

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-123799

(P2012-123799A)

(43) 公開日 平成24年6月28日 (2012. 6. 28)

(51) Int.Cl.	F I	テーマコード (参考)
GO6T 19/00 (2011.01)	GO6T 17/40 A	5B050
HO4N 13/04 (2006.01)	HO4N 13/04	5C061
HO4N 7/26 (2006.01)	HO4N 7/13 Z	5C159

審査請求 未請求 請求項の数 12 O L (全 22 頁)

(21) 出願番号 特願2011-264974 (P2011-264974)
 (22) 出願日 平成23年12月2日 (2011. 12. 2)
 (31) 優先権主張番号 61/419, 501
 (32) 優先日 平成22年12月3日 (2010. 12. 3)
 (33) 優先権主張国 米国 (US)

(71) 出願人 504139662
 国立大学法人名古屋大学
 愛知県名古屋市千種区不老町 1 番
 (74) 代理人 110000578
 名古屋国際特許業務法人
 (72) 発明者 ウィルダブラ メインダード オノ
 愛知県名古屋市千種区不老町 1 番 国立大
 学法人名古屋大学内
 (72) 発明者 楊 路
 愛知県名古屋市千種区不老町 1 番 国立大
 学法人名古屋大学内
 (72) 発明者 パナヒフル テヘラニ、メヒルダド
 愛知県名古屋市千種区不老町 1 番 国立大
 学法人名古屋大学内

最終頁に続く

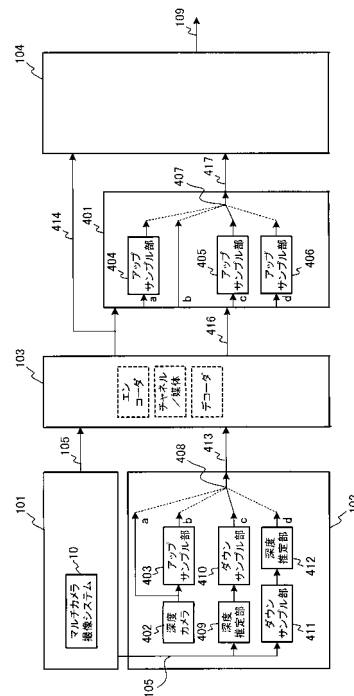
(54) 【発明の名称】 仮想視点画像合成方法及び仮想視点画像合成システム

(57) 【要約】

【課題】 アップサンプリングされた深度マップに遮断作用又は平滑作用を及ぼさないアップサンプリング方法を低いビットレートで実現する仮想画像合成技術を提供する。

【解決手段】 多視点撮影サブシステム 101 にて、複数の視点に設置したカメラから、多視点ビデオ画像 105 を取得し、形状生成サブシステム 102 にて、形状情報 106 を生成するとともに、深度カメラ 402 で多視点ビデオ画像 105 よりも低解像度の深度マップ 413 を取得する。アップサンプリングサブシステム 401 にて、多視点ビデオ画像 105 及び深度マップ 413 に基づき深度マップ 413 をアップサンプリングする。仮想ビュー合成サブシステム 104 にて、多視点ビデオ画像 105、アップサンプリングした深度マップ及び、画像を合成する際の視点となる仮想視点の位置情報と方向情報に基づいて、仮想視点情報に対応する仮想視点画像を合成することにより、仮想視点画像を合成する。

【選択図】 図 1



【特許請求の範囲】**【請求項 1】**

複数の視点から取得した画像情報に基づき、仮想視点から見た仮想視点画像を合成する仮想視点画像合成方法であって、

複数の視点に設置した複数の画像取得手段から、前記仮想視点画像の基準となる基準画像を取得する基準画像取得工程と、

前記複数の画像取得手段と同じ視点における画像の深度を取得する画像深度取得手段によって、前記複数の画像取得手段と同じ視点における画像の深度マップを生成する深度マップ生成工程と、

前記深度マップ生成工程において生成した深度マップをアップサンプリングするアップサンプリング工程と、

画像を合成する際の視点となる仮想視点の位置情報と該仮想視点から前記合成画像を観たときの方向である仮想視点方向の情報を取得する仮想視点情報取得手段から前記仮想視点の位置情報及び方向情報を取得する仮想視点情報取得工程と、

前記基準画像取得工程において取得した前記基準画像、前記アップサンプリング工程においてアップサンプリングした前記深度マップ及び前記仮想視点情報取得工程において取得した前記仮想視点の位置情報及び方向情報に基づいて、前記仮想視点の位置情報及び方向情報情報に対応する仮想視点画像を合成する仮想視点画像合成工程と、

により、仮想視点画像を合成することを特徴とする仮想視点画像合成方法。

【請求項 2】

請求項 1 に記載の仮想視点画像合成方法において、

前記画像深度取得手段は、画像の深度を検出する深度カメラであり、

前記アップサンプリング工程は、

前記深度カメラから入力された深度マップを入力し、

前記深度マップ生成工程において生成された深度マップにおける一組の隣接するピクセルと、前記基準画像における互いに隣接していないピクセルとを対応させ、

前記深度マップにおける一組の隣接するピクセルにおける各ピクセルに重みを割り当て

、前記一組の隣接するピクセルに割り当てた各ピクセルの重みを最適化し、最小重みを算出し、

前記算出した最小重みを得たときの前記一組の隣接するピクセルにおける最適な深度値を選択することを特徴とする仮想視点画像合成方法。

【請求項 3】

請求項 2 に記載の仮想視点画像合成方法において、

前記重みは、前記基準画像のピクセルと前記深度マップにおける前記一組の隣接するピクセルとの色又は強度差及び距離に基づいて割り当てられることを特徴とする仮想視点画像合成方法。

【請求項 4】

請求項 2 に記載の仮想視点画像合成方法において、

前記重みは、前記深度カメラから入力された深度マップ及び / 又は前記基準画像のピクセルと、前記深度カメラから入力された深度マップ及び / 又は前記基準画像における一組の隣接するピクセルとの色又は強度差及び距離の組合せに基づき割り当てられることを特徴とする仮想視点画像合成方法。

【請求項 5】

請求項 2 ~ 請求項 4 のいずれか 1 項に記載の仮想視点画像合成方法において、

前記重みの最適化は、権利独占選択によることを特徴とする仮想視点画像合成方法。

【請求項 6】

請求項 2 ~ 請求項 5 のいずれか 1 項に記載の仮想視点画像合成方法において、

前記最適深度値の選択は、

出力深度値として最小重みを有するピクセルの深度を選択することによることを特徴と

10

20

30

40

50

する仮想視点画像合成方法。

【請求項 7】

複数の視点から取得した画像情報に基づき、仮想視点から見た仮想視点画像を合成する仮想視点画像合成システムであって、

複数の視点に配置した画像取得手段と、

前記複数の画像取得手段から画像構成の基準となる基準画像を取得する基準画像取得手段と、

前記複数の画像取得手段と同じ視点における画像の深度を取得する画像深度取得手段と、

前記画像深度取得手段により取得した、前記複数の画像取得手段と同じ視点における画像の深度マップを生成する深度マップ生成手段と、

前記深度マップ生成手段において生成した深度マップをアップサンプリングするアップサンプリング手段と、

画像を合成する際の視点となる仮想視点の位置情報と該仮想視点から前記合成画像を観たときの方向である仮想視点方向の情報を取得する仮想視点情報取得手段と、

前記基準画像取得手段において取得した前記基準画像、前記アップサンプリング手段においてアップサンプリングした前記深度マップ及び前記仮想視点情報取得手段において取得した前記仮想視点の位置情報及び方向情報に基づいて、前記仮想視点の位置情報及び方向情報に対応する仮想視点画像を合成する仮想視点画像合成手段と、

