパノラマまたは球体画像は、観察者に奥行きを感じさせる2次元画像または立体画像とすることができる。これらの画像はまた、三角測量法等の公知の方法を用いてカメラから画像内の各ピクセルの距離を計算する等、コンピュータ処理された3次元(3D)画像とすることができる。

## 【0007】

例えば、テンペスト・マイクロシステムズ社(Tempest Microsystems Incorporated)に付与された特許文献2は、魚眼レンズおよび電荷結合素子(CCD)等の撮像装置を用いて、例えば、半球体視野の画像等の広角画像を得る画像取得観察システムを教示している。

10

## 【0008】

2001年11月28日に出願された本出願人による同時継続中の特許文献3も参照し、その内容を本明細書中において参考として援用する。この出願は、観察者の周囲の視環境について水平および垂直両方向360度の完全な立体的球体画像を得るための撮像システムを教示している。立体表示に適した手段を用いて画像を表示することにより、観察者は、上下方向を含めた周囲の全てを見渡すことができ、表示された画像の立体的奥行きを感じることができる。本開示は、カメラのレンズが曲面上に配置され、当該曲面上の共通中心Cから突き出たカメラレイを教示している。キャプチャされた画像は編集および処理され、各ペアにおける一方の画像が観察者の右目用で、他方の画像が左目用である数組の立体画像のペアを生成することにより、3次元の知覚を生じる。

20

## 【0009】

3Dコンピュータアニメーションは、近年普及してきた「仮想現実」の分野に関する。3D仮想現実は、実際の画像と、合成により作成された画像とを組み合わせで構築される。また、完全にコンピュータによって生成された仮想現実も存在する。3D仮想現実は、撮影されている要素の3D情報を作成するために写真撮影された画像の3D計算を必要とする。

## 【0010】

これは、能動的方法を用いてリアルタイムで行なうことが可能である。

30

## 【0011】

例えば、3DVシステムズ社(3DV systems Incorporated (<http://www.3dvsystems.com/>))は、場面内の各ピクセルの奥行き値および色値をリアルタイムでキャプチャすることにより、場面のフレーム毎に、距離をグレーレベルで表した奥行き地図を作成するZCam(登録商標)カメラを提供する。ZCam(登録商標)カメラは、適切な幅を有する光の壁を用いる独自設計のカメラである。この光の壁は、例えば、矩形のレーザパルスとして生成してもよい。実写された場面内のオブジェクトに光の壁が当たると、それらオブジェクトの印象を写したZCam(登録商標)カメラに向かって反射される。この印象は、奥行き地図の再構築に必要な全ての情報を搬送する。

40

## 【0012】

実写画像の3D計算はまた、受動的方法を用いて行なってもよい。

## 【0013】

奥行き構築のための受動的方法には、場面に対する視点の少なくとも2つが既知である三角測量技術を用いてもよい。対応する特徴を特定し、光線を交差させて各特徴の3D位置を見つける。時空立体像により、空間照合関数において用いられる近傍に時間次元を加える。時間を渡った多数のフレームを用いて時間的立体像を加えることにより、1つ目の画像の単一のピクセルを、2つめの画像と照合する。これは、通常の特徴に基づく画像間の照合技術において行なわれる対象点(角)の照合とは対照的に、移動するオブジェク

50

トの時空間軌跡を照合することによっても行なうことができる。これらのシーケンスは、対応する時空間軌跡に沿った全ての点を整合的に照合し、また、2つのビデオシーケンス間におけるサブフレームの時間的な一致（同期）を得ることにより空間および時間的に照合される。

【0014】

3Dコンピュータにより生成された画像（CGI）は仮想世界であり、3Dコンピュータにより生成された画像用ソフトウェアを用いて作成された指定領域である。この仮想世界は指定領域に作成され、当該仮想世界内の各点はコンピュータにより生成された点である。この仮想世界には、2Dまたは3Dの現実の画像を組み込む場合もある。

【0015】

ここで、公知技術による仮想世界を示す図1を参照する。

【0016】

この仮想世界における各点の3D位置は既知である。この空間内の特定の点に色、輝度等の詳細を加えることにより、空間内に各種形状を作成する（図1）。仮想的なカメラをこの世界に配置することにより、仮想世界内に時間に基づくシーケンスを作成し、立体画像等を作成することができる。

【0017】

実写画像とコンピュータにより生成された世界の同期は、空間同期を用いた後、時間同期を用いて、現実世界の画像を空間および時間的に仮想世界と合わせることにより行なうことが可能である。

【0018】

ここで、従来技術による仮想的なスタジオを示す図2を参照する。

【0019】

この例では、ブルー/グリーンスクリーンとして公知の技術において、カメラが人物の姿とその背景を分離することができる仮想的なスタジオを用いる。図3に示すように、人物の姿をその周囲と分けることにより、その姿をコンピュータで作成した仮想的な世界と組み合わせることができる。

【0020】

また、全く逆のことが、既知の固定点を予め設定し、コンピュータにより生成された3D世界に同期する固定点を作成したバスケットボール場等の所定空間において、1組のカメラを監視することにより行なうことができる。このような技術により、CGI図形を分離してバスケットボール場に組み込むことが可能である。例えば、ORAD社（ORAD Incorporated）の製品であるCyberSport（登録商標）は、バスケットボール場、フットボール場等で行なわれるスポーツイベントに関して、競技場に密着した生の3D画像を挿入し、挿入された画像オブジェクトがイベントと一体化して、その一部となる幻想を作成する。

【0021】

上述したように、3D撮像および立体撮影術のための伝統的な方法およびシステムは、特殊なカメラ、特殊なレンズ、2つ以上のカメラの所定の位置取り、および専用アルゴリズムに基づいている。

【特許文献1】米国特許第6,906,687号

【特許文献2】米国特許第6,833,843号

【特許文献3】米国特許出願第10/416,533号

【発明の開示】

【発明が解決しようとする課題】

【0022】

それゆえ、撮像および撮影のためのシステムおよび方法に対する要望が広く認識されており、それらを有することは大きな利点となる。

【課題を解決するための手段】

【0023】

10

20

30

40

50

本発明の1つの局面によると、画像データの3D表現のための装置であって、該装置は、前記画像データ内において運動中の構造を特定する構造特定器と、前記構造特定器と関連づけられた骨格挿入部であって、3次元骨格要素が前記構造とともに移動することにより、前記画像データの運動および構造を3次的に解釈させることができるように、該骨格要素と該構造とを関連づける骨格挿入部と、を備える、装置が提供される。

**【0024】**

本発明の第2の局面によると、画像データの3D表現のための方法であって、前記画像データ内の構造を特定する工程と、3次元骨格要素が前記構造とともに移動することにより、前記画像データを3次的に解釈させることができるように、該骨格要素と該構造とを関連づける工程と、を包含する、方法が提供される。

10

**【0025】**

本発明の第3の局面によると、奥行き情報を用いて入力データを記録する記録装置であって、前記画像データ内において運動中の構造を特定する構造特定器と、前記構造特定器と関連づけられた骨格挿入部であって、3次元骨格要素が前記構造とともに移動することにより、前記画像データの運動および構造を3次的に解釈させることができるように、該骨格要素と該構造とを関連づける骨格挿入部と、前記入力データを前記骨格要素のうちの少なくとも1つおよび背景と関連して記録する記憶部と、を備える、装置が提供される。

20

**【0026】**

本発明の第4の局面によると、奥行き情報を用いて入力データを圧縮する圧縮装置であって、前記画像データ内において運動中の構造を特定する構造特定器と、前記構造特定器と関連づけられた骨格挿入部であって、3次元骨格要素が前記構造とともに移動することにより、前記画像データの運動および構造を3次的に解釈させることができるように、該骨格要素と該構造とを関連づける骨格挿入部と、前記入力データおよびその奥行き情報を提供するように、該入力データを前記骨格要素のうちの少なくとも1つおよび背景と関連して出力する圧縮部と、を備える、装置が提供される。

30

**【0027】**

本発明の第5の局面によると、奥行き情報を用いて入力データを記録する記録方法であって、前記画像データ内において運動中の構造を特定する工程と、3次元骨格要素が前記構造とともに移動することにより、前記画像データの運動および構造を3次的に解釈させることができるように、該骨格要素と該構造とを関連づける工程と、前記入力データを前記骨格要素のうちの少なくとも1つおよび背景と関連して記録する工程と、を包含する、方法が提供される。

40

**【0028】**

本発明の第6の局面によると、奥行き情報を用いて入力データを圧縮する圧縮方法であって、前記画像データ内において運動中の構造を特定する工程と、3次元骨格要素が前記構造とともに移動することにより、前記画像データの運動および構造を3次的に解釈させることができるように、該骨格要素と該構造とを関連づける工程と、前記入力データを圧縮し、その奥行き情報を提供するように、該入力データを前記骨格

50

要素のうちの少なくとも1つおよび背景と関連して出力する工程と、  
を包含する、方法。

【0029】

特に別に定義しない限り、本明細書中において用いる全ての技術的および科学的用語は、本発明が属する分野の当業者に共通して理解されるものと同じ意味を有する。本明細書中において示される材料、方法、および例は、全て例示のみを目的とするものであり、限定を目的とするものではない。

【0030】

本発明の方法およびシステムの実施は、選択された特定のタスクまたは工程を手動でまたは自動的に実行または完成すること、あるいはそれらの組み合わせを含む。さらに、本発明の方法およびシステムの好適な実施形態の実際の器具および機器によると、選択されたいくつかの工程は、任意のファームウェアのオペレーティングシステム上のハードウェアもしくはソフトウェアあるいはそれらの組み合わせによって実施することができる。例えば、ハードウェアについては、本発明の選択された工程をチップまたは回路として実施することができる。ソフトウェアについては、本発明の選択された工程を、任意の適切なオペレーティングシステムを用いてコンピュータにより実行される複数のソフトウェア命令として実施することができる。いずれの場合においても、本発明の方法およびシステムの選択された工程は、複数の命令を実行する計算プラットフォーム等のデータプロセッサによって行なわれるものとして説明することが可能である。

【0031】

本明細書中において、本発明のほんの一例を添付の図面を参照して説明する。ここで特に図面を詳細に参照すると、提示した事項は一例であり、本発明の好適な実施形態の例示的な考察のみを目的とするものであり、かつ本発明の原理および概念的な側面の最も有効で容易に理解できると思われる説明を提供するために提示されていることを強調しておく。このため、本発明の基本的な理解を得るために必要である以上に詳細に本発明の構造内容を示してはならず、図面とともに説明を参照することにより、当業者には、いくつかの形態において本発明がどのように実施されるのかが明らかとなる。

【発明を実施するための最良の形態】

【0032】

本発明の各実施形態は、時間に基づく実写画像のシーケンスを空間に基づく3次元(3D)モデルに変換し、3D実画像アニメーション、新たな時間に基づくシーケンス、画像処理操作、2D/3Dモーションキャプチャ等のリアルタイムおよび非リアルタイム用途を可能にする方法および装置を含む。

【0033】

本発明の各実施形態は、2次元または部分的に3次元であるデータ内の構造を特定し、それらと3次元骨格または骨格要素とを関連づける。これらの骨格または骨格要素を元のデータとは別のレベルで適用することにより、それらのレベルを相互に投影して、画像データに正確な奥行き情報を与えることができるようにしてもよい。

【0034】

本発明による方法および装置の原理および動作は、図面および付随する説明を参照することにより理解できる。

【0035】

本発明の少なくとも1つの実施形態を詳細に説明する前に、本発明が、下記の説明に示したまたは図面に例示した構成要素の構造の詳細および構成に限定されるものではないことを理解されたい。本発明は、他の実施形態が可能であり、各種の方法で実践または実施することが可能である。また、本明細書中において用いる語句および用語は説明目的のものであり、限定を目的とするものではないことを理解されたい。

【0036】

ここで、画像データに3次元解釈を与える装置を示す簡略ブロック図である図3を参照する。この画像データは2次元または部分的に3次元である情報であり、上記解釈とは

10

20

30

40

50

、3次元構造および3次元的な運動の統一的解釈である。

【0037】

図3の装置は、画像データ内の構造を特定する構造特定器302を含む。下記でより詳細に考察するように、これらの構造は、人口知能を用いて自動的に特定するか、ユーザの入力により特定してもよく、またはそれらの組み合わせにより特定されてもよい。

【0038】

この装置は、上記固定構造特定器と関連づけられ、3次元骨格要素を画像データ内で特定された構造と関連づけるかまたはそれらに貼り付ける骨格挿入部304をさらに含む。骨格要素は、ブロック、チューブ、球体、楕円または任意の他の要素的もしくはより複雑な3次元幾何学物体であってもよく、これらの要素は、それら自身を接合し、相互に貼り合わせる能力を有する。この要素の3次元形状を上記のように特定した構造に持たせることで、骨格要素が上記構造と移動するか、あるいは混在することにより、当該構造の3次元解釈を与えることができる。

10

【0039】

すなわち、この骨格要素は既知の3次元構造を有しており、これは、X、YおよびZ次元に延在することを意味する。この構造の移動は、XおよびY次元で見ることができ、Z次元における構造の挙動の詳細は、骨格要素との関連から推測することができる。

【0040】

構造特定器は、個々の構造だけでなく、相互に関連する構造から構成された複雑な物体も認識できることが好ましく、「相互に関係する」とは、それら構造間の運動関係が定義されていることを意味する。一例としては、前腕および上腕等の構造からなる人体が挙げられる。前腕は、上腕の端部を中心にして定義された方法で動くが、これは、本発明の各実施形態の骨格要素によってモデル化することができる。

20

【0041】

そのような複雑な物体が画像データから認識された場合、骨格挿入部は、その複雑な物体について骨格要素間の運動関係が定義されている同様に複雑な骨格を構築しようと試みる。下記で説明するように、これを達成する1つの方法は、人体のような複雑な物体を認識し、必要な要素を有し、関係が予めプログラムされた骨格を予め設定することである。

【0042】

このような予め記憶または設定した骨格を用いることで、構造および運動の両方を含む複雑な物体の3次元側面を理解することが可能である。すなわち、画像内の3次元構造および運動は、特定された物体の演繹的知識から理解することができる。さらに、システム内において、骨格に基づいてそのオブジェクトの奥行き情報が分かる場合、画像の3次元処理に関する処理負荷を大きく低減することができる。

