

インタラクティブコンテンツの生成 パノラマ画像と3Dモデリング

(正会員) 蚊野 浩†

1. ま え が き

銀塩写真やビデオテープによる画像・映像と比較して、メディアがデジタル化されたことによって、保存・加工・編集・伝送・共有が、簡単・高速・劣化なしに行えるようになった。一方で、画像・映像コンテンツは、基本的に、撮影した画像・映像をディスプレイやハードコピーに出力して鑑賞するのであって、コンテンツ自身が特段変わったということはない。

一方で、顔検出などの画像認識技術や3次元シーンの解析技術が、民生機器での応用において実用レベルを超えたことで、コンテンツ生成においても新機能が提供されつつある。今回の講座では、通常のカメラでは撮影不可能な広視野パノラマ画像を生成する技術と、物体を全周方向から撮影した画像から、その3DCG (3 Dimensional Computer Graphics) データを自動的に生成する3Dモデリングについて解説する。従来の画像・映像が、そのまま楽しむ受動的なメディアであるのに対して、パノラマ画像と3Dモデルは、ユーザからのインタラクションによって画像が多様に変化するインタラクティブなコンテンツであり、デジタルならではのコンテンツと言える。

2. パノラマ画像合成

2.1 パノラマ画像の概要

パノラマ画像は、広角レンズを利用しても一度に撮影できない、視野の広い画像のことである。カメラの視野角は28mmレンズで75°、21mm超広角レンズで91°程度であることを考慮すると、90°以上の視野の画像はパノラマ画像といえる。

パノラマ画像合成技術は、1枚の画像として撮影するには視野が広すぎる画像を複数の画像に分割撮影し、後処理によって、あたかも1枚の画像であるかのように継ぎ目なく

合成する技術である。もっとも簡単にパノラマ画像を合成する方法は、大きな被写体を手持ちカメラで複数に分割撮影し、画像の重なり部分を手作業で位置合わせし、ブレンディング合成する方法である。この方法は、いかに入念に位置合わせを行っても、多くの場合、若干の位置ずれが生じる。パノラマ画像合成技術の重要なポイントは、位置ずれの生じない撮影条件と画像の幾何変形の性質を明らかにすることである^{4) 10) 12)}。

代表的なパノラマ合成用画像の撮影方法として、円筒パノラマ、平面パノラマなどが挙げられる。円筒パノラマとは、三脚などを用いてカメラの動きを水平回転だけに固定する方法である²⁾。360°全周方向の画像を撮影し、隣り合うすべての画像を接続・合成して、完全に円筒状にしたものを全周パノラマと呼ぶ。この拡張として球体パノラマおよび全天パノラマを考えることができる。

円筒パノラマのようにカメラの投影中心を移動させないことは、被写体によらず合成時に位置ずれを生じない撮影条件であるが、被写体が平面的な場合は、異なった場所で撮影した複数の画像であっても、共通平面に画像を投影することで、位置ずれなくパノラマ画像を合成することができる¹⁵⁾。このパノラマ画像合成手法を平面パノラマと呼ぶ。そのほかの手法として、時間的に連続して撮影した画像を利用できれば、連続画像のストリップ(細長い長方形領域)をつなぎ合わせることで、奥行きのあるシーンでも、横長・縦長パノラマ画像を生成することができる。ラインセンサーで3次元シーンの画像入力をするようなものであるが、この方式をストリップパノラマと呼ぶ。以下、これらの技術について説明する。

2.2 円筒パノラマ

図1に全周パノラマ画像の例を示す。図の左端と右端はつながっており360°の円筒画像である。全周パノラマを作るには広角レンズを使っても10枚近い画像を合成する必要がある。

カメラを水平回転させ、2枚の画像が重なりを持つように撮影した状況を図2(a)に示す。重なり部分に、図のように光線が入射する場合、2枚の撮像面は同じ光線を記録する

†三洋電機株式会社 研究開発本部 デジタル技術研究所
"Technologies Leading to the Next-generation Digital Cameras and Movies (11): Generation of Interactive Contents: Panoramic Images and 3D Modeling" by Hiroshi Kano (SANYO Electric Co., Ltd., Osaka)