を備えたことを特徴とする仮想視点画像合成システム。

【請求項 8】

請求項 7 に記載の仮想視点画像合成システムにおいて、

前記画像深度取得手段は、画像の深度を検出する深度カメラであり、

前記アップサンプリング手段は、

前記深度カメラから入力された深度マップを入力する深度マップ入力手段と、

前記深度マップ入力手段で入力した前記深度マップにおける一組の隣接するピクセルと

、前記基準画像における互いに隣接していないピクセルとを対応させる対応手段と、

前記深度マップにおける一組の隣接するピクセルにおける各ピクセルに重みを割り当てる重み割り当て手段と、

前記重み割り当て手段で前記一組の隣接するピクセルに割り当てた各ピクセルの重みを最適化し、最小重みを算出する最小重み算出手段と、

前記最小重み算出手段で算出した最小重みを得たときの前記一組の隣接するピクセルにおける最適な深度値を選択する最適深度値選択手段と、

を備えたことを特徴とする仮想視点画像合成システム。

【請求項 9】

請求項 8 に記載の仮想視点画像合成システムにおいて、

前記重み割り当て手段は、

前記基準画像のピクセルと、前記深度マップ入力手段で入力した前記深度マップにおける前記一組の隣接するピクセルと、の色又は強度差及び距離に基づいて前記重みを割り当てることを特徴とする仮想視点画像合成システム。

【請求項 10】

請求項 8 に記載の仮想視点画像合成システムにおいて、

前記重み割り当て手段は、

前記深度マップ入力手段で入力した深度マップ及び / 又は前記基準画像のピクセルと、前記深度マップ入力手段で入力した深度マップ及び / 又は前記基準画像における一組の隣接するピクセルとの色又は強度差及び距離の組合せに基づき、前記重みを割り当てることを特徴とする仮想視点画像合成システム。

【請求項 11】

請求項 8 ~ 請求項 10 のいずれか 1 項に記載の仮想視点画像合成システムにおいて、

前記最小重み算出手段は、

10

20

30

40

50

権利独占選択により、前記重みの最適化を行うことを特徴とする仮想視点画像合成システム。

【請求項 1 2】

請求項 8 ~ 請求項 1 1 のいずれか 1 項に記載の仮想視点画像合成システムにおいて、前記最適深度値選択手段は、

前記出力深度値として最小重みを有するピクセルの深度を選択することを特徴とする仮想視点画像合成システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数位置において取得した画像に基づき仮想視点における合成画像を得る仮想視点画像合成方法及び仮想視点画像合成システムに関する。

【背景技術】

【0002】

仮想視点画像とは、仮想位置において実際のカメラによって撮影されたかのように見える画像のことである。例えば、対象及び背景を、2台のカメラによって撮影した場合、2台の実際のカメラの間の位置から撮影されたかのように見える画像が生成される。この画像は「仮想視点画像」と呼ばれる。

【0003】

この画像を生成するプロセスは、「レンダリング」又は「ビュー合成」と呼ばれる。以下で、「視点画像」は、指定された視点から撮影された画像であり、実際のカメラによるかあるいはビュー合成のプロセスを通じて生成される。さらに、本明細書で、「画像」という語は、画像ピクセルからなるデジタル画像を指す。

【0004】

人は、各々の目で異なるビューを見るので深度を感知することができる。最新の3Dビデオシステム(3D-TV又は自由視点TVなど)の原理は、各々の目に対して1つの、2視点画像を生成することである。視点に自由度を持たせるためには多くの視点画像が必要である。3Dシーンの情報は、多くの方法で得られて表現され得る。

【0005】

よく使用される3Dシーン表現は、深度画像がシーン形状を表現するNビュー及びN深度画像に基づいている。図10に、複数のビュー及び形状に基づく多視点ビデオシステムの一般化されたシステム図を示す。

【0006】

複数のカラービューは、一般に、複数の同期カメラによって撮影される。形状情報は、例えば、3Dモデル又はパーピクセル深度画像によって表現され得る。深度画像に基づくレンダリングを使用すると、実際のカメラで撮影されているように思われる無数の仮想視点画像が所与の範囲内で合成され得る(例えば、非特許文献1参照)。

【0007】

深度画像に基づくレンダリングは、パーピクセル深度値を用いて所与の視点画像の画像ピクセルを別の視点画像に投影する仮想ビュー合成プロセスである。この投影は、一般に、3Dワーピングと呼ばれる。

【0008】

Nビュー及びN深度表現の長所の一つは、受信機側における所要の処理が比較的低いことである。さらに、所要の送信/記憶帯域幅が抑制され得る。例えば、3Dディスプレイが20視点画像を必要とする場合、20視点画像を送信しなくても2つ又は3つのビュー及び対応する深度マップを送信すれば十分であり得る。

【0009】

一般に、多視点ビデオシステムでは、複数の深度マップ及びビューが記憶又は送信のために圧縮される。ビュー及び深度の両方の効率的圧縮と信頼性の高い高品質の仮想ビュー合成が、このようなシステムにおいて重要である。

10

20

30

40

50

【0010】

圧縮のために深度をダウン/アップサンプリングするアプローチでは、補間されたサンプルによって低解像度の深度マップのみから推定される従来のアップサンプリングが採用されている（例えば、非特許文献2、非特許文献3参照）。

【先行技術文献】

【非特許文献】

【0011】

【非特許文献1】C. Fehn, 「Depth - image - based rendering (DIBR), compression and transmission for a new approach on 3D - TV (3D - TVに関する新しいアプローチのための、深度画像に基づくレンダリング(DIBR)、圧縮、及び送信)」、Proc. SPIE Stereoscopic Displays and Virtual Reality Systems, XI, pp. 93 - 104 (Jan. 2004)

10

【非特許文献2】S. Shimizu, M. Kitahara, K. Kamikura, 及びY. Yashima, 「View scalable multiview video coding using 3 - D warping with depth map (深度マップを使用した3Dワーピングを用いるビュースケラブル多視点映像符号化)」、IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, pp. 1485 - 1495, November 2007

20

【非特許文献3】K - J. Oh, S. Yea, A. Vetro, Y - S. Ho, 「Depth Reconstruction Filter and Down/Up Sampling for Depth Coding in 3 - D Video (3 - Dビデオにおける深度符号化のための深度再構成フィルタ及びダウン/アップサンプリング)」、IEEE signal processing letters, vol. 16, No. 9, Sept. 2009, p747 - 750

【発明の概要】

【発明が解決しようとする課題】

【0012】

30

ところが、従来のアップサンプリング方法を適用すると、アップサンプリングされた深度マップに強力な遮断作用又は平滑作用を及ぼす。これは、合成されたビューにアーチファクトをもたらすので望ましくないという問題がある。

【0013】

本発明は、こうした問題に鑑みなされたもので、アップサンプリングされた深度マップに遮断作用又は平滑作用を及ぼさないアップサンプリング方法を低いビットレートで実現する仮想画像合成技術を提供することを目的とする。

【課題を解決するための手段】

【0014】

40

上記[発明が解決しようとする課題において述べた問題を解決するためになされた請求項1に記載の発明は、

複数の視点から取得した画像情報に基づき、仮想視点から観た仮想視点画像を合成する仮想視点画像合成方法であって、

複数の視点に設置した複数の画像取得手段(10)から、仮想視点画像の基準となる基準画像を取得する基準画像取得工程と、

複数の画像取得手段(10)と同じ視点における画像の深度を取得する画像深度取得手段(402, 409, 412)によって、複数の画像取得手段(10)と同じ視点における画像の深度マップを生成する深度マップ生成工程と、

深度マップ生成工程において生成した深度マップをアップサンプリングするアップサンプリング工程と、

50

画像を合成する際の視点となる仮想視点の位置情報と仮想視点から合成画像を観たときの方向である仮想視点方向の情報を取得する仮想視点情報取得手段から仮想視点の位置情報及び方向情報を取得する仮想視点情報取得工程と、