30

【0043】

上記装置は、元の画像データ内の相対運動を分析して骨格挿入部304に運動関係に関する定義を与えることができる移動分析部306をさらに含んでもよい。この移動分析器は、画像を構成するピクセル集合内の構造を認識し、公知のトラッキング技術を用いてピクセルグループ内での運動を特定することができる。

【0044】

骨格記憶部308は、認識された複雑な物体とともに使用するために予め設定された骨格を記憶する。この記憶部は、例えば、画像データ内で人間が認識されるたびに用いられる、予め設定された人間の骨格を記憶してもよい。

40

【0045】

構造が予め設定された骨格を有しないと認識されると仮定する。この場合、骨格挿入部は、幾何学的要素を挿入することにより最初から骨格を形成しようと試みる。しかしながら、これらの幾何学的要素は、適合させる前に回転および変形することが必要な場合がある。それゆえ、画像データに適合するまで、選択した要素を回転させることが可能な回転部310、および様々な方法でその要素を変形させてデータに適合させることが可能な変形部312が提供される。これら回転および変形部は、ユーザによる入力を介して動作し

50

ても、自動的に動作してもよい。

【0046】

骨格を適合させると、画像内の構造は3次元モデルとしてモデル化される。トラッキング部314は、最初の画像データ内の運動をトラッキングし、画像と共に骨格を移動させることができるため、その運動の3次元情報を得ることができる。骨格と画像データ間の投影処理を実行することができ、それによって、単一のカメラから3次元の運動情報を得ることができる。

【0047】

アニメーション部316は、骨格を介して運動を適用し、一度モデル化した図形または他のオブジェクトをアニメーションにすることができる。

10

【0048】

上記装置は、アプリケーションによっては、必ずしもトラッキング部およびアニメーション部の両方を必要としないことは言うまでもない。アニメーションアプリケーションは、通常、アニメーション部を有するが、トラッキング部はなくてもよく、一方で、ビデオキャプチャアプリケーションは、トラッキング部を有するが、アニメーション部はなくてもよい。

【0049】

描画部318は、トラッキング部およびアニメーション部のいずれかまたは両方に接続され、モデル化されている場面を要求された方向から観察するために描画する。すなわち、3Dデータを有することには、モデル化したオブジェクト等を、最初に画像を撮影した角度だけではなく、あらゆる角度から観察することが可能であるという利点がある。下記においてより詳細に説明するように、描画部は、3次元モデルを要求された観察方向で平面に投影してテクスチャ等を適用することを必要とするだけであり、所与の方向から場면을観察することが可能である。

20

【0050】

次に、本発明の好適な実施形態による動きデータを含む3次元モデルを得る処理を示す簡略図である図4を参照する。

【0051】

画像データを段階402において獲得するが、このデータは2Dデータであっても、部分的にもしくは完全に3Dデータであってもよい。データ内の要素を特定する。段階406において、骨格要素を、特定した構造要素と関連づけて挿入する。次いで、段階408において、特定した構造要素を適合させるために、骨格要素を回転、変換、サイズ調整する。変換は変形を含む。次いで、段階410において、利用可能な情報に応じて骨格要素間の移動関係を定義する。

30

【0052】

ここで、人体等の複雑な物体が認識される場合について、図4のフローチャートの変形である図5を参照する。段階502において、初期データを再度取得する。上記複雑な物体をこの初期データから特定する。段階504において、適切な骨格をデータ記憶部から取り出し、段階506において、複雑な物体と関連づけて挿入する。次いで、段階508において、骨格を回転、変換またはサイズ調整する。変換は変形を含む。結果として、特定した構造と適合が必要な骨格要素とが適合する。上記図4のように骨格要素を複雑な物体に適合させるといふ試みにより、人体等の複雑な物体を特定し、複雑な骨格を適切に選択できることに留意されたい。

40

【0053】

ここで、図4および図5の処理をモデル化されているオブジェクトの運動にまで拡張する図6を参照する。段階602~608は、前述したとおりである。段階612は、3Dオブジェクトをモデル化する骨格を伴うため、オブジェクトの運動が骨格上に投影され、かつ/または骨格の運動が2D画像データ上に投影される。次いで、段階614では、任意の所望方向からの描画のために画像データが利用可能である。

【0054】

50

ここで、本発明の好適な実施形態において画像データがどのようにして層構造で管理されるかを示す簡略図である図7を参照する。2次元または部分的もしくは完全な3次元画像データは、第1のレイヤ702に記憶される。3次元骨格は、下側レイヤ706に記憶される。3次元骨格の2次元投影は、それらの間の仮想レイヤ704に存在する。見かけ上は2次元的な画像は、レイヤ702の元の2D画像とは異なる方向から観察することができるが、これは3次元骨格をその方向に投影することによって可能となる。この投影は、仮想レイヤ704内に記憶される。

**【0055】**

ここで、骨格要素の割り付けのために、初期データ内のオブジェクトおよび構造がどのように認識されるかを示す簡略図である図8を参照する。

**【0056】**

2つの経路を示しているが、1つ目は、ユーザが要素、物体および複雑な物体を装置に示す手動経路(段階802)である。代替的な方法として、段階804、806および808で構成される自動経路を構造を特定するために設けることができる。

**【0057】**

グループ分けおよび継続するか否かに関する決定は、単一の段階として見なすことができることは言うまでもない。一連の画像上の点またはピクセルを追跡し、共に移動する点を一まとめにする。このグループ分けの処理は、特定が確定するまで反復して繰り返される。

**【0058】**

これら2つの処理の組み合わせを用いてもよいことにも留意されたい。例えば、ユーザは、いくつかの要素または1つの複雑な物体をシステムに提示してもよく、システムは、他の要素を自動的に特定するか、またはその複雑な物体内の個々の要素を特定する。

**【0059】**

ここで、本発明の各実施形態を用いて運動を3次元的に推定する反復的性質を示す簡略図である図9を参照する。初期の2Dまたは部分的もしくは完全な3D画像内のピクセルをトラッキングする(段階902)。下側の骨格は、関連づけられた構造においてトラッキングされた点の動きに従って移動し(段階904)、段階906において、推定を行なって初期の画像内でのピクセルの3次元位置を判定する。段階904および906は同時に行なわれ、相互に影響を及ぼすため、図面内においては2本の矢印を用いて示していることに留意されたい。

**【0060】**

本発明の好適な実施形態では、コンピュータにより生成された時間に基づく写真シーケンスを用いて、3つの3Dモデルを構築する。

**【0061】**

初期データについての入力、例えば、既に記録された公知のビデオ形式のいずれかで圧縮されたビデオクリップから、またはUSBもしくは他の公知のデジタルまたはアナログ接続のいずれかを用いて直接接続された単一または複数のカメラから、実写画像またはビデオクリップを受信およびデジタル記録するモジュールによって行なわれてもよい。

**【0062】**

図4および図5を再度参照して、初期データは、例えば、1つ以上の時間に基づく実写画像のシーケンスから取得してもよい。公知のビデオまたはフィルムシーケンスにより、観察者の脳内において運動について時間的な錯覚を引き起こす。

**【0063】**

入力されたデータは分析される。この分析は、下記のアルゴリズムの項目で説明するように、入力シーケンスの奥行き地図の構築、時間に基づく各シーケンスの奥行き地図の作成、および奥行き地図の処理を伴うことが好ましい。

**【0064】**

好適な実施形態では、本方法は、最終的に、実写画像530のシーケンスによってキャプチャされたオブジェクトの3Dモデルを生成する。

10

20

30

40

50

## 【0065】

これらのモデルは、実際の画像または図形クリップ等から再構築されるが、この時、時間次元は、全ての図形および変化しない背景を3次元モデルとした空間次元に変換される。

## 【0066】

これらの3Dモデルは、以前はコンピュータにより生成された3D仮想世界でのみ可能であった多くの操作を可能にし得る。

## 【0067】

本発明の好適な実施形態では、個々の図形をシーケンス内で特定することができる。一度特定されると、これらは独立した3Dモデルに変換され得る。図形の移動は、シーケンス内の写真のタイミングと比較され、写真上での進行状態から図形の移動を対応させる根拠を提供することが可能である。さらに、各図形毎に時系列を個別に調整して、異なるイベントのシーケンスを提供することも可能である。このように、モデル化されている図形をモーフィングすることが可能である。

10

## 【0068】

例えば、2人の人が通りを渡るシーケンスにおいて、人物Aが人物Bよりも先に反対側に到達する場合、各図形を別々にモデル化しているため、個々の図形のタイミングを変更することが可能である。よって、人物Bが人物Aよりも先に通りを渡るように決定し、実写されたシーケンスの元の時系列を変更することにより、そのシーケンスをモーフィングしてもよい。

20

## 【0069】

上記装置により、ユーザは、生成された3D空間に基づくモデルに基づいて、媒体から異なる種類の出力を数種類作成することができる。ユーザは、作成された3D空間に基づくモデルを、外部画像処理、アニメーション、放送等の公知のプログラムのいずれかとともに用いるか、またはエディタ等の内部ツールを用いてもよい。そのようなエディタにより、ユーザは、主に、線形媒体および非線形媒体の2種類の出力を作成および編集することができる。線形媒体とは、時系列に基づく媒体を指し、特定の時間間隔で撮影された画像のシーケンスを意味する。ユーザは、自身が見せたいと思う時系列のイベントに基づいてクリップを作成することができる。さらに、それらの結果を、例えば、リアルタイムでのライブビデオ画像処理、ビデオクリップ、モーションキャプチャ、静止画像、DVD、球形画像、2D画像、立体画像またはあらゆる公知のフォーマットを含む種々の表示フォーマットで出力することが可能である。

30

## 【0070】

本発明の各実施形態の装置はまた、時系列されていない、すなわち、非線形媒体を作成することも可能である。そのような時系列化されていない出力としては、例えば、1組の画像、アニメーション、およびテキストを含む3D的な周囲環境が挙げられる。

## 【0071】

本発明の各実施形態の装置は、この出力を3次元仮想環境として表す能力を与えるが、例えば、そのような環境において、ユーザは、選択したあらゆる経路を完全に辿ったり、あらゆる地点に到達することができ、その地点の周囲360度を見渡したり、あらゆる図形とインタラクトすること等が可能である。このような非線形的な出力として、コンピュータゲーム、医療手術シミュレータ、飛行シミュレータ等多くの例が挙げられる。

40

## 【0072】

上記装置は、図3のアニメーション部316のようなアニメーションエディタを含んでもよい。アニメーションエディタ316は、ユーザが選択する全てのオブジェクトに生命を与えるツールであり、アニメーションエディタ316はまた、風に揺れる木または歩行する人物の姿などのオブジェクトに特定の動きを与え、さらに、その人物が歩く、または走る時にどのように振舞うか、怒ったまたは悲しい時に、どのような顔をし、どのように唇を動かすか等、固有の特徴を与える。アニメーションエディタはまた、外部ソースからのコンピュータアニメーションもしくはモーションキャプチャ、または上記装置のモーシ

50

ョンキャプチャツールを用いて、予め定義された1組の動きをオブジェクトに与えてもよく、また、例えば、足を少し引きずったりすることや、額の皺等、各オブジェクトを特徴づける1組の運動および特徴を定義することができる。これらの運動は、図形の個性を作る上で役立つ特徴である。アニメーションエディタはまた、上記装置のモーションキャプチャツールを用いてオブジェクトの音声的な特徴を作ることも可能にし、これによって、オブジェクトを喋らせることも可能である。

【0073】

このソフトウェアは、以下で説明および例示する方法およびアルゴリズムを用いることが好ましい。

【0074】

3D空間に基づくモデル作成のさらなる詳細

本発明の好適な実施形態の基本構成は、コンピュータにより生成された3Dアクシスエンジンであり、この3Dアクシスエンジンに対応する3つのベクトル、および下記で説明する時空間ベクトルを用いる。

【0075】

入力画像シーケンスSは、この構成に入力される画像のシーケンスである。

【0076】

好適な実施形態では、以下で示すようなアルゴリズムを実行してもよい。

【0077】

シーケンスSは、シーケンス内の $N_f(s)$ 個のフレーム、例えば、PALビデオ表示規格において25フレーム/秒となるように分割される。

【0078】

$S_0$  (第1シーケンス)の第1フレームは $s(0, 0)$ と示す(第1シーケンスの第2フレームは $s(0, 1)$ 、最終フレームは $s(0, n)$ とする( $N_f(S_0) = n + 1$ フレーム))。

【0079】

多数のアンカー点を用いる。これらのアンカー点は2つの主要な要素を有するが、一方の要素は $S_i$  (ただし、 $0 \leq i \leq N_f(S_i)$  シーケンス数)内の要素間の対応であり、他方の要素は、 $S_i$ と、Fで示す3Dアクシスエンジン間の対応である。

【0080】

入力、奥行き地図およびアンカー点

本発明の好適な実施形態によるアルゴリズムでは、入力として $S_0$ を受け取り、その中の全てのシーケンスフレームを用いて、シーケンスの奥行き地図を生成してもよい。係数Dは、(下記で定義する) $s(0, 0)$ の奥行きベクトルとして定義される。 $z$ がフレーム $s(0, 0)$ からの異なるピクセルの奥行き値の組であると仮定する。 $s(0, 0)$ では、 $d(0, 0)$ は、 $z_0$  (ただし、 $z_0 = \min(z)$ )に等しい奥行き値を有するフレームからの点の集合である。 $d(0, h)$ は、 $z_h$  (ただし、 $z_h = \max(z)$ )に等しい奥行きを有するフレームからの点の集合である。 $\{z_i = \min(z / z_i), i = 0, \dots, h-1\}$  (ただし、 $z / z_i$ は、 $z_0$ から $z[i-1]$ までの要素がない集合 $z$ を意味する)であるため、 $\{z_0, z_1, z_2, \dots, z_h\}$ は、フレーム $S(0, 0)$ ベクトル $D_0$ の奥行きレイヤの集合であり、この集合は、下から上の順序で分類され、この集合内の数は、さらに説明するように、Fの分析係数に対するベクトル $D_0$ のレイヤであることは定義から明らかである。例えば、 $D_0 = \{2, 5, 6, 9, 13, 56, 22, 89\}$ であれば、 $\{z_0 = 2, z_1 = 5, z_2 = 6, z_3 = 9, z_4 = 13, z_5 = 22, z_6 = 56, z_7 = 89\}$ である。 $D_0$ は、 $S(0, 0)$ の奥行きベクトルであり、 $D_0 = \{d(0, 0), d(0, 1), \dots, d(0, h)\}$ である。