図1 全周パノラマ画像の例

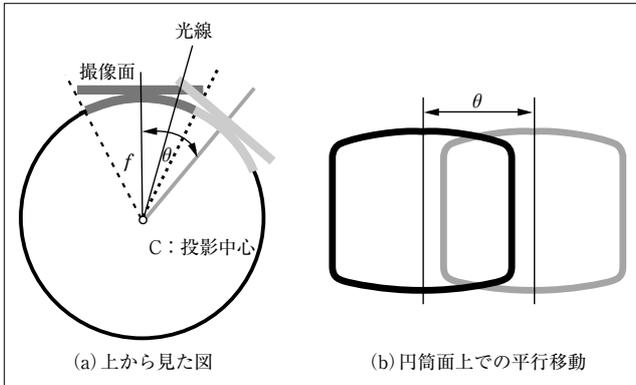


図2 円筒パノラマの原理図

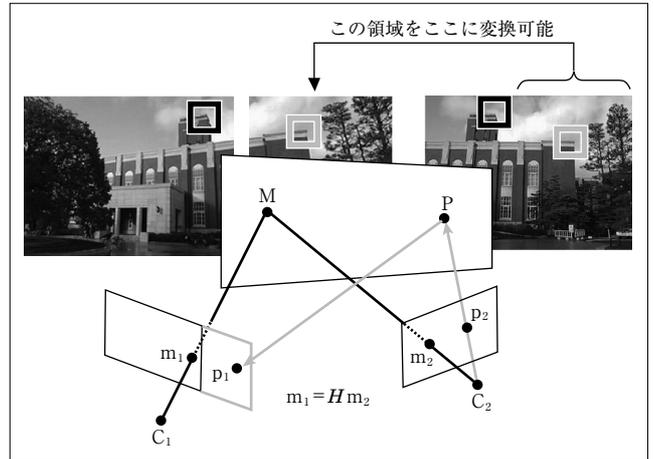


図3 平面パノラマの原理図

が、外側に位置する撮像面は内側に位置する撮像面に比較して、光像が若干拡大されて撮影される。したがって、撮影された2枚の原画像を単純に平行移動させただけでは完全には重ならない。これを解決するために、画像を円筒面に投影する。円筒面投影により、画像は図2 (b) のように、たる型に変形され、撮影時の回転角に応じた量だけ画像を平行移動することで、2枚の画像は完全に重なる。このように処理することで、図1のようにつながりのよい全周パノラマ画像を得ることができる。

2.3 平面パノラマ

円筒パノラマでは、画像の撮影時に、カメラの投影中心の位置が移動しない。カメラの位置が移動する場合、被写体までの距離が十分に遠くない限り、円筒面投影を行っても、画像は完全には重ならない。重なり部分の被写体までの距離情報が全画素で既知であれば、完全に重なるように変形することは可能である。しかし、この条件は現実的ではない。

被写体が平面的であれば、異なる2箇所で撮影した画像に適切な平面射影変換(ホモグラフィ)を施すことで、重なり部分の画像を完全に重ねることができる。このように生成するパノラマを平面パノラマと呼ぶ。図3に平面パノラマの原理図を示す。一方の視点 C_1 で撮影した平面的な被写体と、他方の視点 C_2 で撮影した同じ被写体において、4点以上の画像座標の対応から、 3×3 の行列で記述される平面射影変換行列 H を求めることができる。この H によりお互いの座標値を変換することができるので、 C_2 だけで観察できる領域を、 H を用いて変換し、 C_1 で撮影した画像を継ぎ