基準画像取得工程において取得した基準画像、アップサンプリング工程においてアップサンプリングした深度マップ及び仮想視点情報取得工程において取得した仮想視点の位置情報及び方向情報に基づいて、仮想視点の位置情報及び方向情報情報に対応する仮想視点画像を合成する仮想視点画像合成工程と、

により、仮想視点画像を合成することを特徴とする仮想視点画像合成方法である。

【0015】

このような仮想視点画像合成方法によれば、アップサンプリングされた深度マップに強力な遮断作用又は平滑作用を及ぼさず、合成されたビューにアーチファクトをもたらさないものとなる。したがって、符号化効率の改善及びビュー合成品質の向上に寄与する。

【0016】

つまり、従来型のアップサンプリングとは対照的に、本発明のアップサンプリング方法では、高解像度ビューを使用することによって再構成される深度マップにおける明確な対象境界を再構成する。これによって、対象のエッジがより良く維持されるので、比較的低いビットレートにおける符号化効率が改善されるとともにレンダリング品質が向上する。

【0017】

したがって、アップサンプリングされた深度マップに遮断作用又は平滑作用を及ぼさないアップサンプリング方法を低いビットレートで実現することができる。

基準画像取得工程において取得した基準画像、アップサンプリング工程においてアップサンプリングした深度マップ及び仮想視点情報取得工程において取得した仮想視点の位置情報と方向情報に基づいて、仮想視点の位置情報及び方向情報情報に対応する仮想視点画像を合成する仮想視点画像合成工程と、

により、仮想視点画像を合成することを特徴とする仮想視点画像合成方法。

【0018】

請求項2に記載の発明は、請求項1に記載の仮想視点画像合成方法において、画像深度取得手段(402, 409, 412)は、画像の深度を検出する深度カメラ(402)であり、アップサンプリング工程は、深度カメラ(402)から入力された深度マップを入力し、深度マップ生成工程において生成された深度マップにおける一組の隣接するピクセルと、基準画像における互いに隣接していないピクセルとを対応させ、深度マップにおける一組の隣接するピクセルにおける各ピクセルに重みを割り当て、一組の隣接するピクセルに割り当てた各ピクセルの重みを最適化し、最小重みを算出し、算出した最小重みを得たときの一組の隣接するピクセルにおける最適な深度値を選択することを特徴とする仮想視点画像合成方法である。

【0019】

請求項3に記載の発明は、請求項2に記載の仮想視点画像合成方法において、重みは、基準画像のピクセルと深度マップにおける一組の隣接するピクセルとの色又は強度差及び距離に基づいて割り当てられることを特徴とする仮想視点画像合成方法である。

【0020】

請求項4に記載の発明は、請求項2に記載の仮想視点画像合成方法において、重みは、深度カメラ(402)から入力された深度マップ及び/又は基準画像のピクセルと、深度カメラ(402)から入力された深度マップ及び/又は基準画像における一組の隣接するピクセルとの色又は強度差及び距離の組合せに基づき割り当てられることを特徴とする仮想視点画像合成方法である。

【0021】

請求項5に記載の発明は、請求項2～請求項4のいずれか1項に記載の仮想視点画像合成方法において、重みの最適化は、権利独占選択によることを特徴とする仮想視点画像合成方法である。

【0022】

10

20

30

40

50

請求項 6 に記載の発明は、請求項 2 ~ 請求項 5 のいずれか 1 項に記載の仮想視点画像合成方法において、最適深度値の選択は、出力深度値として最小重みを有するピクセルの深度を選択することによることを特徴とする仮想視点画像合成方法である。

【0023】

請求項 7 に記載の発明は、複数の視点から取得した画像情報に基づき、仮想視点から観た仮想視点画像を合成する仮想視点画像合成システム(100)であって、複数の視点に配置した画像取得手段(10)と、複数の画像取得手段(10)から画像構成の基準となる基準画像を取得する基準画像取得手段(102)と、複数の画像取得手段(10)と同じ視点における画像の深度を取得する画像深度取得手段(402, 409, 412)と、画像深度取得手段(402, 409, 412)により取得した深度に基づき、複数の画像取得手段(10)と同じ視点における画像の深度マップを生成する深度マップ生成手段(102)と、深度マップ生成手段(102)において生成した深度マップをアップサンプリングするアップサンプリング手段(401)と、画像を合成する際の視点となる仮想視点の位置情報と仮想視点から合成画像を観たときの方向である仮想視点方向の情報を取得する仮想視点情報取得手段と、基準画像取得手段(102)において取得した基準画像、アップサンプリング手段(401)においてアップサンプリングした深度マップ及び仮想視点情報取得手段において取得した仮想視点の位置情報及び方向情報に基づいて、仮想視点の位置情報及び方向情報に対応する仮想視点画像を合成する仮想視点画像合成手段(104)と、を備えたことを特徴とする仮想視点画像合成システム(100)である。

10

【0024】

このような仮想視点画像合成システム(100)によれば、請求項 1 に記載の仮想視点画像合成方法と同様な効果を得ることができる仮想視点画像合成システム(100)とすることができる。

20

【0025】

請求項 8 に記載の発明は、請求項 7 に記載の仮想視点画像合成システム(100)において、画像深度取得手段(402, 409, 412)は、画像の深度を検出する深度カメラ(402)であり、アップサンプリング手段(401)は、深度カメラ(402)から入力された深度マップを入力する深度マップ入力手段(601)と、深度マップ入力手段(601)で入力した深度マップにおける一組の隣接するピクセルと、基準画像における互いに隣接していないピクセルとを対応させる対応手段と、深度マップにおける一組の隣接するピクセルにおける各ピクセルに重みを割り当てる重み割り当て手段と、重み割り当て手段で一組の隣接するピクセルに割り当てた各ピクセルの重みを最適化し、最小重みを算出する最小重み算出手段と、最小重み算出手段で算出した最小重みを得たときの一組の隣接するピクセルにおける最適な深度値を選択する最適深度値選択手段と、を備えたことを特徴とする仮想視点画像合成システム(100)である。

30

【0026】

請求項 9 に記載の発明は、請求項 8 に記載の仮想視点画像合成システム(100)において、重み割り当て手段は、基準画像のピクセルと、深度マップ入力手段(601)で入力した深度マップにおける一組の隣接するピクセルと、の色又は強度差及び距離に基づいて重みを割り当てることを特徴とする仮想視点画像合成システム(100)である。

40

【0027】

請求項 10 に記載の発明は、請求項 8 に記載の仮想視点画像合成システム(100)において、重み割り当て手段は、深度マップ入力手段(601)で入力した深度マップ及び/又は基準画像のピクセルと、深度マップ入力手段(601)で入力した深度マップにおける一組の隣接するピクセル及び/又は基準画像との色又は強度差及び距離の組合せに基づき、重みを割り当てることを特徴とする仮想視点画像合成システム(100)である。

【0028】

請求項 11 に記載の発明は、請求項 8 ~ 請求項 10 のいずれか 1 項に記載の仮想視点画像合成システム(100)において、最小重み算出手段は、権利独占選択により、重みの最適化を行うことを特徴とする仮想視点画像合成システム(100)である。

50

【 0 0 2 9 】

請求項 1 2 に記載の発明は、請求項 8 ~ 請求項 1 1 のいずれか 1 項に記載の仮想視点画像合成システム (1 0 0) において、最適深度値選択手段は、出力深度値として最小重みを有するピクセルの深度を選択することを特徴とする仮想視点画像合成システム (1 0 0) である。