【0081】

フレーム $S(0, i)$ のDは、3Dマトリックスである。フレーム $S(0, i)$ のDは、2Dマスクマトリックスのベクトルである。ピクセルの奥行き $d(i, j)$ が、なんらかの理由で定義されない場合、 $d(i, j) = \text{無限大}$ であり、ブーリアン2Dマトリック

10

20

30

40

50

スでは0と定義する。

【0082】

ただし、フレームからのピクセルの奥行き値  $d(i, j)$  を、フレーム  $S(0, i)$  からの2D画像のデータにおいては定義できない場合、上記アルゴリズムは、同じシーケンス  $S(S(0, i))$  を取るシーケンス)からの複数のフレームのデータを用いて定義しようと試みる。 $s(0, 0)$ の奥行き地図、またはその一部が、例えば、照明が不良であるために定義できない場合、 $SP$ は、 $d(0, 0)$ 、または定義した部分を、一時的に「 $\infty$ 」(無限大)として扱い、 $s(0, 1) \dots s(0, i) S_i$ (ただし、 $0 \leq i \leq S$ シーケンス数)を用いて、 $s(0, 0)$ を計算しようとする。

【0083】

フレーム  $s(S(0, i) | i = 0, 1, 2, \dots, n)$  ( $S$ (シーケンス番号、フレーム番号))の奥行き地図を計算する場合、 $D0 = \{d(0, 0), d(0, 1), \dots, d(0, h)\}$  ( $D$ (フレーム番号、奥行き  $enc$  点))である。ソフトウェアは、奥行き地図マトリックスの値の上限(奥行き値の集合の最大値)を探す。

【0084】

$SP$ はまた、奥行き地図マトリックスの値の下限を探す。シーケンス  $S0$ の奥行き地図の全てのフレーム処理が成功すると、 $SP$ は、定義されたあらゆる時点において、シーケンス  $S0$ のアンカー点の上限および下限を探す。

【0085】

$s(0, 0)$ では、 $d(0, 0)$ が、奥行き方向において最も近い点である。 $s(0, 0)$ において奥行き方向で最も深い点を  $d(0, h)$ で示す( $D0 = \{d(0, 0), d(0, 1), \dots, d(0, h)\}$ )。

【0086】

係数  $D$ は、アルゴリズム内の奥行きベクトルの集まりであり、いくつかの  $D$ ベクトルを用いて、画像の奥行き構造マップを関連づける加工ツールとしてデータを分析する。 $SP$ 構造マップは、 $F$ 内で展開されて静止した要素および移動する要素をモデル化し、要素の一部(例えば、手、足等)を表す新たなマトリックスを多数用いて  $F$ 内に作られる。 $D$ は、このベクトルに沿った各点が現在の奥行きで、対応する奥行き情報を含み、さらに、位相地図における高度線のように各点で切り取った奥行きの面をに沿って奥行き値を表すように構成される。

【0087】

$D$ は3Dマトリックスであり、 $D$ に沿った  $Z$ 点毎に2Dブーリアン画像マトリックス( $x, y$ )として構成され、各2D画像マトリックス( $x, y$ )において「1」を示し、この情報のみに対応する奥行き点( $Z$ )における画像内に含まれる。

【0088】

ビデオシーケンス  $S0$ のフレーム間の基準となるアンカー点を探すために、システムは以下のように進行する:

$\{x, y, z\} = \{\text{水平、垂直、奥行き}\} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ は、 $F$ からの空間ベクトルであり、これを「世界座標系」と呼ぶ。

【0089】

ここで、あるフレーム  $S(0, t)$  ( $t = \text{任意の数}$ )に着目すると、このフレームは、自身の局所座標系{列の指数、行の指数、奥行き}を有する。

【0090】

このフレーム上の  $k$ 番目のアンカー点について考える。フレームの局所座標系における位置は、( $u_{t,k} = i, v_{t,k} = j, \text{奥行き}(i, j)$ )である。

【0091】

このフレーム局所座標系の範囲(基礎)は、世界座標系では、 $\{i_t, j_t, k_t\}$ である(ただし、 $i_t, j_t, k_t = (i_t \times j_t)$ は、世界座標系のベクトルである)。

【0092】

フレーム局所座標系では、 $i_t$ は  $(1, 0, 0)$ であり、 $j_t$ は  $(0, 1, 0)$ であり

10

20

30

40

50

、 $k_t = (0, 0, 1)$ である。

【0093】

ビデオシーケンス $S_0$ のフレーム間の基準となるアンカー点を見つけるために、システムは、各フレーム( $S(0, 0)$ 、 $S(0, 1)$ ... $S(0, n)$ )を、フィールド $F$ より上のベクトル空間 $S_0$ の(ベクトル空間自体である)部分空間として扱う。システムは、ベクトルの基礎 $W_0 = S(0, 0)$ を計算する(ただし、 $W_0$ の範囲( $W_0 = \text{Sp}\{w_1, w_2, \dots, w_m\}$ )は、部分空間 $S(0, 0)$ 上に広がる。各部分空間毎に異なる基礎がいくつか存在し得るが、各基数の範囲は $W_0$ 上を広がり、ベクトルの基礎ベクトルの役割は、数学の分野で公知のものと同様である。これらのベクトルは、部分空間 $W(0, 0) = S(0, 0)$ を作成する。

10

【0094】

奥行き的位置合わせ

画像の背景等の固定されたオブジェクトの奥行き的位置合わせは、2段階で行なわれる。第1の段階では、システムは、ベクトルの基礎 $W(0, 0)$ を見つける(ただし、 $\text{Sp}(w_1, \dots, w_m) = W(0, 0) = S(0, 0)$ )。位置合わせを速く行なうために、システムは、ベクトルの基礎 $W(0, 0)$ から水平、垂直、奥行き、空間/時間ベクトルの4つの基準ベクトルを作成する。

【0095】

第1のベクトル $Z = S(0, 0)$ は、各点の基礎ベクトル数 $d_0 = S(0, 0)$ を反映しており、フレーム内の基礎ベクトルの奥行き情報を表す $Z$ ベクトルを作成する。 $Z$ の中点を $d(\text{anc})$ とも表し、これはフレーム自体の中点である。

20

【0096】

$$d(\text{anc}_1) = d(0, 0) + d(0, m) / 2$$

$d(\text{anc})$ は、システムが一時的に0点軸 $XYZ = F$ として示す点とすることができる。水平および垂直ベクトルは、 $Z$ ベクトルに沿った画像マトリックスの水平および垂直方向の各点のベクトルを表す。第4の基準ベクトルは、時間次元から空間次元への変換ベクトルとして用いられる空間/時間ベクトルである。この時点で、システムは、 $S(0, 0)$ と $S(0, 1)$ の間で用いる位置合わせのための3つの基準ベクトルを作成しており、これらを $D'$ として統合している。

30

【0097】

フレーム間の誤差は、例えば、照明係数、フレーム内を移動する要素、およびカメラの動向(例えば、寄せ/引き、左寄せ/右寄せ、上/下方向移動、上/下方向への傾け、左/右パンおよびズームが挙げられる(光学またはデジタル写真に関しては、インチ毎のピクセル量に誤差が見られ、デジタルズームではより低くなる))。フレーム間での誤差の変化は、大抵の場合、ピクセルの位置として見られるため、フレーム間でのなんらかのピクセルの移動が見られ、 $SP_1$ は、 $S_0$ のフレームの3つの基準ベクトルを空間/時間ベクトルの関数として計算する。3つの対応するベクトルは、画像の3Dに関する位置合わせのために構築され、垂直および水平ベクトルは空間窓( $X, Y$ )に対応し、 $Z$ ベクトルは奥行きベクトルに対応する。

40

【0098】

空間( $X, Y$ )ベクトル内の各係数は、画像の各点における $Z$ ベクトルに沿った空間領域内の画像の基礎ベクトルを反映する。

【0099】

マッチング関数は、誤差がゼロか、またはできるだけ多くの点において誤差をゼロとすることを目的として、最大でも画像1のベクトル $V(h_0/v_0/z_0)$ と画像2のベクトル $V(h_1/v_1/z_1)$ の間の所定の最小点とする傾向がある。ベクトルの統合された部分の位置合わせに関して、両方のベクトルのそれぞれの両端では、フレーム間における相違点は一致しないことがある。これらの相違点は、新たなフレームに加えてもよいが、先行するフレームには現れない異なる情報を参照する場合がある。

50

## 【 0 1 0 0 】

$$\begin{matrix} V' 0 & V 0 & V 0 & V 1 \\ V' 1 & V 1 & V 0 & V 1 \end{matrix}$$

これら3つのベクトルは、画像の3次元位置情報の結果であり、視覚情報情報とは関係なく、各点における画像の基礎ベクトルを表す。

## 【 0 1 0 1 】

好ましくは、水平、垂直、および奥行きベクトルは、できる限り多くの点において最小の誤差をみつけるために、個別に各ベクトルと比較される。

## 【 0 1 0 2 】

フレーム間の相違点は一致しない場合がある。これらの不一致は、あるフレームにおいて、先行するフレームでは見られない異なる情報が現れることを示すことがある。

## 【 0 1 0 3 】

変形およびカメラの動きへの対処

カメラのレンズ等の光学要素は、実写画像に変形を生じさせ、同じオブジェクトが変形することにより奥行き地図内に微細な誤差を生じることがある。

## 【 0 1 0 4 】

ここで、2つの画像を示し、カメラが動くことによって生じる変形について例示する図14を参照する。

## 【 0 1 0 5 】

第1のフレーム1401では、石柱1403がフレームの中心に見られる。第2のフレーム1405では、同じ柱1403がフレームの右側に位置する。ここでは、光学処理の結果、なんらかの変形が生じ、奥行き地図は、これらの変形による微細な誤差を生じる場合がある。これによって、位置合わせされた画像に対する最終的な結果としてなんらかの途切れが生じるが、この途切れはこれらの変形によるものである。換言すれば、ブロック構造が、フレームの中心から横に移動する際に光学的な理由により変化するように見える。

## 【 0 1 0 6 】

解決策としては、所与のオブジェクトが比較的フレームの中央に位置するものをそのオブジェクトの最良の画像として特定することである。ピクセルは、この最も正確な測定により得られた3D位置を受け取る。

## 【 0 1 0 7 】

Zベクトルを位置合わせするときには、画像間のズーム係数（光学またはデジタル）に誤差を生じることもある（すなわち、カメラの「寄せ」または「引き」による）。

## 【 0 1 0 8 】

フレーム間のカメラの移動が少なくなるほど、ベクトル間の相関度が高くなり、より良好な結果が得られる。画像間の誤差が大きくなるほど、ベクトル間の相関度が低くなる。SPは、そのずれに対して閾値を設定し、より大きな誤差を有する場所を異なるオブジェクトに関するものと見なす。

## 【 0 1 0 9 】

水平、垂直および奥行きベクトルの統合部分の位置合わせ後：

$$\begin{matrix} V' h/v(0) & V h/v(0) & V h/v(0) & V h/v(1), \\ V' h/v(1) & V h/v(1) & V h/v(0) & V h/v(1) \\ V' z0 & V z0 & V z0 & V z1 \\ V' z1 & V z1 & V z0 & V z1 \end{matrix}$$

## 【 0 1 1 0 】

この時には、統合部分を部分空間として扱い、ベクトルをこの部分空間基準として再計算する。Fフィールドに関して「スカラー」係数を用いてベクトルを計算することによりズーム係数を考慮するが、スカラーによってベクトルを乗算するかまたは除算することにより、カメラのズーム/寄せまたは引きを模倣し、フレーム内の要素間の関係は同じに保つが、分解能が異なる場合がある。スカラーの力を借りることにより、ベクトルS(0,

10

20

30

40

50

0) および  $S(0, 1)$  を位置合わせすることができる。この処理によって、画像の位置合わせをするとともに、次のフレームの位置合わせ方向を示すこともできる。空間/時間ベクトルは、時間領域から空間領域への変換に関連し、ここでは、新たな位置合わせを統合フレーム F として見なし、この先行する統合フレームと次のフレームとを位置合わせする。これはまた、特に複数のフレームがすでに位置合わせされたエリアを繰り返す場合に計算量を低減することができる。空間/時間ベクトルは、時間次元から空間次元への変換のための基準ベクトルである。

#### 【0111】

奥行き情報の不一致への対処

奥行き情報に不一致が生じた場合、上記装置は、F 1 で示すような新たなベクトル平面を開くことが好ましい。この新たな平面は、このアルゴリズムの処理が最初から開始される空の XYZ 座標系である。

10

#### 【0112】

そのような処理の最後に、システムのユーザは、F 0 および F 1 を異なる位置のままにするか、またはそれらの位置合わせを行ないたいかを問われる。次いで、ユーザは、それら 2 つの構造を手動で位置合わせするために回転、ズーム、反転等のツールを用いて、F 0 および F 1 を手動で位置合わせしたいか問われる（あるいは、ユーザは、システムに F 0 および F 1 を自動で位置合わせするように指示してもよい）。

#### 【0113】

ユーザが F 0 および F 1 を手動で位置合わせした後、システムはこの位置合わせを計算するように指示し、システムは、位置合わせアルゴリズムを用いてフィールドを位置合わせしようと試みる。フィールドが上手く位置合わせされると、システムはその旨を伝えるが、そうでなければ、SP は、ユーザに、位置合わせずれ係数の基準をより低く設定（より精度の低い位置合わせ）するように問う。システムは、さらに、ユーザに、公知の画像処理ツールを用いて画像平面内の途切れを直すためのツールボックスを用意する。

20

#### 【0114】

SP 分解能

システム 1 は、「F 0」（XYZ 軸のフィールド）サインの一時的な分解能を「R 0」と定義する。R は、インチ毎の基準点の数によって定義される。この分解能は、時間次元におけるピクセルに関する画像分解能の組み合わせ係数、空間次元における奥行き点の組み合わせ係数の結果である。分解能ツールは、例えば、異なる距離から同じ位置を撮影した 2 つのビデオクリップの位置合わせに役立つ。

30

#### 【0115】

例えば、テーブルは、テーブルの部品間（例えば、1 本の脚から次の脚まで）により多くの基準点が存在し得る高分解能クリップで撮影されるか、またはより分解能の低い第 2 のクリップと比較してより近い位置から撮影されるか、あるいはデジタルズームを用いるかもしくは基準点の数が少なくなるより離れた距離から撮影される。