目なく拡大することが可能になる。

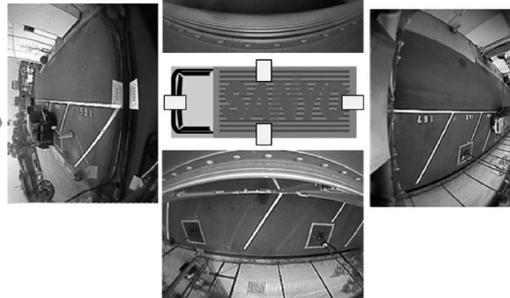
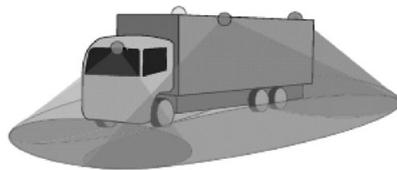
平面パノラマ合成においては、平面射影変換行列を高速かつ高精度に求めることが重要である。これは、2画像の重なり部の相関が最大になる平面射影変換行列を求める最適化問題になる。単純に重なり部の画素値の相関が最大になるように繰り返しの求める解法は、計算量が膨大であり、また局所解に陥る可能性がある。著者らはKLT法¹³⁾であらかじめ重なり部の特徴点对応を求めた後、線形解法で変換行列を求める手法を提案している¹⁵⁾。

平面パノラマは横長や縦長の建築物の撮影に有効である。また、手持ち撮影した画像であっても、遠景シーンは平面パノラマ、円筒パノラマの両方が可能である。また、図4 (a) のように車両上部に4台のカメラを取り付け、道路面を共通平面として平面パノラマ合成を行うことで、あたかも車の上空から撮影した航空写真のような画像を生成することができる。

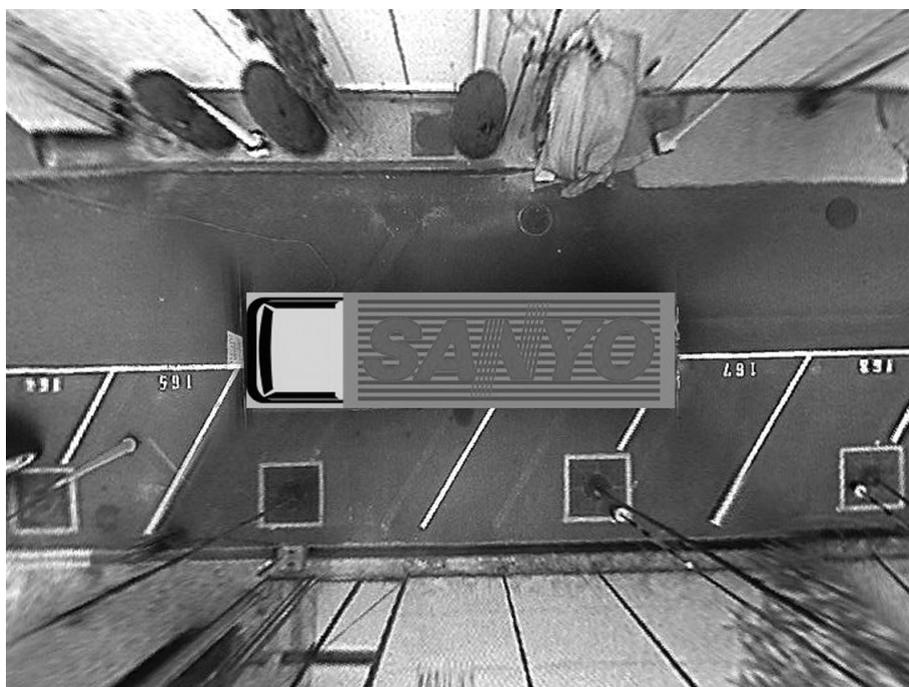
2.4 その他のパノラマ画像

衛星や航空機から撮影した画像をつなぎ合わせて、googleマップのように、非常にサイズの大きな航空写真が生成される。ズーム倍率が十分に大きく、被写体である地面と撮像面が並行であれば、画像を変形することなく、回転・平行移動だけで画像を重ね合わせることができる。この手法で莫大なサイズのデジタル画像が生成されている⁶⁾。

長尺絵巻物を分割撮影した複数の画像から、1枚のデジタル絵巻物が生成できる¹⁶⁾。基本的な考え方は絵巻物の上端・下端の境界線を手がかりに平面射影変換を用いてつなぎ合わせることである。しかし、絵巻物自身のたわみに