【 図面の簡単な説明 】

【 0 0 3 0 】

【 図 1 】 例示的な N ビュー及び N 深度に基づく 3 D システムのシステム図である。

【 図 2 】 ダウンサンプリングの概略図である。

【 図 3 】 3 つのビュー及び対応する深度マップの例である。

【 図 4 】 アップサンプル部の概略の内部構成を示す図である。

【 図 5 】 アップサンプル部におけるアップサンプリングの概略図を示す

【 図 6 】 アップサンプリング処理の流れを示すフローチャートである。

【 図 7 】 アップサンプリング処理中のサブルーチンであるアップサンプル深度処理の流れを示すフローチャートである。

【 図 8 】 アップサンプル部におけるアップサンプリングの結果を示す。

【 図 9 】 ビデオシーケンス「 Champagne _ tower 」に関する PSNR 比較である。

【 図 1 0 】 複数のビュー及び形状に基づく多視点ビデオシステムの一般化されたシステム図である。

【 発明を実施するための形態 】

【 0 0 3 1 】

以下、本発明が適用された実施形態について図面を用いて説明する。なお、本発明の実施の形態は、下記の実施形態に何ら限定されることはなく、本発明の技術的範囲に属する限り種々の形態を採りうる。

【 0 0 3 2 】

図 1 は、仮想視点画像合成システム 1 0 0 の概略の構成を示すシステム図である。仮想視点画像合成システム 1 0 0 は、図 1 に示すように、多視点撮影サブシステム 1 0 1 、形状生成サブシステム 1 0 2 、送信 / 記憶サブシステム 1 0 3 及び仮想ビュー合成サブシステム 1 0 4 を備えている。

【 0 0 3 3 】

なお、本実施形態における、形状生成サブシステム 1 0 2 及びアップサンプリングサブシステム 4 0 1 では、単に一つの処理経路を有するのではなく、アプリケーションの違いによって、種々の経路 (以下、アプリケーション経路とも呼ぶ) を採ることができるようになっている。

【 0 0 3 4 】

つまり、スイッチ 4 0 7 及びスイッチ 4 0 8 を切り替えることにより、4 つのアプリケーションが経路「 a 」 ~ 「 d 」を選択し、アプリケーションの違いによって、アップサンプリングサブシステム 4 0 1 における処理内容を異なるものとすることができるようになっている。

【 0 0 3 5 】

なお、スイッチ 4 0 7 , 4 0 8 は、常に同じ経路を選択し、例えば、両方を経路「 a 」に設定するかあるいは両方を経路「 b 」に設定するといった具合である。

多視点撮影サブシステム 1 0 1 は、図示しない複数のカメラにより静的又は動的のいずれかの画像を取得するマルチカメラ撮像システム 1 0 を含んでいる。

【 0 0 3 6 】

マルチカメラ撮像システム 1 0 で撮影されたビデオ画像は、較正又は修正によって前処理され、多視点ビデオ画像 1 0 5 として、形状生成サブシステム 1 0 2 及び送信 / 記憶サブシステム 1 0 3 に出力される。

【 0 0 3 7 】

多視点撮影サブシステム 101 における較正又は修正の方法は、従来の技術によればよいため、説明を省略する。

形状生成サブシステム 102 は、深度マップ 413 を生成し、送信 / 記憶サブシステム 103 へ出力する。

【0038】

深度マップ 413 の生成は、4つの方法（4つのアプリケーション経路「a」～「d」に対応）によって行われる。

アプリケーション経路「a」では、深度カメラ 402 で取得した深度画像をその深度マップ 413 として出力する。

【0039】

アプリケーション経路「b」では、深度カメラ 402 で取得した深度画像をアップサンプル部 403 でアップサンプリングして出力する。このアップサンプリングについては、詳細を後述する。

【0040】

ここで、アプリケーション経路「a」、「b」で用いられる深度カメラ 402 は、画像をその深度とともに取得することができるカメラであり、マルチカメラ撮像システム 10 で撮影する多視点ビデオ画像 105 と同じ画像の深度を取得する。

【0041】

通常、深度カメラ 402 からの深度マップは、撮影された多視点ビデオ画像 105 よりも解像度が低い。

アプリケーション経路「c」では、多視点撮影サブシステム 101 から入力される多視点ビデオ画像 105 から、深度推定部 409 において多視点ビデオ画像 105 の深度が推定された深度マップが生成され、生成された深度マップがダウンサンプル部 410 でダウンサンプルされて出力される。

【0042】

アプリケーション経路「c」では、このように、深度推定部 409 からの高解像度の深度マップがダウンサンプル部 410 においてダウンサンプルされることによって符号化効率を向上させることができる。

【0043】

深度推定部 409 における深度推定は複雑であるので、深度を比較的低い解像度で推定し、この後に本発明の低複雑度のアップサンプリング方式が続くようにする方がはるかに速い。これには深度マップの品質が僅かに低下するという犠牲を伴うが、このアプリケーションは実時間深度推定が要求される場合に有用である。

【0044】

アプリケーション経路「c」におけるダウンサンプル部 410 では、ダウンサンプリング中に深度境界を維持するために、下記式 1 で定義されるピクセルドロッピングによってダウンサンプリングを適用する。

【0045】

【数 1】

$$Depth_{down}(x, y) = Depth_{in}(x \cdot step, y \cdot step) \quad \dots \text{式 1}$$

【0046】

ここで、step は受信機側に送信されるパラメータであるダウンサンプリング係数である。図 2 は、このダウンサンプリングの概略図を示す。

このダウンサンプリングでは、規則的な格子上のすべてのサンプルは低解像度の深度マップに記憶され、格子から外れたピクセルは捨てられる。図 2 では、文字 A ~ K で示されたピクセルの深度値のみがダウンサンプリングされた深度マップに記憶され、他の深度値は捨てられる。例えば、ピクセル 501 は、ピクセル 503 にマップするが、ピクセル 502 は捨てられる。

10

20

30

40

50

【 0 0 4 7 】

アプリケーション経路「d」では、多視点撮影サブシステム101から入力される多視点ビデオ画像105がダウンサンプル部411においてダウンサンプリングされ、深度推定部412において、ダウンサンプリングされた多視点ビデオ画像の深度が推定されて、深度マップが生成され、出力される。

【 0 0 4 8 】

仮想視点画像合成システム100においては、3Dシーンの情報は多くの方法で表現され得る。よく使用される3Dシーン表現は、Nビュー及びN深度マップ画像に基づくものであるため、図3に、形状生成サブシステム102において得られる、N=3の場合のNビュー及びN深度マップ画像の例を示す。

10

【 0 0 4 9 】

図3は、3つの深度ビュー201~203及び対応する深度マップ204~206を示す。3つの深度ビュー201~203は、並べた3台の深度カメラ402で撮影されている。

【 0 0 5 0 】

深度マップ204~206は、所定の深度範囲内の深度カメラ402に対するパーピクセル深度を表すグレースケール画像である。8ビットスケールでは、グレースケール値0(黒)は最大距離を表し、255(白)は最小距離を表す。

【 0 0 5 1 】

深度マップ204~206は、対象内で滑らかに変化する強度と対象エッジにおける急峻な変化を含む傾向がある(例えば、図3参照)。したがって、深度マップ204~206は、深度のダウンサンプリング及びアップサンプリングによって利用され得るテクスチャビデオに比べてより大きな空間的冗長性を有する傾向がある。

20

【 0 0 5 2 】

深度マップ204~206は、対象境界付近の重要なエッジ情報を伝達する。最近隣法(nearest neighbor)、線形補間、又は多項式補間など、従来型のアップサンプリング方法は、この重要な情報を平滑化するか又は除去する。これは合成されたビューのアーチファクトにつながる。