#### 【0116】

分解能の問題に対処するための基準点は、現実の位置に対する各ピクセルの 3D 位置である。よって、空間次元におけるピクセルの 3D 位置は、時間次元からの変換後の計算された位置である。この分解能により、D 0 と S 0 とを対応させることができる。D 0 = d 0 (anc) の中点は、点 (0, 0, 0) (XYZ) F 0 において軸フィールドの中心に一時的に置かれる。

40

#### 【0117】

基準点の視覚情報は、下記でさらに説明するように、F 0 において、情報の視覚レイヤとしてレイヤ化される。

#### 【0118】

移動する要素の特定および再構築

移動する要素が画像に現れる場合。骨格は、移動する図形要素からなり、移動する要素の非常に正確な 3D 形状モデルを構築し、そのモーションキャプチャを保存し、実写され

50

た視覚情報をその3Dモデルに貼りつけるように、それらの相対的位置および移動パターンを定義する。システムは、特定および再構築処理の自動化を可能にする。

【0119】

まず、システムは、データ画像シーケンス内に移動する要素を有することを学習する必要がある。次の段階は、予め分類されたか、またはユーザによって定義された骨格要素を用いてデータ画像内の要素を特定することである。最後に、システムは、所定の3D構造および骨格を用いて要素の3D構造の再構築を実行するか、またはシステムは、徐々に作られた独立した新たな3D構造を要素の特徴に基づいて作成する。

【0120】

カメラに対して背景とは異なる情報を付加する移動要素は、風に揺れる木またはフレームを横切った後向きを変えて、フレームの反対側から出て行く人物等、時間の経過につれて僅かな情報を付加する半固定オブジェクトとすることができる。

10

【0121】

上述したように、システムは、まず、シーケンス内に移動するオブジェクトを有することを学習する。次に、システムは、1組の所定の3D要素または骨格を用いてこのオブジェクトを特定する。あるいは、ユーザが、骨格または要素を定義してその図形に貼りつけてもよい。

【0122】

次いで、システムは、所定の3D要素もしくは骨格、または新たなユーザが定義した要素を用いて、その図形の3D構造を構築する。

20

【0123】

フレーム内に移動する要素が存在することを特定するために、システムは、空間および時間に関して、シーケンス内に奥行きピクセルの途切れを探す。すなわち、空間次元におけるカメラおよび背景に対するS0の固定点との一貫性はないが、その情報を時間次元に関して変更するある特定の3D構造がS0に存在することがある。

【0124】

換言すれば、時間の経過につれて空間的な位置ずれが存在する。例えば、右側に移動するカメラを用いてテーブルを撮影する場合、テーブルは、まず、フレームの右側に現れ、次いで、フレームの左側に向かって移動する。

【0125】

情報が時間次元に関して変化するテーブルの前に3D要素が存在する場合、システムは、フレーム内に移動するオブジェクトが存在すると結論づけることがある。システムは、その移動する要素の3Dモデルを再構築するが、ここでは、テーブルは静止した要素である。

30

【0126】

次元に基づく画像処理における3D要素の生成のために、マッチングベクトルを、着目する要素の周囲の領域から構築することができる。

【0127】

$N \times M \times Z$ の大きさの矩形ウィンドーを選択することができるため、3Dマトリックスとすることができる。NおよびMは、ウィンドーの空間サイズであり、Zは奥行き寸法である。第4のベクトルを与えて、要素またはオブジェクト次元の時間次元から空間次元への変換を定義し、3D要素および図形の構築に導くことができる。時間および空間次元の両方において要素をマッチングすることにより、要素またはオブジェクトごとに個別に作る事が可能な対応する3D構造マップに沿った全ての点について一貫性のあるマッチングが可能である。

40

【0128】

本発明の各実施形態のアルゴリズムは、3D情報構造の2D画像上への投影に基づいて、背景および完全な3D環境に関してフレーム内の移動する要素のトラッキングに役立つ。

【0129】

50

### 要素の特定

画像データから現在の要素を特定することは、1組の予め分類された3D構造を活用して実行される。システムは、当該要素またはその一部の形式を決定することから全体構造の特定までの工程を行ない、また、ユーザが新たな構造を構築する際に役立つ。

#### 【0130】

システムは、球状、箱型、パイプ状等の単純な3D形状3Dモデルから始まり固定および非固定物体の完全な骨格に至るまでの骨格要素に関する1組の3D構造を含むデータベースを備えていてもよい。骨格は、例えば、車から、動物や図12に示すような人間の骨格に至るまでを特定し、モデル化するための矩形領域とすることができる。

#### 【0131】

骨格は、3つの重要な要素を含む複雑な3Dデータ構造である：

1. 骨格の物理的集合（すなわち、構成する骨格要素の形状および相関関係）
2. 入力された3D情報に従った骨格の形状情報
3. 物体の物理構造（骨、筋肉、関節等）および物体の物理的動向等の内部情報の組み込み

#### 【0132】

上記3つの側面は、本発明の各実施形態のアルゴリズムによる特定および再構築処理に必要である。

#### 【0133】

1. 骨格の集合とは、骨格の構造形態を取り、その部品からシステムが骨格要素を理解し、当該物体または新たな物体をユーザの要求により組み立てることができるように物体の部品の最小単位まで定義するものを意味する。

#### 【0134】

例えば、人間の腕は、3D的な円筒形状に基づいており、関節を介して、手を表し得る別の円筒と接続され得る。他の例では、頭部は、3D的な球形の単純な図形を元にして、首を表し得る関節に接続され得る。この首は、次に、胴体を表し得る大きな円筒に接続される。骨格の部品の異なる物理的動向および人間、動物等の個々の要素の物理的動向を組み込んで、当該基本構成を再構築することにより、システムが図形を特定、再構築する際に役立つ。

#### 【0135】

2. 入力された3D入力に従って骨格を形作る能力は、その一部が変形する骨格を示す図13に関して後述するように、特定処理および再構築処理において用いられる。

#### 【0136】

3. 物体の物理的構造（骨、筋肉、関節等）およびその物理的動向等の内部情報は、以下で説明するように、特定処理および再構築処理において用いられる。

#### 【0137】

3D構造および骨格の組を用いて、システムは要素の特定を行なう。この処理は、自動的に行なうか、または図8に示すようにユーザによって手動で行なうことができ、当該要素をシステムに対して特定し、内部骨格を図形に貼りつけるか、または新たな構造を作ることを伴う。

#### 【0138】

移動するオブジェクトがフレームのシーケンスに位置するとき、システムは、それを特定して、以前に定義された骨格要素の組から選択された1組のマッチング用骨格要素、または好ましくはシステムにより与えられた1組のツールを用いてユーザが移動するオブジェクトについて定義した特定の骨格に貼りつけようと試みる。好ましくは、貼りつけられた骨格要素は、自動的に、移動するオブジェクトのサイズ、形状および移動パターンに合わせて調整され、移動するオブジェクトをサイズ、形状および移動パターンに関して適合させる。システムは、適切に重ねられたテクスチャを用いて骨格要素の組を完成する。

#### 【0139】

本発明の好適な実施形態では、システムは、移動するオブジェクトを、所望する任意の

10

20

30

40

50

視点の2D平面上において推測するツールをさらに提供する。

【0140】

3D構造に基づく位置合わせの性質を利用することにより、異なるビデオシーケンス間、異なる縮尺(ズーム)、異なる検知方法(IRおよび可視光線カメラ)におけるマッチング等、様々な状況において情報をマッチングすることができる。

【0141】

移動するオブジェクトから要素の3D構造を生成することは、システムが異なるシーケンス間の外見上の誤差を処理する際に役立つ基本的要素である。

【0142】

1つの要素は、円筒で構成された基本骨格に貼りつけ、関節は腕、足および体に貼りつけ、球状の要素は頭部に貼りつけて、奥行き的位置合わせにより、3D要素構造の生成および3D図形の基本骨格の物理的動向、腕、体、足に対する円筒の長さおよび厚さ、頭部に貼りつけた球状要素のサイズ等の相互関連づけに新たな情報を加えてもよい。

10

【0143】

これらの個別の円筒および球状要素から完全な3D図形を構築することにより、それらの相互の動向(それらがどのように相互に貼りつけられるか、または移動するか)が表される。この段階では、システムは、どのような種類の要素に直面するかを判定したり、その要素が何であるかを判定できないと決定し、ユーザに当該図形を判定する手助けを求めるか、またはユーザが所定の図形の基本組にはない新たな構造を形成してもよい。

【0144】

20

上述したように、ソフトウェアは、移動するオブジェクトの構造を、その奥行き情報を用いてできる限り特定しようとする。以前に定義された3D要素の組を活用して、ソフトウェアは、視覚情報に存在しないものがあつたとしても、オブジェクト部品の形態を1つずつ判定して完全な構造を完成する。

【0145】

第1の工程は、オブジェクトを特定し、その基本形態を判定することである。次いで、システムは、それをできる限り完成しようとする。

【0146】

その基本形態を用いて、空間領域および奥行き領域における要素について学習することが可能である。SPは、1組の3D骨格要素(球状、箱型、パイプ等)を用いて、オブジェクトの詳細を再構築しようとする。

30

【0147】

単一のカメラを用いた移動する要素の3D構造マップ

システムは、1つの画像の完全な奥行き3Dマップを受け取ってもよい。移動する要素を含む画像の奥行き地図を構築するための公知のアルゴリズムが存在する。例えば、空間時間立体アルゴリズムを用いた奥行き構造マップは、少なくとも2つのカメラを利用する。

【0148】

例えば、単一のカメラで空間時間立体アルゴリズムを用いて静止した背景画像の奥行き地図を構築する公知のアルゴリズムが存在する。

40

【0149】

1つのカメラを用いて、また奥行き地図を推測することなくビデオシーケンスから静止したモデルを生成する公知のアルゴリズムが存在する。

【0150】

本アルゴリズムは、上述したように、単一のカメラを用いて移動する要素の奥行き地図を推測するために用いることができる。

【0151】

移動する要素の奥行き値の構築のために、システムは、公知のアルゴリズムを用いて静止した固定背景の予め獲得した奥行き値を用いてもよく、基準点を用いて、移動する要素をその背景に対して独立した4Dマトリックスとして参照してもよい。

50

## 【 0 1 5 2 】

2 D 画像平面上に 3 D 情報構造（予め作成された 3 D 骨格等）を投影することにより、奥行き軸に関する各フレームにおいて、移動する要素のトラッキングに役立つ。3 D 空間への 2 D 画像平面の投影に加えて、要素の奥行き地図を生成する能力が与えられる。骨格および器官の貼りつけにより、画像を奥行き地図に適合させ、合成により任意のオブジェクトをコピーし、その運動をキャプチャする。後者の処理は、要素のテクスチャを再構築された骨格上に重ね、下記においてさらに説明する再構築処理を完成することをさらに伴ってもよい。

## 【 0 1 5 3 】

このようにして、本方法は、フレーム内で移動する要素の 3 D マップの生成を行なう。

10

## 【 0 1 5 4 】

奥行き推測の第 1 の工程は、各フレームの各ピクセルの 2 D 位置をトラッキングし、ピクセルごとに軌跡を作成することである。

## 【 0 1 5 5 】

このトラッキングは、公知のトラッキングアルゴリズムを用いて行なわれる。

## 【 0 1 5 6 】

前述したように、2 つの画像上において時間次元に沿った同じピクセルを見つけるための受動的方法は、ピクセルの色、階調、輝度、および陰影の調整を用いて、2 つのフレームにおいて時間軸に沿った同じピクセルを突き止める。

## 【 0 1 5 7 】

本トラッキングアルゴリズムは、時間における 2 D 画像平面では、特定に役立つ投影パターンを有さないため、時間の経過につれて誤りを収集する傾向にある。

20

## 【 0 1 5 8 】

上記の見解は、上述した奥行き推測アルゴリズムを実行して、1 つのカメラから、情報（現れる体の後ろ側にある手、照明の変化等）が隠れている、移動する 3 D 要素の映画クリップにおいて同じピクチャまたはピクセルをトラッキングしようとする場合特に正しい。背景と一致しないピクセルの独立した動きにより、1 つのフレームから次の数フレームまで誤ったトラッキング点を見つける等の問題が生じる。さらに、1 つのカメラを単一の角度で用いることは、カメラが同時に起こる光景を完全に捉えない点で不利である。人物を撮像する際には、人物は、通常、2 本の足、および 2 つの手を有する。しかしながら、2 D に基づくトラッキング技術では、いくつかのフレームにおいてカメラの角度から隠れており、連続性がなく突然（再度）現れる 1 人の人物の足または手を識別できないことが多い。（ ）よって、2 D におけるトラッキングは複雑で困難となり、そのような 2 D トラッキングから奥行き値を推測しても、実際の奥行き地図を得ることはできない。

30

## 【 0 1 5 9 】

しかしながら、3 D 骨格を 2 D 画像平面上に投影されたデータ構造として用いる場合、各フレームは、3 D データ構造の 2 D 投影である一方、例えば、手、左足、右足等の体の器官が特定された 2 D データ構造の 3 D 投影となる。

## 【 0 1 6 0 】

結果として、トラッキングの間違いが発生しなくなる。骨格の 3 D 姿勢は時間軸に沿って 2 D 平面上に投影されるため、3 D 投影から各フレーム内に新たなトラッキング点が無限に存在する正確なトラッキングを行いことで、システムは、3 D 物体の隠れた部分がどの部分であるのか、それらが 3 D 空間のどこにあるのかを知る。システムは、それらの要求された部分が元の 2 D 画像において現時点で見られない場合でも、3 D 部品を要求された任意の 2 D 画像平面上に単に投影してもよい。

40

## 【 0 1 6 1 】

所定の 3 D 骨格が、2 D 画像平面上に投影される。システムは、図 7 に関して上述したように、影のような画像を情報のさらなるレイヤに効果的に作成する。この情報のさらなるレイヤは、トラッキングする必要がある画像の一部を直接探しあて、すぐにそれらの誤りを切り取り、それらが広がることを妨げる。このような段階により、固定および非固定

50

要素の両方を含む歩行者等、移動する要素の奥行きトラッキングおよび推測が可能になる。3D骨格は、次いで、すでにトラッキングされた運動中の2D要素の3D奥行き地図を推測するために用いられてもよい。3D骨格上に位置する無限の点を用いて、基準点を用いた2D画像内の時間に沿ったピクセルの軌跡を生じることが、上記で参照した4Dマトリックスが2D空間に対する3D空間またはその逆の投影点として要素を囲むことを可能にする。このようにして、システムは、時間に沿った三角測量法および3D骨格データ構造を用いた点の3Dトラッキングを用いて、必要とされる正確な超分解能の生成を各フレーム内を移動する要素の奥行き地図情報により行なうことが可能である。一例として、ワークフローは、以下のように進行する：