(a)4台のカメラの設置位置と入力画像



(b)道路平面での合成画像

図4 平面パノラマの車載カメラへの応用

よって部分的なずれが発生するため、局所的にオプティカルフロー計算を用いた位置合わせが必要である。

移動しながら撮影した動画の連続フレームから、細長い縦領域(ストリップ)を切出し、それらを接続することで横長のパノラマ画像を生成することができる。この手法をストリップパノラマと呼ぶ。一定速度で走る車窓から道路沿いの住宅を撮影し、街並みを横長のパノラマ画像として表現することに利用される。取得されるパノラマ画像は縦方向には透視投影、横方向には正投影になる。このような縦横で異なったカメラモデルで取得されるストリップパノラマ画像は、違和感のある写真になる。これを解決して、違和感の少ないストリップパノラマを生成する手法が提案されている¹⁾。

3. デジタルカメラを用いた3Dモデリング

3.1 3Dモデリングの概要

3DCGは、コンピュータゲームのキャラクタや工業デザインのCADなどに利用されており、インタラクティブ性と写実性に富む画像メディアである。3DCGデータを生成するプロセスを3Dモデリングと呼ぶが、通常、経験豊富な専門家が高機能なCGソフトやCADソフトを用いて、多大な時間をかけて行う作業である。

3DCGの写実性を向上させる一つの方法として、照明条件を制御しながら、実在する物体を全周方向から撮影し、取得された多数の画像から、自動的な処理により、3DCGデータを生成する方法がある。このイメージベースドモデ

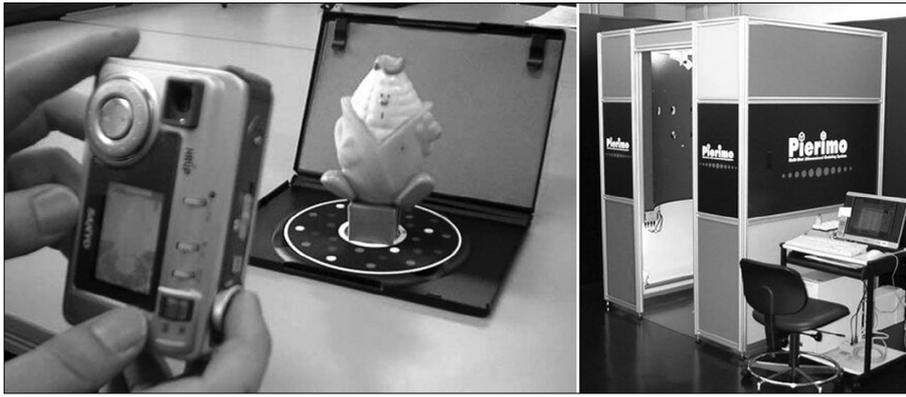


図5 デジタルカメラを用いた3Dモデリング装置

リング¹⁴⁾あるいは、イメージベースドレンダリングと呼ばれる手法を用いれば、デジタルカメラのような安価な入力装置を用い、照明条件の制御を簡単化することで、一般のユーザでも利用できるレベルに3Dモデリングを簡単化することが可能である。この分野の先駆的な研究として Virtualized Reality⁸⁾、ボクセルカラーリング¹¹⁾、スペースカービング⁷⁾などがある。現在でも活発に研究されている分野であるが、いくつかの実用システムも見られる。

デジタルカメラを用いた3Dモデリング装置に、図5に示すものがある^{19) 20)}。図5の左は、1台のデジタルカメラと安価なソフトだけで3Dモデリングを可能にしたものである。回転ステージと位置決めパタンを兼ねたCD-ROMの上に被写体を置き、被写体を一回転させながら10枚程度の画像を撮影する。視体積交差法により形状復元し、撮影画像をテクスチャマッピングすることで、3DCGデータを生成する。図5の右は、28台のデジタルカメラと2台のプロジェクタを用いて、頭髪を含む人間の頭部全体の3Dモデリングを可能にした装置である。視体積交差法とアクティブステレオ法のハイブリッド技術により高精度化と高速性を達成し、各種の頭部造型ビジネスやビジュアルシミュレーションに利用された。また、両機器の中間的な装置として、7台のデジタルカメラと1台のプロジェクタを用いて人間の顔面形状を入力する装置が、愛知万博で利用された²¹⁾。