【 0 0 5 3 】

送信/記憶サブシステム103は、多視点ビデオ画像105及び深度マップ413を記憶又は送信する。送信/記憶サブシステム103では、多視点ビデオ画像105及び深度マップ413を圧縮する場合も圧縮しない場合もあり、チャンネルは単にワイヤーであってよい。

30

【 0 0 5 4 】

したがって、送信/記憶サブシステム103から出力される復号化ビュー414及び復号化深度マップ416は、それぞれ多視点ビデオ画像105及び深度マップ413と全く同じであってよい。

【 0 0 5 5 】

また、多視点ビデオ画像105及び深度マップ413を圧縮するには、標準的な最新式のビデオエンコーダ、例えば、AVC/H.264、MVC、MPEG2、MJPEGなどを使用してもよい。

40

【 0 0 5 6 】

アップサンプリングサブシステム401は、スイッチ407、アップサンプル部404~406及びワイヤーを備えている。スイッチ407は、前述のようにアプリケーション経路を選択するスイッチであり、ワイヤーは、送信/記憶サブシステム103からの入力を、スイッチ407を介して、直接仮想ビュー合成サブシステム104へ出力するためのものである。

【 0 0 5 7 】

また、アップサンプル部406は、経路「d」で示されるように、実時間深度推定が要求されるアプリケーションに使用され得る。多視点ビデオ画像105は、ダウンサンプル

50

部 4 1 1 によってダウンサンプリングされ、深度は、深度推定部 4 1 2 によって比較的低い解像度で推定され、この後にアップサンプル部 4 0 6 が用いられている。

【 0 0 5 8 】

ここで、アップサンプリングサブシステム 4 0 1 のアップサンプル部 4 0 4 ~ 4 0 6 及び形状生成サブシステム 1 0 2 におけるアップサンプル部 4 0 3 について、図 4 ~ 図 8 に基づいて説明する。

【 0 0 5 9 】

図 4 は、アップサンプル部 4 0 3 ~ 4 0 6 の概略の内部構成を示す図である。アップサンプル部 4 0 3 ~ 4 0 6 の各々は、図 4 に示すように、画像入力ユニット 6 0 1、重み計算ユニット 6 0 2、最小値最適化ユニット 6 0 3、深度選択ユニット 6 0 4 及び後処理ユニット 6 0 5 を備えている。

10

【 0 0 6 0 】

そして、多視点ビデオ画像 1 0 5 及び深度マップ 4 1 3 の集合から、画像入力ユニット 6 0 1 が入力ビュー 6 0 6 及び対応する深度マップ 6 0 7 を選択する。

アップサンプリングを必要とする各ピクセル u に対して、重み計算ユニット 6 0 2 が、ピクセル u と一組の隣接するピクセルとの色差又は輝度差及びこれらの間の距離に基づいて一組の重み 6 0 8 を計算する。

【 0 0 6 1 】

最小値最適化ユニット 6 0 3 が最小重み 6 0 9 のピクセルを選択し、対応する出力深度マップ 6 1 0 が深度選択ユニット 6 0 4 によって選択される。

20

オプションとして、後処理ユニット 6 0 5 が出力深度 6 1 1 を生成するためにフィルタリング（例えば、メジアンフィルタリング）などの後処理動作を適用してもよい。

【 0 0 6 2 】

図 5 は、アップサンプル部 4 0 3 ~ 4 0 6 におけるアップサンプリングの概略図を示す。入力画像は、最高解像度の入力ビュー 6 0 6（アップサンプル部 4 0 3 では、多視点ビデオ画像 1 0 5 に該当）と対応する低解像度の深度マップ 6 0 7（アップサンプル部 4 0 3 では、深度マップ 4 1 3 に該当）とからなる。入力ビュー 6 0 6 のサンプル格子と深度マップ 6 0 7 のサンプル格子の間の倍率は、パラメータステップ 7 0 9 によって定義される。

【 0 0 6 3 】

深度マップ 6 0 7 内のピクセル 7 0 1 「A」は、入力ビュー 6 0 6 のピクセル 7 0 5 「A」に対応し、ピクセル 7 0 1 の深度値は出力深度マップ 6 1 0 のピクセル 7 0 3 にコピーされる。

30

【 0 0 6 4 】

深度マップ 6 0 7 で定義される値を持たない出力深度マップ 6 1 0 のピクセル、例えば、出力深度マップ 6 1 0 のピクセル 7 0 8 「u」の場合、深度は隣接するピクセルから推定される。

【 0 0 6 5 】

隣接するピクセルの集合 N は予め定義される。例として、 N は入力ビュー 6 0 6 において図 5 で A、B、C、D と標記される 4 つの隣接するピクセルからなり、下記式 2 で定義される。

40

【 0 0 6 6 】

【 数 2 】

$$N \in \{A, B, C, D\} \quad \dots \text{式 2}$$

【 0 0 6 7 】

このとき、集合 N の各隣接するピクセルに対して、重みが、集合 N 内のピクセル u と隣接するピクセルとの間の色差と、ピクセル距離とに基づいて、下記式 3、式 4 及び式 5 により計算される。

【 0 0 6 8 】

50

【数 3】

$$W_N(u) = f(u, N)g(u, N) \quad \dots \text{式 3}$$

【0069】

【数 4】

$$f(u, N) = \|u - N\| \quad \dots \text{式 4}$$

【0070】

【数 5】

$$g(u, N) = \|C_u - C_N\| \quad \dots \text{式 5}$$

10

【0071】

式 3 ~ 5 中、関数 $f(\quad)$ はピクセル距離であり、関数 $g(\quad)$ は集合 N 内のピクセル u と隣接するピクセルとの色差である。 C_u はピクセル u の色を示し、 C_N は集合 N 内のピクセルの色を示す。

【0072】

ここでは、 $\|\cdot\|$ は任意のノルム（例えば、 L_1 又は L_2 ノルム）であり得るし、色空間 C は任意の空間（例えば、 RGB 、 YUV 、 HSI 、グレースケールなど）であり得る。

20

【0073】

さらに、境界付近のノイズを抑制するために、例えば、下記式 6 及び式 7 などのように平滑項が式 3 に追加され得る。

【0074】

【数 6】

$$W_N(u) = f(u, N)g(u, N) + V_N \quad \dots \text{式 6}$$

【0075】

【数 7】

$$V_N = \lambda_1 |d_N - d_{left}| + \lambda_2 |d_N - d_{up}| \quad \dots \text{式 7}$$

30

【0076】

式 6、7 中、 λ_1 及び λ_2 は平滑項の強度を制御するパラメータであり、 d_N は集合 N におけるピクセルの深度であり、 d_{left} 及び d_{up} は、それぞれ、ピクセル u の左及びピクセル u の上のピクセルの深度値である。

【0077】

ピクセル 98「 u 」の出力深度の深度値は、下記式 8 のように、獲得深度 ($winning\ depth$) を選択するために式 3 又は式 6 の $W_N(u)$ を最適化することによって決定される。

40

【0078】

【数 8】

$$depth_u = \arg \min(W_N(u)) \quad \dots \text{式 8}$$

【0079】

(アップサンプリング処理)

次に、アップサンプル部 403 ~ 406 の各ユニット 601 ~ 605 の処理を CPU で

50

実行する場合のアップサンプリング処理について説明する。図6は、アップサンプリング処理の流れを示すフローチャートであり、図7は、アップサンプリング処理中のサブルーチンであるアップサンプル深度処理の流れを示すフローチャートである。

【0080】

図6に示すように、アップサンプリング処理では、ステップS800において、入力ビュー606及び深度マップ607が入力される。

続くステップS805において、入力ビュー606のうち、図5において左上隅に位置するピクセル705が選択される。以下、ピクセルの選択は、左上隅で始まり、右下隅で終わる従来のスキニング順（ラスタースキャン）に処理される。