Mが2D(X, Y)ピクセルのR・G・Bマトリックスであるとする、  
 n - フレーム数、  
 A - 背景に対して移動する要素、  
 B - 影のようなレイヤ、グレースケール「影」図形の2Dマトリックス、  
 Q - 閾値が定義された各フレームの特徴点、  
 T - 軌跡(Qの2D点位置ベクトル)、  
 - Qに関するTの推移関数、  
 K - Qが有するフレーム数(=Tの長さ)、  
 Z - 3D推測。

10

## 【0162】

よって、入力は、n個のフレームの「M」である。システムは、本明細書中の各所で説明しているように、移動する要素を特定し、次いで、位置合わせ処理F(b) G(A)、Aに関する3D骨格の2D投影Bが行なわれる。位置合わせは、変更可能な初期定義された閾値のためのものである。

20

## 【0163】

処理は継続され、Qを検索およびトラッキングすることにより、各QのTを作成するが、ここで、 $f(a, b)_i$ であり、 $q_{i+1}$ は軌跡ベクトルの関数である。トラッキングにより、画像aおよび影b上においてフレームi内の特徴 $Q_j$ の位置を突き止め、(k個の)フレームに関して、フレームi+1等においてその位置を付加する。I+1内の新たな点を画像Bに貼り付け、このフレームi+1に関する新たな情報により、画像Bを画像Aの動きに応じて移動させることが可能になる。よって、B内の足はA内の足の後を追う。次いで、この処理は、新たなフレームの各々に正確に位置する無限数のQを新たに受け取る。各Tについて、Zを推測し、出力は、移動する要素の正確な超分解能奥行き地図となる。

30

## 【0164】

次いで、t、t+1等において要素を囲む4Dマトリックスからの基準点を用いてZ次元が推測してもよい。これは、カメラの運動、および背景に対する焦点について行なってもよい。次いで、基準点からの光線により、以下の計算が可能となる：

3Dデータ構造に関する2D変換、

ピクセルの2Dデータ構造に関する3D変換、または三角測量法によるZの推測もしくはピクセルの3D位置のトラッキングのための特徴点、または

40

単一のカメラのビデオクリップから、移動する要素の正確な超分解能奥行き地図を作成するための特徴点。

## 【0165】

奥行き推測処理は、例えば、以下のようにすることが可能である。

{a, b, c, . . .}が既知の3D座標の3D基準点であると仮定する。

tは、(見つけたい3D座標を有する)時間「t」における要素の3D点(ピクセルまたは特徴)であり、

「t+1」は、時間「t+1」で同じ3D点である(これについても3D座標を求めることが望まれる)。

## 【0166】

50

基準点から  $t$  および  $t + 1$  への投影光線により、三角形  $[t + 1, t, a]$ ,  $[t + 1, t, b]$ ,  $[t + 1, t, c]$ , . . . が作成される。

【0167】

各三角形において、その基準点の3D座標は既知である。未知のものが6つ存在する（各三角形は未知の座標  $\{t, t + 1\}$  で構成され、 $t, t + 1$  は3D点である）。

【0168】

3D空間では、全ての三角形が「 $t$ 」と「 $t + 1$ 」の間に同じ距離を有する必要がある。この等式システムを解くことにより、各ピクセルの  $t$  および  $t + 1$  または各フレーム内の各  $Q$  の特徴を見つけ、画像平面上の三角形の2D投影から、三角形の3D座標を計算する。

10

【0169】

運動中の要素の3D構造マップを縮小することにより、システムが要素の3Dモデルをさらに完全に再構築し、異なるシーケンス間の3D形状を再生し、異なるシーケンス間の外見上の相違を処理することに役立つ。

【0170】

3D再構築

以下、3D奥行き地図上への3D骨格の投影および3DSP空間への奥行き地図の投影を用いながら、奥行き地図形成に適合する骨格器官を貼りつけ、合成によるコピー、およびその運動をキャプチャして、要素のモデルを作成することに関して説明する。最後の処理は、要素のテクスチャを再構築した骨格上に重ねることにより、さらに説明するように再構築処理を完成させることである。

20

【0171】

本方法は、フレーム内を移動する要素の3Dモデルの作成を可能にする。

【0172】

本アルゴリズムは、概念的には空間に基づくものである。2D画像平面上への3D情報の投影により、3D情報の推測を可能にする一方、空間次元に基づくアルゴリズムを用いて、その3D奥行き地図とともに2D世界を空間に基づく3D世界に投影する。

【0173】

上記で参照した運動中の要素の周囲に選択された  $N \times M \times Z$  ウィンドーは、実際には、（4Dマトリックスになる）3Dマトリックスであり、ユーザが予め定義された3D図形の内部骨格または部品（円筒、接続部等）を貼りつけることができる新たな（XYZ）軸フィールド「 $f$ 」である。

30

【0174】

奥行き推測の処理はまた、各ピクセルおよびフレーム間の特徴的な動きを特定し、時間の経過とともにピクセルおよび特徴の2D運動フローを作成することを含む。システムは、2D運動フローを3D運動フローに変換する。

【0175】

$S(0)$  の3D奥行き構造マップ、または奥行き推測の処理中において説明されるように1組みの3D軌跡を考慮すると、システムは、再構築アルゴリズムを用いて各フレームの係数  $D$  間（および統合されたフレーム）を、時間の経過とともに内部に貼りつけられた骨格（それ自身の係数  $D'$ ）と対応付けることにより、内部骨格の割合を定義、構築し、上記で参照した第4のベクトルを用いて空間および時間次元に関して3Dマトリックスを4Dマトリックスに変換する（すなわち、空間と時間の間の変換）。

40

【0176】

奥行き推測および再構築の処理は、直感的には、3D構造の2D的反映が2D画像マトリックスの下でレイヤー化される層状の機構である。この3D構造自体は、フレーム内の要素の合成的な3D再構築を行うために用いられる3D構造の2D的反映の下でレイヤー化される。この3層構造は、図7に関して上述したとおりである。空間次元の下で作用することにより、本明細書中において説明するように、3D構造およびテクスチャの再構築をして、要素の運動を確保することさえ可能であるが、ここで、出力は、フレーム内の要

50

素のモーションキャプチャを3Dモデル化したものである。また、これに関しては、出力は、特定の2D投影とすることもできる。

【0177】

移動する要素の再構築処理について、システムは、完全な3D構造奥行き地図を作成するか、または移動する要素の完全な3D構造奥行き地図を受け取ってもよい。この固定された周囲環境は、前述したように、移動する要素とは個別にモデル化される。

【0178】

本発明の各実施形態は、運動中の人間等の2D物体を、元の2Dまたは部分的な3D画像から3Dモデルへと完全な3D超分解能で再構築することを可能にする。この処理はまた、内部3D骨格に基づいて構築された3D構造テクスチャおよび運動をキャプチャすることを伴う。骨格は、完全な骨格物理データベースを用いて、完全な内部の骨および筋肉によって作ってもよい。システムにより、例えば、アニメーション、モーションキャプチャ、リアルタイムモデル化等、3D再構成モデルに関する無限の操作が可能になる。図12は、個々の骨格要素から作った完全な解剖学モデルを示す。

【0179】

投影は、基準点を用いた背景に対する要素の周囲の4Dマトリックスによる、画像および物体の3D奥行き地図を用いて行われる。

【0180】

図8に関して上記で説明したように、要素は、前述したように自動での特定または手動での特定を行い、3D奥行き地図への3D骨格またはその部品の貼り付けを用いて特定される。システムは、3Dデータ構造、すなわち、3D骨格を3D奥行き地図に投影する。

【0181】

フレーム上での要素の3D的な動きのトラッキングは、ピクセルおよび軌跡のDTM光学フローに基づく。トラッキングにより、要素の3D形成に関してできる限り多くの3D情報を学習することができ、奥行き地図上に骨格の3D構造を補間することにより、その3D形成を獲得する。画像内の要素をできる限り正確に提供するように骨格の3D構造を設計するために、骨格の3D構造は、時間の経過とともに学習した要素の情報を保存する。

【0182】

例えば、実写した人について考察する。第1の段階は、骨格をそのオブジェクトに割り付けることを伴う。第2の段階は、骨格を用いて、顔の構造、目、鼻等のオブジェクトの詳細を学習することを伴う。骨格から構造の種類を学習することにより、特定の詳細を受け取ることを求め、図形に合わせて正しい3Dの詳細を調整する処理となる。

【0183】

例えば、画像内の構造が人であることを認識した後、システムは、設定した方針に従って、例えば、目、鼻等の3Dおよび視覚情報を受信することを求める。SPが、例えば、頭部の特定の部位に、目および眉毛を求めるため、SPは3D図形に関して、この情報をより容易にかつより早く分析できる。

【0184】

移動するオブジェクトの器官間の位置および距離を計算することは、SPがより正確にオブジェクトの他の器官の関係を予想し、この情報を3D図形の情報と調整する際に役立つ。

【0185】

画像処理において空間次元を用いることにより、異なるフレームの各々が3D要素および図形の構築のために、より多くの3Dおよび視覚情報を供給するという点で、時間次元上の要素の動きに付加価値が付けられる。

【0186】

再構築処理と平行して、本発明の各実施形態は、要素の動きを獲得するモーションキャプチャツールとして用いることができ、ユーザがフレーム内の要素を、2D画像としてだけでなく、テクスチャを有する3Dモデルとしてもモーションキャプチャすることが可

10

20

30

40

50

能になる。

【0187】

再構築の処理は、次のように実行できる。

【0188】

初期構成は、4Dマトリックスであり、入力はDTMである。

【0189】

DTMは、外部アルゴリズムからのものとしてすることができる。DTMはまた、本発明の各実施形態の奥行き推測アルゴリズムからのものとしてすることもできる。処理時間に関しては、モデル化処理は、奥行き推測処理と平行する処理であり、直感的には、いくつかのマトリックスが他の1つ下に位置し、1つ目のマトリックスは、画像2Dマトリックスであり、その下に、(3D構造の)2D投影マトリックスがあり、さらにその下に3Dデータ構造がある。入力はまた、ピクセル、特に、フレームに沿って設定され、動きを連続させることを可能にする特徴点の2Dトラッキングに基づく3D軌跡を含む。特徴点は、色またはトラッキングが容易な他の性質に基づいてもよい。

10

【0190】

軌跡は、DTMに変換され、システムは、それらをフレームに沿ったピクセルおよび特徴点の3D位置を示す3D軌跡に変換する。

【0191】

システムは、3D骨格の投影と入力奥行き地図との間に時間軸に沿って拘束を設定し、要素の特定された器官を含む3D投影から生成された各フレーム内の無限の新たなトラッキング点を用いて正確なトラッキングを行なう。このようにして、3D物体のどの部分が隠れているか、およびそれらが3D空間のどこにあるかが分かる。システムは、3D骨格データ構造を用いて上記点の3Dトラッキングを行ない、移動する要素の正確な超分解能3Dモデルの作成を行なう。

20

【0192】

ワークフローは以下の通りである：ただし、

E3d - DTMのシーケンス、移動する要素の(x, y, z)の3Dマトリックス、

n - フレーム数、

S3d - 3D骨格、

Q3d - 各フレーム内の3D特徴点、

T3d - 軌跡(Q3dの3D点位置ベクトル)、

3d - Q3dに関するT3dの推移関数、

K - そのQ3dが有するフレーム数(=T3dの長さ)、

モデル - 再構築された3Dモデル。

30

【0193】

システムは、3D骨格SがDTM-Eと位置合わせされるように、 $F(E) \quad G(S)$ の位置合わせを行なう。システムは、T3dを用いて、(3dが軌跡ベクトルの関数 $3d(f(s, e) \quad i, q3d \quad i + 1)$ 、骨格SおよびDTM-Eに関するフレームi内の特徴 $Q3d \quad j$ の位置である)次のDTM-E $t + 1$ のE $t$ 上のQ3dを3Dトラッキングし、Q3dごとに(k個の)フレームのフレームi+1等における位置を加える。各フレームごとに、I+1内の新たな点をSに貼りつけ、フレームi+1に関する新たな情報により、Eの新たな位置に従ったSの位置合わせが可能になる(例えば、Sの足は、Eの足と位置合わせされる)。

40

【0194】

結果として、新たなDTMの各々に正確に位置する新たな無限数のQ3dを受け取る能力が得られる。係数DおよびD'により、システムは、tおよびt+1(等)における要素および骨格を囲む4Dマトリックス内において、E3dの全体的な3D情報のフォーメーションに従ってSのフォーメーションを変更することが可能になる。システムは、モデルから、次のフレームにおいて手足および他の要素が現れると予想される場所を予測する。Dは、合成により3D要素の形態でS3dをコピーする複雑な数学的構造における主要

50

因である。システムは、3D骨格を新たなデータ構造に変換し、移動する要素の正確な超分解能3Dモデルのコピーを作成するためのピクセルまたは特徴点の3D位置をトラッキングすることにより、3D骨格データ構造に関するDTMのシーケンス内に収集したフォーメーションを集めて保存する。

【0195】

{a, b, c, . . . .}がE3di上の3D座標を有する点であると仮定する。S3di{a', b', c'}上の対応する点をE3diの点に貼りつけるための予想が行なわれる。システムは、4Dマトリックスの部分空間の係数D'をE3diの係数Dと位置合わせし、S3diを統合した単位として位置合わせし、また、S3dを、各々が1つのD'係数を独立した4D部分空間として保持する所定の小規模4Dマトリックスに分割することによって、S3diのフォーメーションをE3diのフォーメーションに合わせて再構成し、次いで、i+1 . . . 上に構築する。ここで、出力は、要素の正確な超分解能再構築（さらに説明されるような形状およびテクスチャ）および移動する要素の3Dモーションキャプチャである。

10

【0196】

テクスチャの重ね合わせは、以下においてさらに説明するように、モデル化処理の一部である。

【0197】

上述した拘束は、要素の3Dモデルの完全な再構築を可能にする。システムが異なるシーケンス間の3D形状を再生すること、および異なるシーケンス間の外見上の相違を処理することも可能にする。3D構造に基づく位置合わせの性質を利用することにより、異なるビデオシーケンス間、異なる縮尺（ズーム）、異なる検知方法（IRおよび可視光線カメラ）でのマッチング等の非常に困難な状況において、情報をマッチングすることが可能になる。