3.2 拡張視体積交差法による3次元形状の復元¹⁸⁾

視体積交差法は、図6に示すように、物体全体を複数の位置から撮影し、画像中の物体の輪郭線(シルエット)を、カメラパラメータ(カメラ位置および焦点距離、画素サイズなど)を基に、3次元空間に逆投影することにより、複数の視体積(投影錐体)の交差部分として、物体の形状を取得する。ボクセル手法は最も基本的な計算アルゴリズムの一つである。ボクセルは物体が存在する3次元空間に設定される処理空間を分割した単位立方体のことである。ボクセル手法のアルゴリズムでは、各ボクセルが全視体積に含まれる場合に物体内、一つの視体積からでも外れれば物体外と判定するのが基本である。ここでは、ボクセル空間での投票に基づく、拡張視体積交差法について説明する。

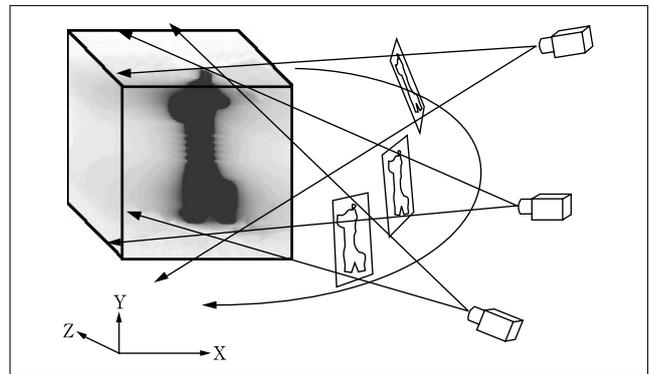


図6 ボクセル空間投票による視体積交差法

図6は、拡張視体積交差法の概念を示したものである。全体の処理は以下のように進む。

- (1) 適切な解像度のボクセル空間を設定する
- (2) 全ボクセルの値をゼロに初期化する
- (3) 各ボクセルの中心位置と、各カメラの投影中心を結ぶ直線が対応するシルエットの内部に含まれる場合、そのボクセルに1を加算する
- (4) 全ボクセルとシルエットの組合せについて(3)の処理を実行して、ボクセル投票を終了する
- (5) あらかじめ設定した閾値によりボクセル単位で2値化処理を行い、3次元形状を得る

シルエットに背景の写りこみなどによる欠損が存在することを考慮して、(5)の閾値は撮影枚数よりも小さい値を設定する。これによってロバスト性が高くなる。

ボクセルベースの手法は処理が単純で実装が容易であるが、最終的に得られる形状の精度がボクセルの分解能に依存するという課題があった。しかし、デジタルカメラの高画素化とパソコンの性能向上により、この課題は自然と解消されている。例えば、人体頭部を $1,024 \times 1,360 \times 1,024$ のボクセル空間でモデリングした場合、ボクセルの分解能は0.4mm程度になり、目鼻の特徴を再現するのに必要な精度での復元が可能になっている。

視体積交差法は照明条件の変化に頑健で安定に動作す

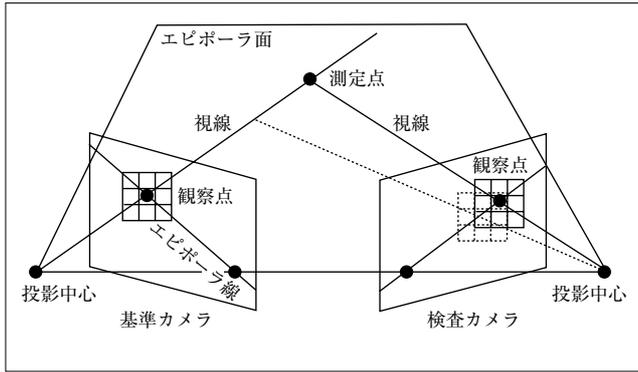


図7 ステレオ計測

る。しかし原理的に、シルエットに現れない凹凸を復元することができないため、この手法だけで最終形状を得ることは難しい。したがって、視体積交差法で得た3次元形状を初期形状として、多眼ステレオ法や照度差ステレオ法などのアルゴリズムによって高精度化を図る必要がある⁵⁾。