【0081】

続くステップS810において、入力ビュー606におけるピクセルuの位置が、ピクセル701の深度の比較的低い解像度格子にマップされる。

続くS815では、ピクセル701の深度の値がその位置で定義されるか否かが判断される。ピクセル701の深度の値がその位置で定義される場合（S815：Yes）、処理がステップS820へ移行され、定義されない場合（S815：No）には、処理がアップサンプル深度処理ステップS900へ移行される。

【0082】

ステップS820では、深度マップからピクセル701の深度の値がコピーされる。深度値が定義されない場合（S815：No）、深度値は、後述するアップサンプル深度処理によって決定される。

【0083】

その後、ステップS825において、入力ビュー606内の次のピクセルuを選択し、続くステップS830において、入力ビュー606のすべてのピクセルについて、ステップS810～S825処理が終了したか否かが判定される。

【0084】

全ピクセルに対してステップS810～S825が終了していないと判定された場合（S830：No）、処理がステップS810に戻され、全ピクセルに対してステップS810～S825の処理が実行される。また、全ピクセルに対してステップS810～S825が終了していると判定された場合は（S830：Yes）、本処理が終了となる。

【0085】

この処理の出力はアップサンプリングされた出力深度マップ610である。

次に、アップサンプル深度処理ステップS900について図7に基づき説明する。アップサンプル深度処理は、図7に示すように、ステップS905において、隣接するピクセルNの集合及び深度マップ607における対応する深度値が決定される。

【0086】

この後、ステップS910～S925において、すべてのピクセル（ $i = 0 \sim N$ ）について、重みが前述した式3～式7を用いて計算される。

そして、ステップS930において、集合Nにおける最小重みが決定される。隣接する獲得ピクセル（winning neighboring pixel）の深度値がステップS935において決定される。

【0087】

このような、アップサンプル部403～406における入力ビュー606に基づくアップサンプリング方法は、急峻な対象境界を維持し再構成することができる。これは、低解像度深度マップをアップサンプリングするときに対応する高解像度カラービュー画像が使用されるからである。

【0088】

図8は、アップサンプリングの結果を示す。図8（a）は最初の深度マップを示す。この後、最初の深度は、ダウンサンプリングされ、さらに、2つの従来の方法及び本発明の方法でアップサンプリングされる。

【0089】

10

20

30

40

50

図 8 (b) は、従来型の最近隣アップサンプリングの結果である。図 8 (c) は、従来型の双線形アップサンプリングの結果である。図 8 (d) は、アップサンプル部 4 0 3 ~ 4 0 6 の出力画像であり、これは平滑作用も遮断作用もさほどなくかつより良く再構成された対象エッジを鮮明に示す。

【 0 0 9 0 】

仮想ビュー合成サブシステム 1 0 4 では、形状生成サブシステム 1 0 2 において取得した基準画像、アップサンプリングサブシステム 4 0 1 においてアップサンプリングした深度マップ及びマルチカメラ撮像システム 1 0 を構成する各カメラの位置情報と方向情報に基づいて、仮想視点情報に対応する仮想視点画像を合成し、仮想ビュー 1 0 9 として出力する。

10

【 0 0 9 1 】

ここで、すべてのアプリケーション経路 (「 a 」 ~ 「 d 」) において、アップサンプル部 4 0 3 ~ 4 0 6 への入力、ビュー画像、及びビュー画像よりも低い解像度の対応深度マップである。再構成された深度マップ 4 1 7 及び復号化ビュー 4 1 4 は、新規なビューをレンダリングするために、仮想ビュー合成サブシステム 1 0 4 で使用される。

(仮想視点画像合成システム 1 0 0 の特徴)

以上に説明した仮想視点画像合成システム 1 0 0 によれば、アップサンプリングされた深度マップに強力な遮断作用又は平滑作用を及ぼさず、合成されたビューにアーチファクトをもたらさないものとなる。したがって、符号化効率の改善及びビュー合成品質の向上に寄与する。

20

【 0 0 9 2 】

つまり、従来型のアップサンプリングとは対照的に、仮想視点画像合成システム 1 0 0 のアップサンプリング方法では、高解像度ビューを使用することによって再構成される深度マップにおける明確な対象境界を再構成する。これによって、対象のエッジがより良く維持されるので、比較的低いビットレートにおける符号化効率が改善されるとともにレンダリング品質が向上する。

【 0 0 9 3 】

したがって、アップサンプリングされた深度マップに遮断作用又は平滑作用を及ぼさないアップサンプリング方法を低いビットレートで実現することができる。

ここで、本発明の方法の長所の証拠として、ビデオシーケンス「 Champagne _ tower 」に関する PSNR 比較を図 9 示す。

30

(その他の実施形態)

以上、本発明の実施形態について説明したが、本発明は、本実施形態に限定されるものではなく、種々の態様を採ることができる。

【 0 0 9 4 】

例えば、本実施形態では、アップサンプリングサブシステム 4 0 1 は、送信 / 記憶サブシステム 1 0 3 の後段で用いたが、アプリケーションに応じて送信 / 記憶サブシステム 1 0 3 の前に使用してもよい。

【 符号の説明 】

【 0 0 9 5 】

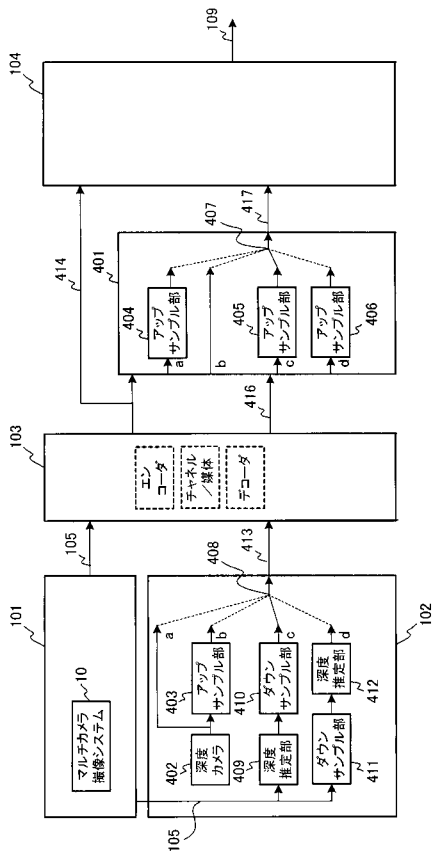
1 0 ... マルチカメラ撮像システム、 1 0 0 ... 仮想視点画像合成システム、 1 0 1 ... 多視点撮影サブシステム、 1 0 2 ... 形状生成サブシステム、 1 0 3 ... 送信 / 記憶サブシステム、 1 0 4 ... 仮想ビュー合成サブシステム、 1 0 5 ... 多視点ビデオ画像、 1 0 6 ... 形状情報、 1 0 9 ... 仮想ビュー、 2 0 3 ... 深度ビュー、深度マップ 2 0 4 ~ 2 0 6、 4 0 1 ... アップサンプリングサブシステム、 4 0 2 ... 深度カメラ、 4 0 3 ~ 4 0 6 ... アップサンプル部、 4 0 7、 4 0 8 ... スイッチ、 4 0 9 ... 深度推定部、 4 1 0、 4 1 1 ... ダウンサンプル部、 4 1 2 ... 深度推定部、 4 1 3 ... 深度マップ、 4 1 4 ... 復号化ビュー、 4 1 6 ... 復号化深度マップ、 4 1 7 ... 深度マップ、 5 0 1 ~ 5 0 3 ... ピクセル、 6 0 1 ... 画像入力ユニット、 6 0 2 ... 重み計算ユニット、 6 0 3 ... 最小値最適化ユニット、 6 0 4 ... 深度選択ユニット、 6 0 5 ... 後処理ユニット、 6 0 6 ... 入力ビュー、 6 0 7 ... 深度マップ、 6