20

【0198】

手に持った風船を膨らませる例において要素が時間および空間に関して個別に速度を変えるときにS0を撮影する場合、単独または統合された解決策として3つの選択肢が存在する。システムは、時間の経過とともに速度が変化する風船をモデル化し、構成に問題があるため奇妙に振舞い時間領域に問題を残すフレームもしくは一連のフレームまたはオブジェクトに印しをつけることができる。あるいは、ユーザは、自動システムがフレーム内の3D図形を定義することを手助けすることにより、どの3D情報および視覚情報を用いるべきかを伝える。

30

【0199】

処理の最後に、オブジェクトのより正確な3Dモデルのための再構築処理に関してより多くの3Dおよび視覚情報を与えるという点で異なるフレームごとに付加価値を有する上記の画像処理において空間次元を用いて、要素の動きによって、モデルとその3Dモーションキャプチャとを関連づけてもよい。

【0200】

時間次元上でのこれらの個別の円筒、球状、および他の骨格要素から完全な3D図形を構築することにより、それらの相互の動向（それらがどのようにして相互に貼りつけられるか、またはともに移動するか）を表し、図形のさらなるアニメーション化に役立ててもよい。

40

【0201】

モデルが完成すると、その元の図形および背景とは個別の図形として維持され、さらなるアニメーション化に用いられることができる。その元の動きは、モーションキャプチャに用いることができる。異なる時間または位置からより多くの視覚情報を同じ図形に付加してもよく、多角形内部骨格上にコンピュータにより生成された画像がそうであるように、ユーザの行動に応じて新たな3D図形に変化させることもできる。

【0202】

さらに、図形は、それが有する背景とは独立して、さらなるアニメーション化に用いる

50

ことができる。オブジェクトが、異なる時間または位置で実写されたまたは映像化された場合、SPは、異なる時間または位置から得た情報を組み合わせる。例えば、視覚情報は、異なる時間または位置で得ることができ、算出した情報は、図形の元の3Dモデルに付加される。オブジェクトの3D構造を用いて、特定の撮影された背景とは独立して、オブジェクト自体のレベルにおいて、アニメーションを生成したり、顔を模倣したり、声を加える等を行なうことができる。

**【0203】**

システムはまた、移動する要素内の運動をキャプチャーすることにより、完全な物体のアニメーションまたは顔の模倣等その一部のアニメーションのモーションキャプチャを用いて既存の3Dモデルをアニメーション化する能力を用いることもできる。

10

**【0204】**

上記で示した手に持った風船を膨らませる例のように、時間および空間の両方に関して相対的速度を変化させる要素が存在する場合、ユーザの支援が必要となることがある。ユーザは、システムがフレーム内の3D図形を定義することを支援するように問われ、どの3D情報を用いるべきか、アニメーションの貼り付けを行なうことなく、この図形を時間に基づくシーケンスとして残すべきかを示す。後者の場合、この要素は、通常の画像処理ツールを用いて編集してもよい。本例およびおよび先行する例の唯一の違いは、本例が時間に基づく3Dオブジェクトを維持しており、空間に基づく3Dオブジェクトではない点である。

**【0205】**

20

画像処理ツールにより、ユーザは、異なる時間および位置の環境を一緒に貼り付け、画像内の変形を修正し、要素を取り除き、入力によって作成された情報に基づいて新たな要素を作成し、さらに、3Dコンピュータにより生成された図形を生成するかまたは異なる3Dコンピュータアニメーションプログラムからコンピュータにより生成された図形を入力することが可能である。

**【0206】****視覚情報**

空間に基づく3次元モデル内の3次元位置を受け取った後、判定された分解能に応じて、各点は、1つまたは複数の視覚情報レイヤー（画像のデジタル情報に記録されているような色および輝度の値）を受け取る。

30

**【0207】**

実写画像の分解能と比較したモデルの分解能、各ピクセルの球状に関する情報、および異なるカメラから、または異なるクリップ等からの異なる質の視覚情報等、いくつかの視覚パラメータが考慮される。

**【0208】**

異なる画像分解能に関して、2つのケースが考えられる。

**【0209】**

1つ目のケースでは、画像分解能がFの制止分解能よりも高く、よって、3Dモデル内の各ピクセルに必要な以上の情報が存在する。例えば、インチ毎のピクセル数に関して、実写画像が5倍分大きい場合、システムは、5ピクセルごとに、視覚情報を合計し、1つのピクセル分として平均し、この新たな計算値を用いて3Dモデル内に新たなピクセルを作成する。

40

**【0210】**

2つ目のケースでは、3Dモデルの分解能が実写画像の分解能よりも高い場合である。ビデオシーケンスを用いて、各フレームは、フレーム内にテクスチャピクセルを生成し、カメラが少し移動すると、ピクセルは、近傍の3D点を撮影し、統合されたモデルについてより、画像内のピクセルの総量よりも多くの視覚情報を収集することを可能にする。このような場合は、例えば、遠い位置から画像を撮影しているか、またはデジタルズーム等を用いている場合に起こることがある。この場合、システムは、時間次元に沿って付加された近傍のピクセルから、ピクセルごとに情報を抽出する。ここでは、各ピクセルの視覚

50

情報の複数のレイヤーの主要要素が重要であり、さらに考察する。

【0211】

このとき、新たなピクセルが作成されており、システムの分解能のレベルでモデルの表面に重ねられている。新たなピクセルの各々は、ここでは、3D空間に基づくモデルにおける3次元位置を有しており、実生活とちょうど同じように、完全な360度から観察することができる。

【0212】

一般に、個々のピクセルは、360度から観察されることはない。例えば、壁の点は、180度から見ることができ（壁の背面は、異なるピクセルの3D位置に応じて、異なる情報を有する）、石の角は、270度から観察されること等が挙げられる。

10

【0213】

ここで、特定の方向からの視覚情報を与える実写された3D画像を示す図15を参照する。実写画像の各々は、特定の方向からの視覚情報を与える。SPが特定の方向のみからピクセルの視覚情報を受け取る場合、ピクセルを扁平にして、180度から見ることを可能にする。この場合、このピクセルを横方向から見ると視覚的な質に歪みが生じる。

【0214】

本発明の好適な実施形態は、空間次元における3D位置に対して、ピクセルごとに視覚情報の複数のレイヤーを統合する半球状のピクセルを形成する。無限数のピクセルを加えることが可能であり、視覚的な質に関しては、超分解能を作成している。

【0215】

超分解能はまた、統合されたモデル内に収集することができ、超分解能3D点を作成することができる奥行き点の数に関係する。奥行き点は、表面の変形を最も正確な方法で可能にする。

20

【0216】

3次元視覚情報に関して、システムにより大きな角度をカバーする視覚情報が提供されるほど、ソフトウェアのピクセルを必要とされるあらゆる方向から見る能力を模倣する能力がより良好となり、このピクセルに関する球状情報がより良好となり、ピクセルを必要とされる実現可能なあらゆる方向から見る能力が得られる。よって、ピクセルの周囲の様々な角度から複数の画像を撮影することができる。各ピクセルは、各クリップに沿った多くのフレームで実写することができる。この情報の全てが必要とされることはなく、これは、この情報の全てが同じ品質を有するものではないことを意味する。しかしながら、多層化した視覚情報をピクセルごとに記録することは、画像処理に必要な計算を低減し、画質を向上させることに役立つ。

30

【0217】

視覚的な質に関して、各ピクセルは、各クリップに沿って多くのフレームで実写することができる。それにもかかわらず、この情報の全てが同じ質ではないのと同様に、情報の全てが必要とされるわけではない。質が低い情報を破棄することは、画像処理に必要なとされる計算を低減することに役立つが、各情報は、不良な画質、照明、カメラ分解能等による不良な画質を向上するために用いられることが好ましい。

【0218】

システムは、品質等級Qを作成し、新たなフレームの各々から情報を得る新たな情報のレイヤーの各々を、その視覚情報の質および分解能について検査する。視覚情報は、2つの要因によって等級付けされるが、1つは時間次元における画質であり、他方は空間次元における画質である。

40

【0219】

例えば、SPは、室内および外庭の撮影に関して、建物の同じ位置から異なる絞りをを用いて撮影された2つのクリップを受け取る。

【0220】

カメラが高露出の絞りをを用いる1つ目のクリップでは、カメラは、画像の内部部分に関して良好な視覚情報を得ることができるが、庭の外部部分は、露出が過剰であり、画像上

50

では焼けたように、または過度に明るく見える。

【0221】

カメラが低露出の絞りをを用いる2つ目のクリップでは、画像の内部部分に関して非常に暗い視覚情報が作成されるが、画像の外部部分は非常にバランスがよく露出が良好である。

【0222】

これらのクリップの各々は、個々としてのバランスが良くなく、それら各々のヒストグラムはバランスが悪い結果を示す。

【0223】

しかし、システムが時間次元から変換したとき場合、各フレームが空間領域において個別に検査され、周囲環境が全体的に検査されるとき、システムが新たな視覚情報を受け取ると、以下の2つの要因に基づいてクリップを確認する：

1つ目の要因は、フレームごとにヒストグラムを個別の単位として獲得する時間次元、およびFに対する品質に基づくものであり、

2つ目の要因は、空間領域からのものであり、すでに構成された画像がフレーム内の特定のエリアを参照して、たとえFにおいて、Qが不良であったとしても、より高い品質を達成する。

【0224】

システムは、新たなクリップ内において、隣接するフレームピクセルとの相関関係とは関係なく、図14に関して上記で説明したように、ピクセルと隣接するFの不良なQとの相関関係において、SPにとって必要な特定の部分において視覚的により良好な品質を求める。システムは、本例において、同じ画像において同じ時間に異なる絞りをを用いて撮影されたように、可能な範囲で最高の品質で外庭および内部の部屋を示す、露出が非常に良好なイメージを与えるバランスがとれた画像を作成する。

【0225】

システムは、ある最小レベルのQまで画像情報を評価するが、これは、画像が上記2つの要因の両方の最小値よりも低い場合には、この情報を用いたり、その値をピクセルのテクスチャの既存の値に加えても意味がないことを意味する。

【0226】

この新たな情報を加える処理は、Q SPに基づく。

【0227】

Qが高くなるほど、情報がピクセル値に關与する値がより高くなり、Qが低くなるほど、その情報がピクセル値に關与する値が低くなる。

【0228】

システムは、両方のクリップからの情報を統合し、バランスがとれ、露光が良好な光景を部屋の内側および外庭の両方に生じる。

【0229】

システムは、品質Qについて閾値を設定し、それに応じて視覚情報を破棄してもよい。

【0230】

画像処理はまた、画像のバランスをとり、隣接する画像間での一貫性を達成するために、標準的なカメラ制御部(CCU)によって用いられるような処理方法を含んでもよい。

【0231】

観察

空間に基づいて構築され、視覚情報を含んだ3Dモデルは、SPに与えられるすべての画像シーケンスについてキャプチャされた全体的な結果である。

【0232】

集合的なフィールドにおける任意の点は、当該分野に限定されない任意の観察方法を用いて任意の視点から観察することができる。以下に例をいくつか示す。

【0233】

本発明の1つの実施形態では、2つの隣接するレンズの視野が水平軸に対してレンズ側

10

20

30

40

50

にある2つの隣接するレンズの視界と広範囲にわたって重なるように仮想的なカメラが配列される。これによって、立体画像を生成することができる。

【0234】

好適な実施形態により、完全に時間に基づくシーケンス、ライブシーケンス、非線形出力、立体/3D球形画像等の生成が容易になる。

【0235】

立体画像を提供する好適な実施形態では、仮想的なカメラが特定の構成で配列され、いずれのレンズの視野も、そのレンズの周囲に隣接する全てのレンズの視野と所望する範囲で重なり、全体的な視野が、完全に円形である画像の集合を含み、各視野内の任意の点が、立体球形画像を作成するための少なくとも2つの仮想的なレンズによって、または2D球形画像の作成のための1つの仮想的なレンズからもしくは任意の視点から3D球形画像を作成するための少なくとも2つの仮想的なレンズからキャプチャされる。

10

【0236】

結果として、単一のカメラによってフィルムに撮られた場面を観察するために立体データを利用することが可能になる。

【0237】

SPによって作成された画像は、静止画、ビデオ、立体視、仮想現実等の様々な形態で観察者に表示することが可能である。形成した画像は、TVまたはコンピュータスクリーン等の平面のスクリーン上で、または表示された画像の一部がユーザの視点に応じて変化する仮想現実用のヘッドセット等の仮想現実用の表示デバイスを用いて表示することができる。仮想現実による表示に適切な手段を用いて、水平および垂直方向において観察者の周囲360度を取り囲むことにより、観察者の周囲のあらゆる場所および上下方向を見る能力を与える一方、表示された画像の3D的な奥行きを知覚させる。

20

【0238】

公知の仮想現実のための手段を用いて、仮想現実による視覚の線形および非線形情報がユーザに提供される。このような手段は、観察者の頭部位置を検出するセンサを有するヘッドセット、または手の位置を検出するセンサを有する仮想的なグローブ、または任意の公知の表示用ソフトウェアであってもよい。

【0239】

TVまたはコンピュータスクリーン等の平面なスクリーン上に表示するために、ユーザの観察パラメータを、ユーザが持っているこの目的のためにプログラムされたポインティングデバイス(例えば、マウスまたはジョイスティック)から取得する。システムは、例えば、この発明のリアルタイムモーションキャプチャ能力、または外部デバイスからのモーションキャプチャを用いてユーザ自身の動きを集めることができる。

30

【0240】

仮想現実用のヘッドセットを着用して実際に頭部を回すことにより、またはユーザが持っているコンピュータデバイスに接続されたポインティングデバイスにより、観察者が特定の光景を選択すると、観察パラメータが表示システムにより検出され、受け取られる。観察者の観察パラメータは、観察者の観察方向および観察者の水平方向を含む。これらのパラメータに従って、観察者の周囲の座標に関して、観察者の視野が判定され、画像が観察手段に投影される。

40

【0241】

カメラの種類

本発明は、SPに与えられる画像または画像のシーケンスをキャプチャするために用いられるカメラの種類に関して限定されるものではない。カメラは、任意の公知のデジタルまたはアナログビデオカメラから選択されてもよい。これらのカメラはまた、デジタルでなくてもよく、その場合、公知の技術を用いて画像をデジタル形式に変換してもよい。