3.3 ステレオ視による3次元計測

ステレオ視は両眼視差を手がかりとして被写体までの距離を計測する技術である。図7に基本的な観察状態を示す。一方のカメラを基準カメラ、他方のカメラを検査カメラとし、基準カメラから被写体までの距離を計測するものとする。基準カメラの撮像面上の1点と投影中心を結んだ直線(視線)は、検査カメラの画像面においてエピソード線と呼ばれる直線を形成する。ステレオ視の問題は、基準カメラの各観察点に対して、検査カメラの対応点をエピソード線上から探索し、対応点の座標とカメラパラメータから測定点の3次元座標を求めることである。

ウィンドウベースの領域マッチングは、対応点探索の代表的なアルゴリズムである。図7に示すように、基準カメラにおいて観察点を中心とする小領域(ウィンドウ)を設定する。基準カメラの視線上で距離を計測する範囲と計測の分解能を設定する。設定された視線上の位置ごとに、検査カメラの対応する座標を求め基準カメラと同じ大きさのウィンドウを設定する。計測範囲内で相関値があらかじめ決めた閾値よりも大きく、かつ、最大になる位置を測定点の位置とする。相関計算には、ウィンドウ間の画素値の差の2乗和をとるSSD(Sum of Squared Difference)や、正規化相互相関係数を用いられる。検査カメラのウィンドウ領域算出における座標値は、サブピクセル精度で求める必要がある。これらの計算を高速化するには、基準カメラの画像面と検査カメラの画像面を共通平面に投影するレクティブイクエーションと呼ばれる幾何変化を用いる。ウィンドウの大きさは5×5~9×9程度が使われる。計測を高精度化するために、探索の分解能を段階的に細くすることや、離散的な距離間隔で得られる相関値から相関値が最大になる位置を補間するなどの手法がある。

ステレオ計測を本質的に高精度化するには、カメラの台

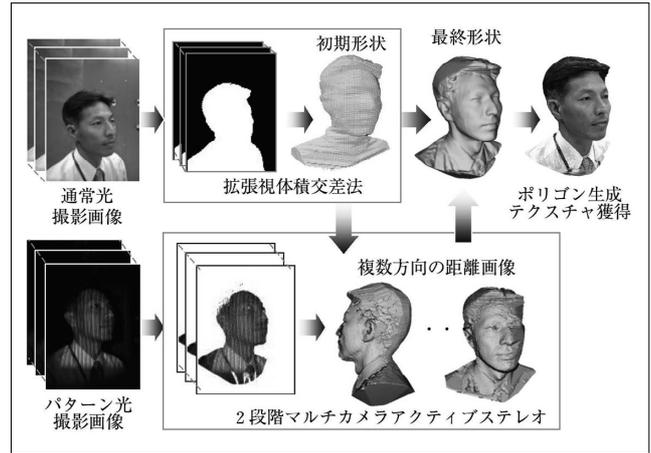


図8 ハイブリッドモデリング手法

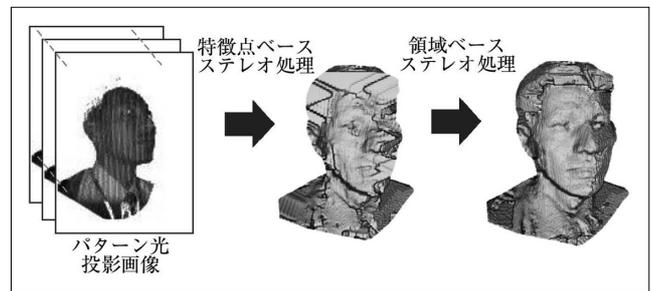


図9 2段階マルチカメラアクティブステレオ法

数を増やしてマルチベースライン手法⁹⁾をフルに活用すること、パタン光を投影するアクティブステレオ法を用いることが有効である。文献3)に、比較的単純なマルチベースラインアルゴリズムであっても、被写体を全周的に撮影した300枚以上の画像を用いることで、密で高精度な計測が可能であることが示されている。次節では、筆者らのグループが開発したアクティブステレオ計測技術の例を示す¹⁸⁾。