40

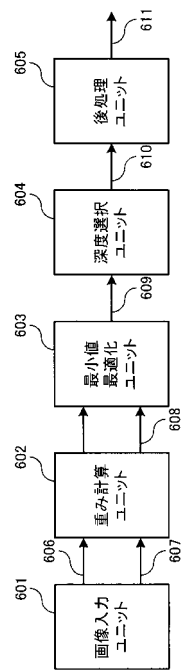
50

10 ...出力深度マップ、611...出力深度、701,703,705...ピクセル、708...深度値、709...パラメータステップ。

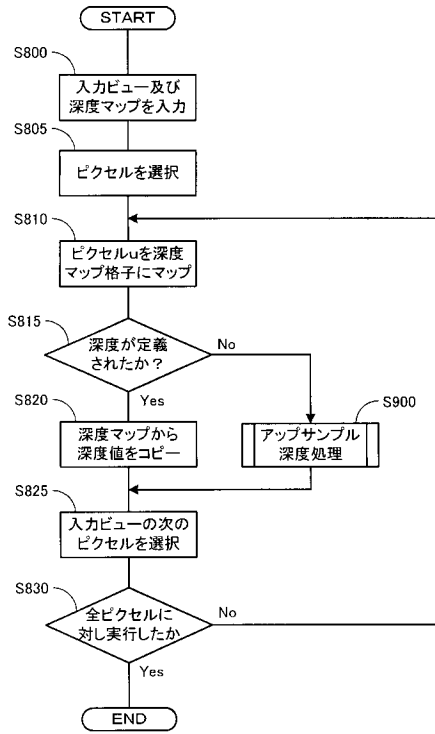
【図1】



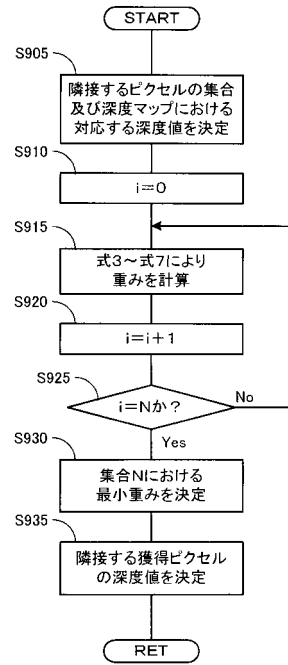
【図4】



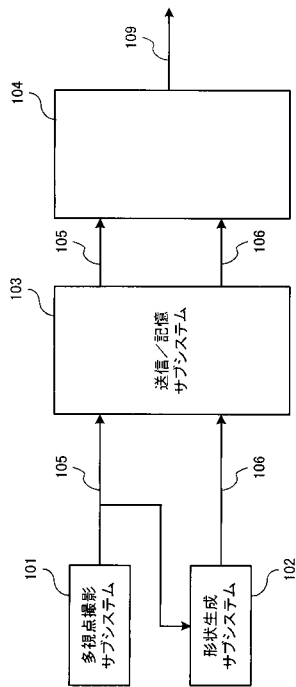
【 図 6 】



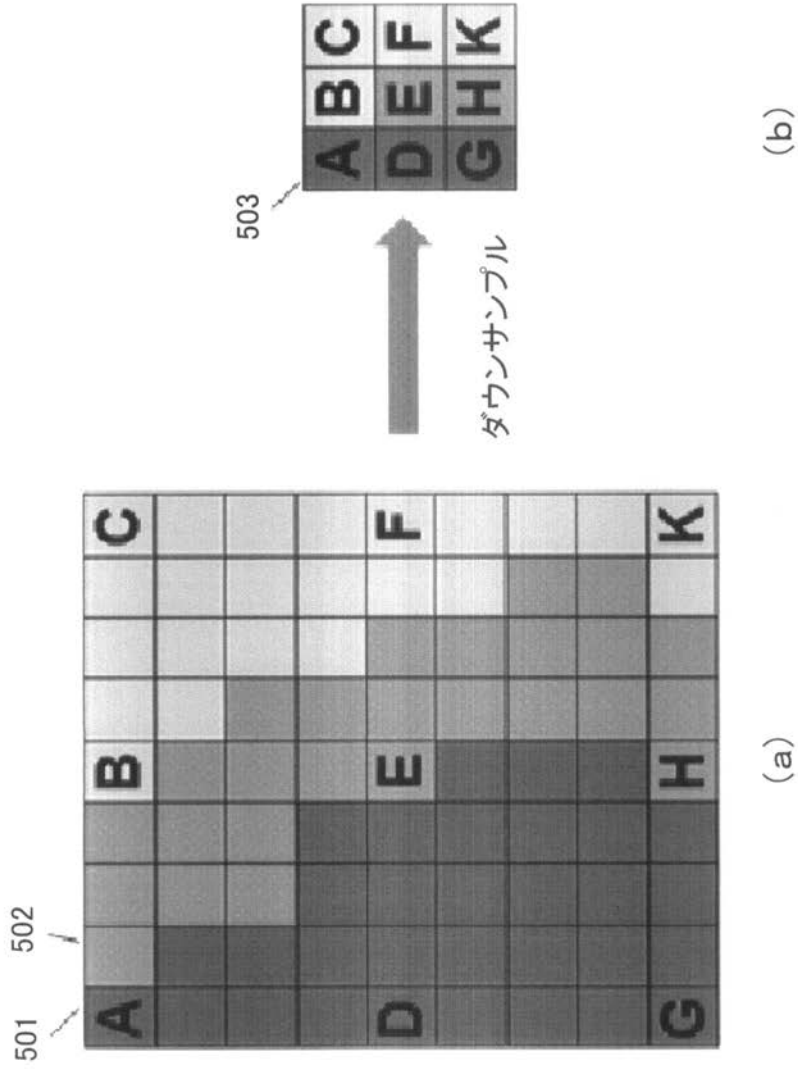
【 図 7 】



【 図 10 】



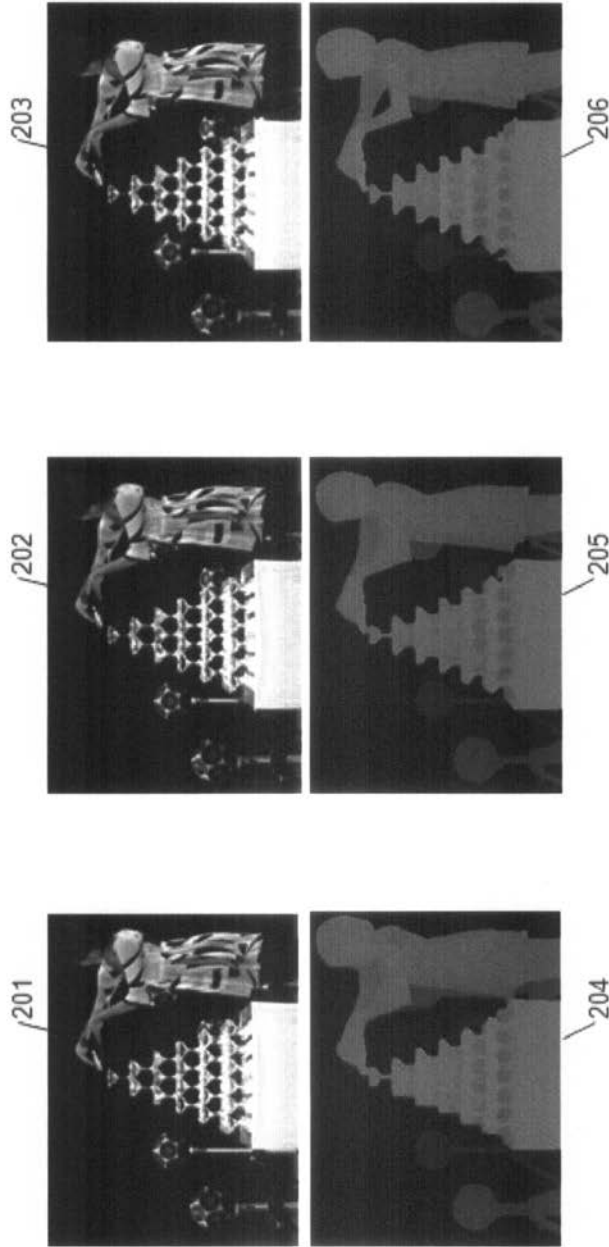
【 図 2 】



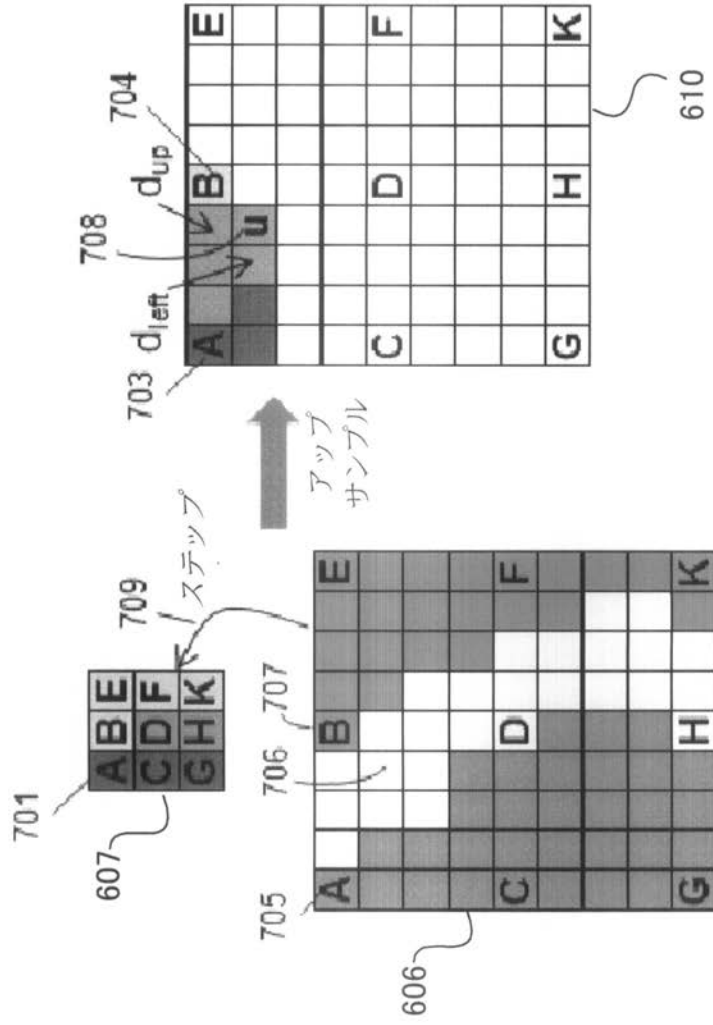