【0242】

好適な実施形態では、デジタル画像は、記憶する前または3Dモデルに基づく空間への変換前に品質を向上するために操作されてもよい。

50

## 【0243】

## 用途

ここで、本発明の異なる用途を示す吹き出し図である図10を参照する。

## 【0244】

好適な実施形態によると、上記で説明したように構築された完全に空間に基づく3Dモデルが利用可能な場合、ユーザは、仮想的な環境内に仮想的なカメラを配置して、元のシーケンス内にカメラがなかった視点から場面を実質的に再度撮影することができる。さらに、これはまた、リアルタイムで行なうことができる：例えば、バスケットボールの試合では、仮想的なカメラを配置して、実際にはカメラがない視点からゲームを撮影することができる。この場合に必要なのは、アリーナおよび個々の選手を予めモデル化することである。実際、このモデル化は、前もって行なう代わりに、放送の始めにリアルタイムで行なわれる。

10

## 【0245】

好適な実施形態に従って、上記のモデル化を用いることにより、シーケンスから一度キャプチャされた各図形を、例えば、外部ソースからのモーションキャプチャまたはSPのモーションキャプチャも用いることができるユーザによって再度アニメーション化することにより、元のクリップにおける図形の動き方を変更することができる。すなわち、ユーザは、元の実写画像からモデルを再構築する一方で、リアルタイムで図形の他の動きを出力することができる。

## 【0246】

好適な実施形態に従って、ユーザは、画像の元の図形を修正したり、または、その図形を、操作することにより完全に新しくなった図形と置き換えることさえ可能である。

20

## 【0247】

この実施形態によれば、本明細書中において論じた技術を用いて、各図形をその図形の3Dモデルと置き換えて、その図形自身を含む新たな映画クリップの作成を可能にすることにより、実写されたクリップ中での各図形の元の動きに依存することなく、各図形に関して新たなアニメーションを作成することができる。図形はまた、コンピュータゲーム、家庭用ゲーム機用ゲーム、TVゲーム等において、ユーザがリアルタイムで操作することができる。

## 【0248】

好適な実施形態では、アニメーションの場面またはビデオクリップの製造中に照明を加える公知の技術を用いて、3Dモデルに新たに照明を加える。

30

## 【0249】

好適な実施形態では、要素および背景の3Dモデル化の一部として、アリーナ内の各要素および背景について、任意の所望する基準点までの奥行きを推測が含まれる。奥行きを推測は、システムに入力された写真図形のシーケンスに関する奥行き地図の解析を含んでおり、これは、下記でより詳細に説明されるような多くの方法で行なうことが可能である。

## 【0250】

好適な実施形態の各々により、画像上の動きをぼやけさせるなど、種々の操作が可能にすることができる。

40

## 【0251】

本明細書中に記載の技術を用いて、場面を撮影中に行なうことができる異なる種類の操作の全ては、焦点の変更、ズームおよび照明の修正等、3Dの仮想的なアリーナでも行なうことができる。

## 【0252】

本明細書中に記載の技術を用いて、ユーザは、図形および背景から完全なモーションピクチャを作成することができる。

## 【0253】

本明細書中において記載した技術を用いて、ユーザは、全ての図形が実際の画像に基づ

50

く3Dモデルである3D空間に基づくモデルを用いて完全なコンピュータゲーム(家庭ゲーム機用のゲーム、TVゲーム等)を作成することができる。

【0254】

コンピュータにより生成された画像を、3次元環境およびその中の3次元モデルに付加することができるのが好ましい。これらの画像は、モデルのスキンを変更する、コンピュータにより生成された要素をモデルまたは背景等にさらに付加するといった効果を有し得る。

【0255】

この実施形態に従って、ユーザは、シーケンス内の個々の図形と関連付けられた時系列情報を用いて、モーションキャプチャ時に図形の運動を再構築することができる。本技術は、単一のカメラからの画像のシーケンス、または2つ以上のカメラからの画像を用いて作用する。

【0256】

本明細書中において記載した手順では、2次元および3次元トラッキングを、時間に基づくクリップ内での動きに基づいて、特定された図形および背景のいずれにも適用することができる。このトラッキングは、リアルタイムで、またはクリップを再度アニメーション化する工程の一部として後から行なうことができる。

【0257】

本発明の好適な実施形態に従って、ユーザは、移動しているまたは静止している要素を空間に基づく3D環境内の図形または背景に追加してもよい。

【0258】

本発明の好適な実施形態に従って、ユーザは、元々実写されたものではない新たなアリーナを作成することができる。例えば、ユーザは、いくつかの異なる周囲環境を合成して、統合したアリーナを作成したり、実写されたアリーナをコンピュータにより生成された合成アリーナと組み合わせてもよい。

【0259】

本発明の好適な実施形態に従って、ユーザは、3Dモデル内において本発明の実施形態を用いて再構築される図形を用いて、その背景からその図形を取り除き、異なるアリーナに移動するか、またはコンピュータにより生成された任意のプログラムにエクスポートすることができる。

【0260】

本発明の好適な実施形態に従って、ユーザは、再構築された図形に基づいて新たな図形を作成することができる。ユーザは、テクスチャ、器官等をさらに追加または変更してもよい。

【0261】

好適な実施形態に従って、ユーザは、例えば、古い映画等の既存の映像を用いてもよく、その映画のデータを用いて映画の図形および背景をモデル化してもよい。これは、図形の完全な3D空間に基づく環境もしくはアリーナおよびその内部の位置を作成することにより行なわれてもよく、作成した3D環境に基づいて、元の図形および周囲環境から作られた新たな映画を作成してもよい。

【0262】

本発明の好適な実施形態に従って、ユーザの仮想的な3D複製物を用いて仮想的な集会を行なうことができる。そのような仮想的な集会は、ユーザのモーションキャプチャを伴ってもよい。アプリケーションにより、ユーザは、先生がユーザの3D図形を見て、ユーザの動きを直し、各生徒が他の生徒を3D図形として見るような仮想的な武術レッスンに参加することができる。モーションキャプチャは、ユーザ自身のウェブカメラを用いて行なうことができる。

【0263】

このようなアプリケーションはまた、他の教育目的、仮想的な身体的トレーニング、仮想的なビデオ会議等に用いられてもよい。3Dモデルおよびモーションキャプチャはまた

10

20

30

40

50

、仮想的な展覧会、多人数が参加するゲーム、またはさらに仮想的なデートに用いられてもよい。

【0264】

本発明の好適な実施形態に従って、空間に基づく3Dモデルは、兵士をトレーニングするための戦闘アリーナのシミュレーション、飛行シミュレーション等のシミュレーションに用いられてもよい。

【0265】

本発明の好適な実施形態に従って、3Dアリーナは、医療デバイスに用いることができる。1つ以上のセンサから獲得した画像を操作するために用いてもよい。画像は、実際の手術中にリアルタイムで用いるか、またはシミュレーション目的で用いる体の器官の3Dモデルを作成するために用いてもよい。

10

【0266】

本明細書中において記載した3Dモデルおよび環境は、例えば、建築および建設工学等の計画および設計に用いてもよい。

【0267】

本発明の1つの特定の用途では、本明細書中において記載したモデルおよび環境はまた、PALおよびNTSC等の異なるビデオ基準間の変換に用いてもよい。

【0268】

本明細書中において提供される技術の1つの用途は、ビデオ圧縮である。この用途では、実写されたクリップを用いた空間に基づく3Dモデル化により、モデルの伝達が可能になり、その後に必要なことは、実質的には、動き情報の伝達だけである。このような技術は、ビデオフレームの伝送に関して、帯域幅を大幅にセーブする。この用途は、モーションピクチャから携帯電話用ビデオクリップまで、様々なビデオの使用法および様々な質の規格に適用可能である。

20

【0269】

さらに、本発明の各実施形態は、記録したものを直接、本発明の各実施形態の3D空間に基づくモデルにしたりまたはそれに適用したりする、ビデオレコーディングの新たな方法を提供する。ビデオフレーム自体は、情報がモデルに抽出された後に再生することができる。

【0270】

本発明の各実施形態の3Dモデルは、単一のソースから移動する要素をリアルタイムでキャプチャしてモデル化し、任意の方向からそれらを観察するために用いることができる。1つの用途では、複数のユーザが、異なるスクリーンにおいて、これらの図形を任意の方向または倍率でリアルタイムに見ることができる。

30

【0271】

本発明の好適な実施形態によるデバイスでは、ユーザの3D的な動きをキャプチャするためにリアルタイムでシステムを用いてもよく、それを用いて、任意のコンピュータプログラムに関して、手または体の3D的な動きによってコンピュータを完全に操作してもよい。この実施態様では、特定のカメラ、通常ビデオカメラ、スチールカメラまたは携帯電話用カメラ等の通常カメラを用いてもよい。例えば、ユーザは、ゲーム内に存在する2Dまたは3Dキャラクターの1つがユーザの動きに従って動くコンピュータゲーム内に入ることも可能である。これはまた、携帯電話または任意の他の携帯用機器のユーザインターフェイスで行なうこともできる。

40

【0272】

本発明の好適な実施形態に従って、ユーザは、完全なまたは部分的な3Dモデルとして自己をモデル化することができ、コンピュータゲームまたは任意の他の関連するコンピュータプログラム内に入る事ができる。

【0273】

本発明の各実施形態を適用することにより、コンピュータゲーム、シミュレータ、またはそのようなプラットフォームの任意の変形において、全くの現実の画像から、2D/3D

50

図形および背景を作成することができる。

【0274】

本発明の好適な実施形態に従って、3Dモデル化は、赤外線等、あらゆる種類のセンサによって収集された情報を用いて行なうことができる。

【0275】

本発明の好適な実施形態に従って、顕微鏡からの情報も、適切なセンサから集めたデータを用いて、新規な3D空間に基づくモデルにモデル化することができる。

【0276】

本発明の好適な実施形態に従って、光学情報を有さないセンサから集められたデータによって、新たなユーザ定義の2D/3Dアリーナ、例えば、素粒子、遠くの星、またはセンサがキャプチャできないエリア（例えば、壁の裏側）でさえも作成するために3Dモデルおよびテクスチャを用いることができる。

10

【0277】

本発明の好適な実施形態に従って、3D SP処理は、マシンビジョンを可能にするために用いてもよい。例えば、ロボットにある場面の3次元空間的な解釈をさせるために用いてもよい。これによって、ロボットは、多数のフレーム内の部分的な画像としてではなく、統合された3次元物体として人間を関連付けることができる。これによって得られるロボットは、例えば、障害者を補助する等の用途に用いてもよい。

【0278】

用途の必要に応じて、3D SP処理は、インチ毎のテクスチャピクセル数、およびモデルの3Dフォーメーションを構築する奥行き点の数に関して、超分解能で再構築された3Dモデルを作成してもよい。

20

【0279】

本特許の有効期間中に、多くの関連する撮影術ならびに撮像デバイスおよびシステムが開発されることが予想されるが、特に、「3Dモデル」、「画像キャプチャ」、「奥行き地図」、「クリップ」、「仮想現実」、および「コンピュータ」等の本明細書中における用語の範囲が、そのような新たな技術の全てを予め含むものとする。

【0280】

本発明のさらなる目的、利点、および新規な特徴は、以下の限定を目的としない例を考察することにより当業者には明らかとなる。また、本明細書中において上述したおよび請求の範囲に記載した本発明の種々の実施形態および局面の各々に関しては、以下の例において実験に関する裏づけが得られる。

30

【0281】

明瞭にするために、個別の実施形態の文脈において記載した本発明の特定の特徴は、組み合わせる1つの実施形態として提供してもよいことは言うまでもない。逆に、簡潔にするために、1つの実施形態の文脈において記載した本発明の種々の特徴は、個別にまたは任意の適切な部分的組み合わせとして提供してもよい。

【0282】

本発明をその特定の実施形態に関して説明したが、当業者には多くの代替例、修正物および変形例が明らかであることは明白である。従って、請求の範囲の精神および一般的な範囲に該当するような代替例、修正物および変形例の全てが含まれるものとする。本明細書中において記載した全ての文献、特許および特許出願は、本明細書中において個々の文献、特許または特許出願の各々を特定的かつ個別に参考として援用すると示した場合と同様にして、それらの全てを参考として本明細書中において援用する。また、本出願におけるあらゆる参考文献の引用または表示は、そのよう参考文献が本発明の従来技術として適用されることを了承したものとして理解されるべきものではない。

40

【図面の簡単な説明】

【0283】

【図1】図1は、従来技術による3Dコンピュータにより生成された仮想的な図形の写真である。

50

【図 2 a】図 2 a は、従来技術による仮想的なスタジオの第 1 の写真である。

【図 2 b】図 2 b は、従来技術による仮想的なスタジオの第 2 の写真である。

【図 3】図 3 は、本発明の第 1 の好適な実施形態による 3 D 画像分析のための装置の簡略ブロック図である。

【図 4】図 4 は、本発明の好適な実施形態に従って 1 つの画像または一連の画像から特定された構造要素に骨格要素を挿入する手順を示した簡略フローチャートである。

【図 5】図 5 は、一連の要素が単一の物体として認識される図 4 a の手順の変形を示した簡略フローチャートである。

【図 6】図 6 は、本発明の好適な実施形態による骨格の挿入、および 2 d 画像データに 3 次元解釈を与えるその後の使用法について示す簡略フローチャートである。