3.4 ハイブリッドモデリング手法¹⁸⁾

視体積交差法は頑健な計測が可能であるが、精度が不十分である。一方、ステレオ視はテクスチャが少ない部分の計測が困難であるため、物体の全体形状を取得することが難しい。そこで、両者を組合せることで、物体の全体形状を高精度に取得することが可能なハイブリッドモデリング手法が有効である。このハイブリッドモデリング手法は、図8に示すように、拡張視体積交差法を用いて人体頭部の初期形状を取得し、これを、二段階マルチカメラアクティブステレオ法によって高精度化し、最終形状とするものである。最後に、ポリゴン生成とテクスチャ獲得によって3DCGデータを完成する。

図9に、二段階マルチカメラアクティブステレオ法の概要を示す。7色程度のカラーストライプパターンを投影し画像を取得する。カラーストライプを領域分割し、ストライプの境界をエッジ抽出する。各エッジ点はその左右の色情報の特徴量として持つ特徴点とする。特徴点ごとに、基準

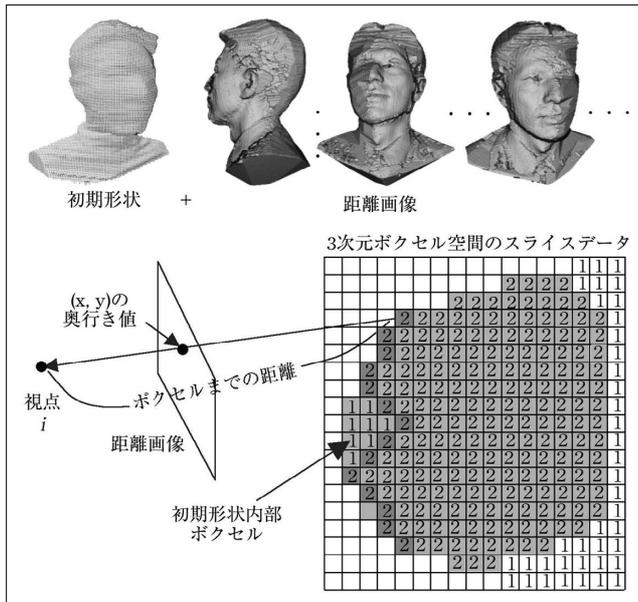


図10 ボクセル空間での初期形状と距離画像の統合

〔文 献〕

- 1) A. Agarwala, M. Agrawala, M. Cohen, D. Salesin and R. Szeliski: "Photographing Long Scenes with Multi-Viewpoint Panoramas", Proceedings of SIGGRAPH'06, pp.853-861 (2006)
- 2) S.E. Chen: "QuickTimeVR - An Image-based Approach to Virtual Environment Navigation", Proceedings of SIGGRAPH'95, pp.29-38 (1995)
- 3) M. Goesele, B. Curless and S. M. Seitz: "Multi-View Stereo Revisited", Proceedings of CVPR 2006, 2, pp.2402-2409 (2006)
- 4) R. Hartley and A. Zisserman: "Multiple View Geometry in Computer Vision", Cambridge University Press (2000)
- 5) C. Hernandez and F. Schmitt: "Silhouette and Stereo Fusion for 3D Object Modeling", Computer Vision and Image Understanding, 96, 3, pp.367-392 (2004)
- 6) J. Kopf, M. Uyttendaele, O. Deussen and M. F. Cohen: "Capturing and Viewing Gigapixel Images", ACM Trans. On Graphics, 26, 3, Article 93 (2007)
- 7) K.N. Kutulakos and S.M. Seitz: "A Theory of Shape by Space Carving", IJCV, 38, 3, pp.199-218 (2000)
- 8) P.J. Narayanan, P. Rander and T. Kanade: "Constructing Virtual Worlds Using Dense Stereo", Proceedings of ICCV, pp.3-10 (1998)
- 9) M. Okutomi and T. Kanade: "A Multiple-baseline Stereo", IEEE Trans. On PAMI, 15, 4, pp.353-363 (1993)
- 10) H.Y. Shum and R. Szeliski: "Construction of Panoramic Image Mosaics with Global and Local Alignment", IJCV, 36, 2, pp.101-130 (2000)