(a)

(b)

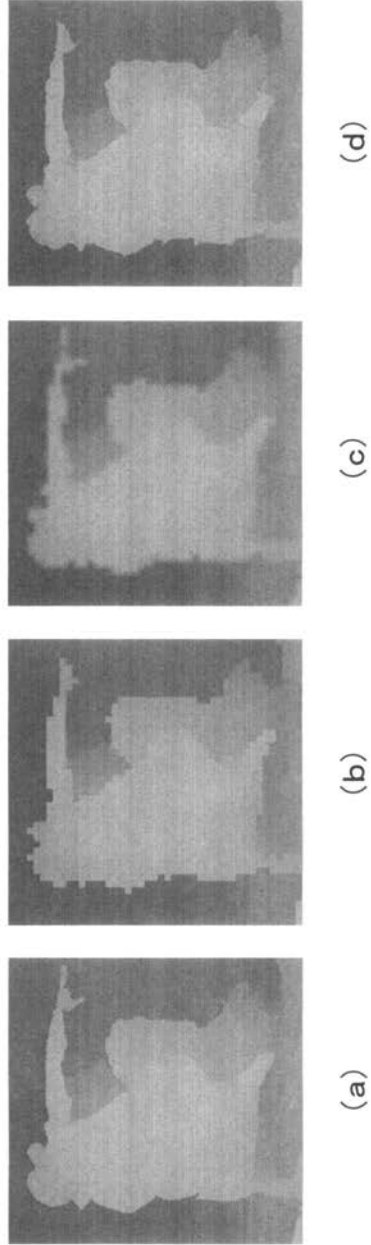
【 図 3 】



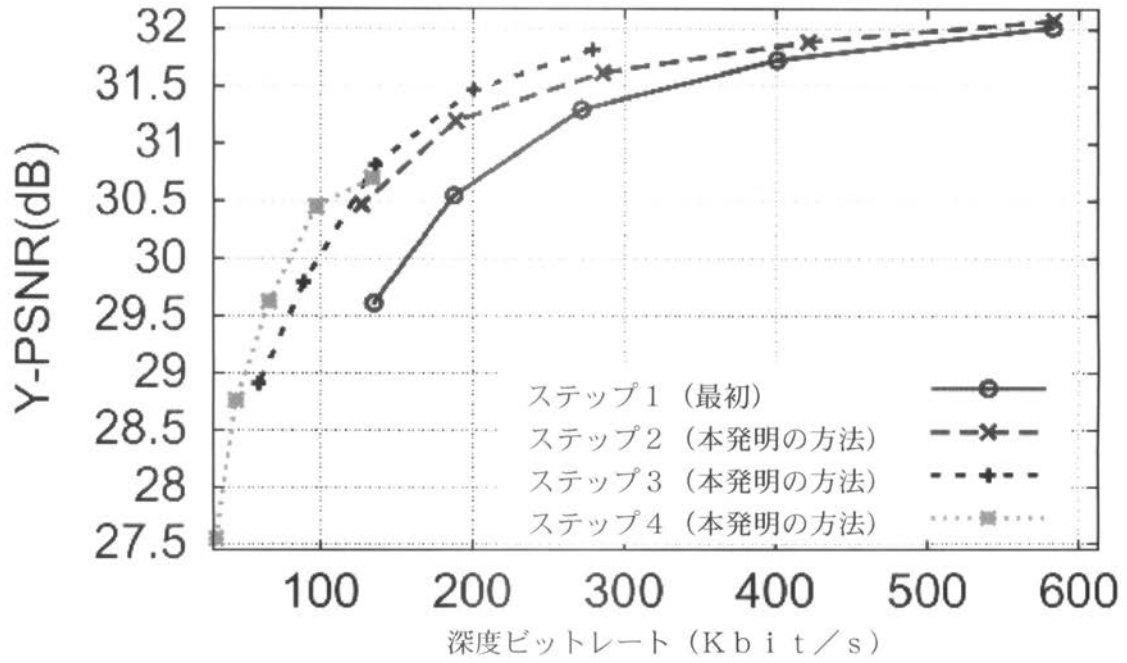
【 図 5 】



【 図 8 】



【図9】



フロントページの続き

(72)発明者 圓道 知博

愛知県名古屋市千種区不老町1番 国立大学法人名古屋大学内

(72)発明者 谷本 正幸

愛知県名古屋市千種区不老町1番 国立大学法人名古屋大学内

Fターム(参考) 5B050 AA08 BA09 DA01 DA04 DA07 EA09 EA17 EA19 EA26 FA02
FA05
5C061 AA29 AB12 AB14 AB17
5C159 KK37 LB05 LB15 MB16 NN41 PP03 PP13 RC19 TA06 TA08
TB10 TC35 TC42 UA02 UA05