10

【図 7】図 7 は、2 D 画像に 3 D 的な解釈を与えるために図 4 a ~ 図 4 c の骨格を多層フォーマットで使用することを示す簡略図である。

【図 8】図 8 は、本発明の好適な実施形態に従って骨格要素を挿入する 2 次元画像からオブジェクトを特定する 2 つの方法を示した流れ図である。

【図 9】図 9 は、画像内の移動する 2 D 要素に 3 D 的解釈を与えるために、本発明の各実施形態に従って骨格を用いる処理を示す簡略フローチャートである。

【図 10】図 10 は、本発明の各実施形態の一連の例示的用途を示す吹き出し図である。

【図 11】図 11 は、本発明の好適な実施形態に従って実行可能な撮像処理を示す奥行き地図である。

【図 12】図 12 は、本発明の好適な実施形態に従って実行可能な撮像処理を示す奥行き地図に貼りつけられた骨格である。

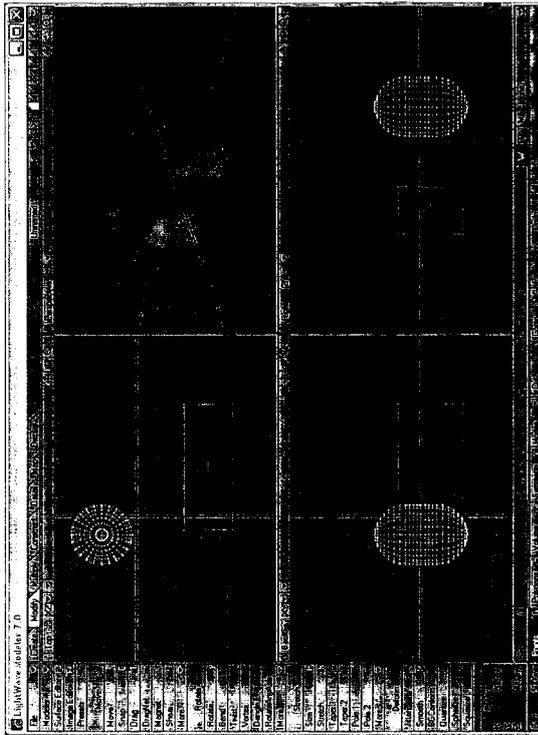
20

【図 13】図 13 は、構造の変形処理を明確に示す骨格であり、本発明の好適な実施形態に従って実行可能な撮像処理を示す。

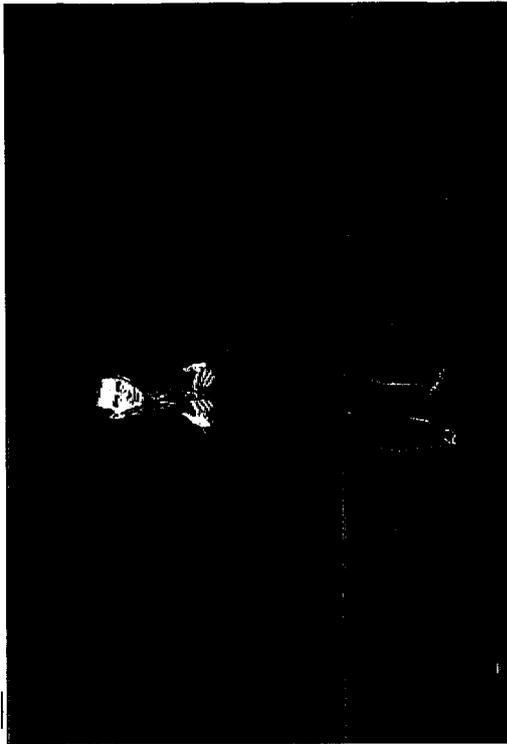
【図 14】図 14 は、一連の画像内の構造要素が変形が最小となる位置のフレームからどのように処理されるべきかを示す。

【図 15】図 15 は、特定の方向から 3 D 情報を供給する実写画像を示す。

【図 1】



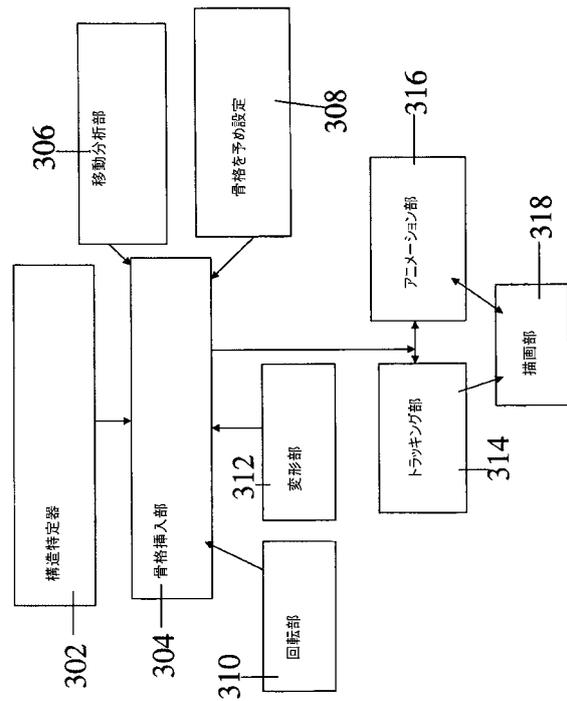
【図 2 a】



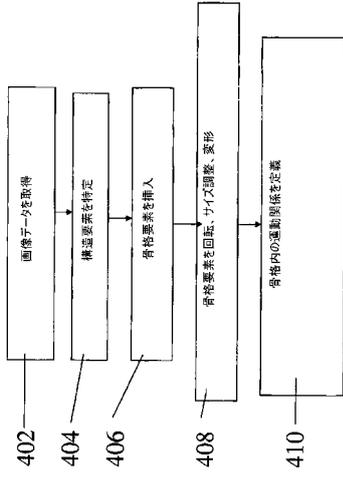
【図 2 b】



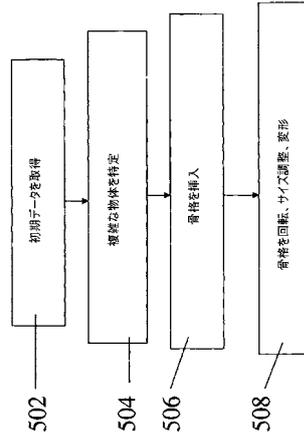
【図 3】



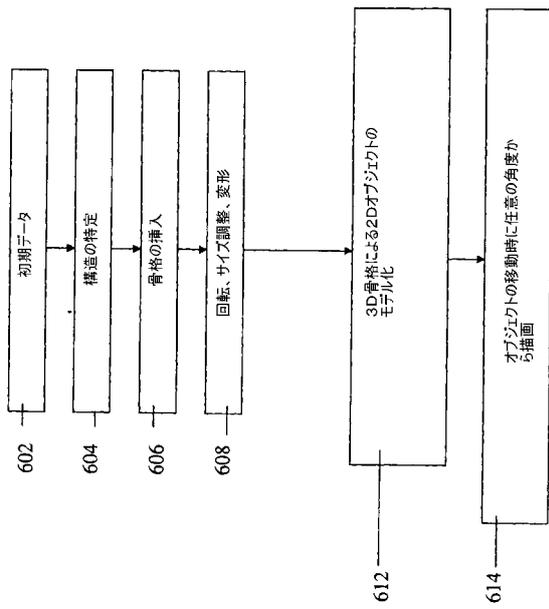
【 図 4 】



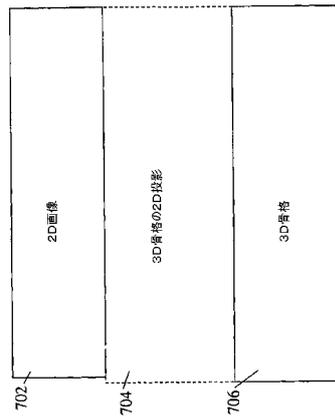
【 図 5 】



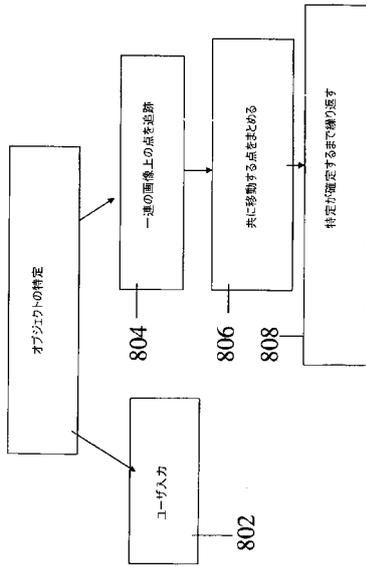
【 図 6 】



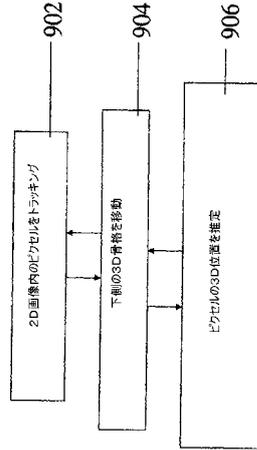
【 図 7 】



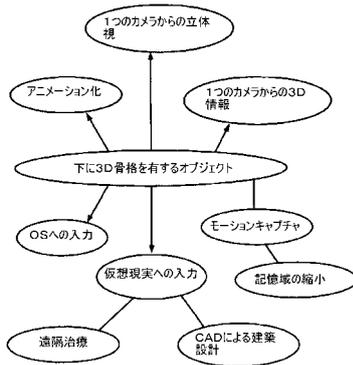
【 図 8 】



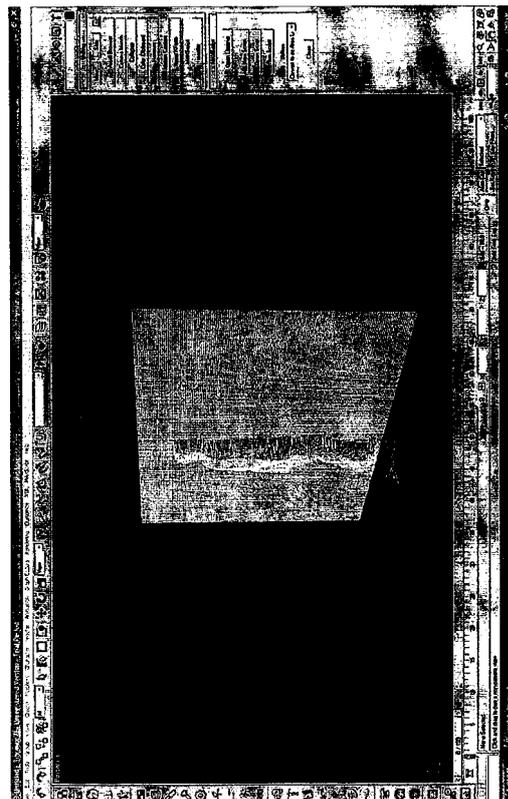
【 図 9 】



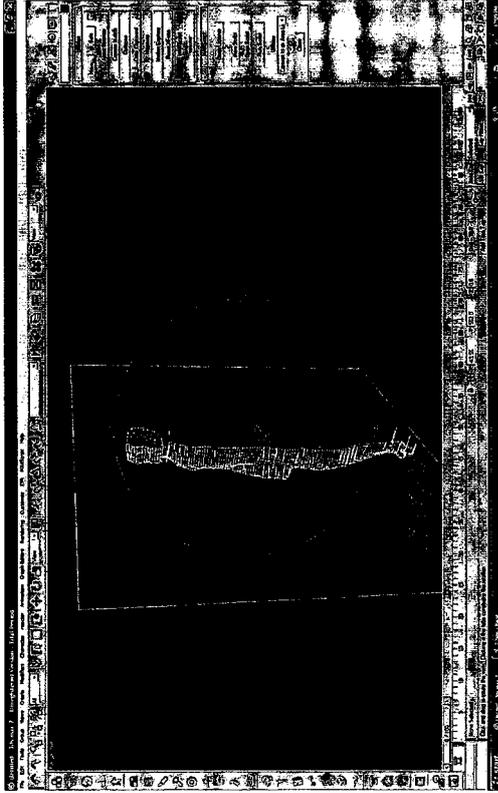
【 図 10 】



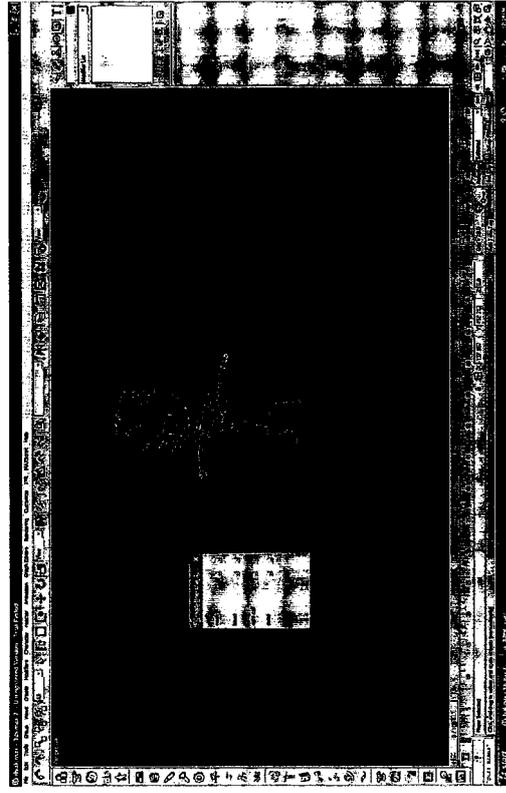
【 図 11 】



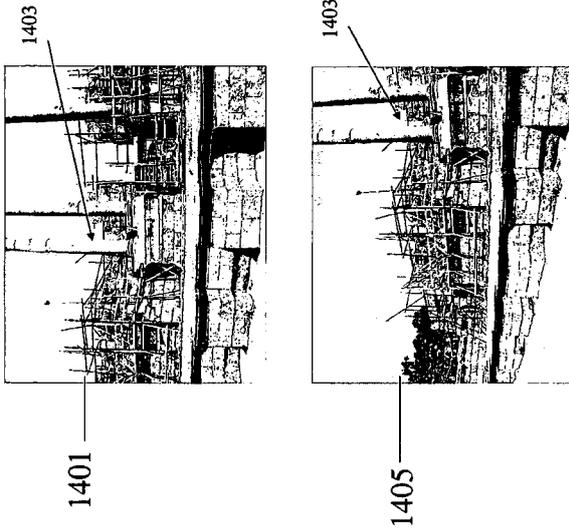
【 1 2 】



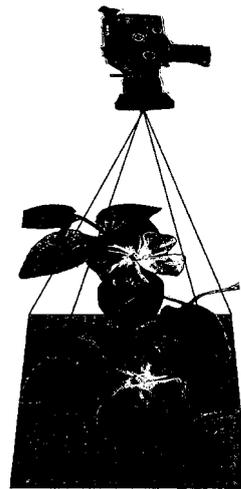
【 1 3 】



【 1 4 】



【 1 5 】



---

フロントページの続き

審査官 西出 隆二

(56)参考文献 国際公開第2004/013814(WO, A1)  
特開2002-216146(JP, A)  
特開2004-062692(JP, A)

(58)調査した分野(Int.Cl., DB名)  
G06T 13/40  
G06T 1/00