- 11) S.M. Seitz and C.R. Dyer: "Photorealistic Scene Reconstruction by Voxel Coloring", Proceedings of CVPR, pp.1067-1073 (1997)
- 12) R. Szeliski and H. Shum: "Creating Full View Panoramic Mosaics and Environment Maps", Proceedings of SIGGRAPH'97, pp.251-258 (1997)
- 13) C. Tomashi and T. Kanade: "Detection and Tracking of Point Features", Technical Report CMU-CS-91-132, Carnegie Mellon University, Pittsburgh, PA (1991), <http://vision.stanford.edu/~burch/klt>
- 14) 佐藤洋一: "イメージベースドモデリング", 人工知能学辞典, 8.13章, 共立出版 (2005)
- 15) 千葉, 蚊野, 美濃, 安田: "画像特徴に基づくイメージモザイクング", 信学論誌, J82-D-II, 10, pp.1581-1589 (1999)
- 16) 千葉, 蚊野: "デジタル絵巻のための画像モザイク手法", 情処学論誌, 44, SIG_9 (CVIM_7), pp.1-10 (2003)
- 17) 千葉, 畑中, 飯田: "画像特徴に基づく高速・高精度なパノラマ画像合成ソフトウェア", 三洋電機技報, 35, 1, pp.75-82 (2003)
- 18) 藤村, 大上, 寺内: "マルチカメラを用いた人体頭部3次元モデリングシステムとその応用", 情処学論誌, 45, SIG_8 (CVIM_9), pp.1-9 (2004)
- 19) 藤村, 大上, 寺内, 江見, 福崎: "デジタルカメラ画像から簡単に3次元データを生成するソフトウェア技術", 三洋電機技報, 35, 1, pp.68-74 (2003)
- 20) 藤原, 興水, 藤村, 藤田, 野口, 石川: "3D似顔絵フィギュア製作の実用化の試み", 情処学論誌, 43, SIG_4 (CVIM_4), pp.85-94 (2002)
- 21) 森島繁生: "フューチャーキャストシステム『三井・東芝館』", 映像学誌, 59, 4, pp.522-524 (2005)

カメラ・検査カメラのペアによってステレオ計測を行うとともに、第3のカメラを用いて計測結果を検証することで誤計測を除去する。この特徴点ベースステレオでは、十分な点数を十分な精度で計測することが難しいため、二段階目の処理として領域ベースステレオを導入する。領域ベースステレオは、基準カメラの全エッジ点について、第一段階で得られる計測結果を基に探索範囲を限定することで、高精度化を実現するとともに、第一段階で計測できなかったエッジ点についても再探索し、計測点数の増加を図る。

拡張視体積交差法で取得する初期形状と、二段階マルチカメラアクティブステレオ法で取得する距離画像を、図10に示すように、ボクセル空間での投票を用いて統合する。まず、初期形状として得られたボクセルの値を1に初期化する。次いで、各ボクセルについて、視点とボクセルの距離が対応する距離画像の奥行きよりも大きい場合に1を加算する。これをすべての視点とボクセルの組合せについて計算し、あらかじめ決めた閾値によって最終形状を抽出する。

4. む す び

パノラマ画像は一般ユーザのホームページでも見かけることが多くなり、デジタルカメラならではの楽しみ方として普及が進んでいる。市販のパノラマ画像合成ソフトウェア¹⁷⁾を手に入れることも容易である。一方、デジタルカメラを利用した3Dモデリングは、一般ユーザが利用するには完成度が不十分であり、未来的な技術にとどまっている。当面は、3DCGデザイナーが高度に写実的なCG映像を生成するためのツールとして利用されるであろう。

(2008年8月1日受付)



蚊野 浩 (ひろし) 1984年、京都大学大学院工学研究科情報工学専攻修了。同年、三洋電機(株)入社。現在、同社研究開発本部デジタル技術研究所担当部長。デジタルカメラ、車載カメラなどの画像処理技術を担当。DVDフォーラムPCC議長、工学博士。正会員。