

令和元年6月18日現在

機関番号：62615

研究種目：基盤研究(B) (一般)

研究期間：2016～2018

課題番号：16H02851

研究課題名(和文) モバイル環境におけるRGBDカメラを用いた3次元モデル生成と複合現実感への応用

研究課題名(英文) 3D modeling using an RGBD camera in the mobile environment and its application to mixed reality

研究代表者

杉本 晃宏 (Sugimoto, Akihiro)

国立情報学研究所・コンテンツ科学研究系・教授

研究者番号：30314256

交付決定額(研究期間全体)：(直接経費) 12,300,000円

研究成果の概要(和文)：モバイル環境にあるRGBDカメラから得られるストリーミングデータを入力として、リアルタイムに省メモリで3次元モデルを高精度に生成する技術の開発を目指し、パラメトリックな曲面による3次元モデル表現とモデルあてはめ、正確で頑健な奥行データ統合法を開発し、大規模な3次元モデルを実時間で生成することが可能となった。また、屋外環境における複合現実感への応用を示した。

研究成果の学術的意義や社会的意義

省メモリでリアルタイムに3次元モデルを高精度に生成する技術は、スマートフォンの写真撮影や映像撮影の要領で、エンドユーザがだれでも思い立った時にその場で、簡単にリアルタイムで3次元モデルを生成し、それを手軽に楽しめる環境を提供するものである。通信技術と連携することで、訪問が困難な博物館に保管された展示物や未公開遺産のモデル化・伝送を通じた文化遺産の3次元情報発信、遠隔講義における教材提供の充実、3次元地図データ生成などへの応用が見込まれる。

研究成果の概要(英文)：We aimed at developing a method for constructing high fidelity 3D models in real time with less memory consumption from streaming input data captured by an RGBD camera in the mobile environment. We developed a parametric surface representation for dense 3D modeling, a model fitting method, robust and accurate depth fusion, and so on. As a result, we succeeded in accurately modeling large-scale indoor scenes in real time with cheap memory consumption. We could also show the effectiveness of our developed methods for mixed reality in the outdoor scenes.

研究分野：数理工学

キーワード：3次元モデル 複合現実感 モバイル環境 コンピュータビジョン

1. 研究開始当初の背景

スマートフォンやタブレットの急速な普及は目覚ましい。スマートフォン、タブレットの利用率は2015年にはそれぞれ、62.3%、21.3%にまで達しており(総務省)、ビッグデータと融合したスマート革命に向けて躍進を続けている。スマートフォンやタブレットといった携帯型端末に搭載されたCPU、GPUの性能も飛躍的に向上しており、モバイル環境でのデータ処理技術の向上には目を見張るものがある。事実、携帯型端末に搭載されたカメラをつかって、ユーザはいつでもどこでも手軽に画像や映像を撮影・加工し、その場でそのままSNSにアップして共有するという状況が日常的になっている。一方、従来の奥行きセンサ(VividやCyraxなど)は、重厚、高価でスキャンに時間を要し、3次元をスキャンするのは専門家のみに限定されていたが、Microsoft KinectやAsus Xtion Proに代表される軽量で廉価な3Dカメラ(いわゆるRGBDカメラ)の出現により、誰でも手軽にRGBデータ付で3次元をスキャンすることができるようになった。しかも、従来とは異なり、リアルタイムのスキャンが可能になっている。今後、RGBDカメラの小型化・廉価化が推し進められ、きわめて近い将来、スマートフォンやタブレットにも(さらには、ウェアラブル端末にも)RGBDカメラが搭載され、現在の画像や映像と同様に、3D画像や3D映像がいつでもどこでも撮影・加工されるという状況が当たり前になることは確実である。2014年にGoogleは、奥行きセンサと動きセンサを装備したスマートフォンで3次元空間を認識することを目的としたプロジェクトTangoを、2015年にマイクロソフトは、スマートフォンの(通常の)カメラからの映像を入力とする3次元復元プロジェクト(MobileFusion)を発表していることからこのような方向性が注目されていることがわかる。このようにエンドユーザがモバイル環境で3次元データを手軽に入手し、データを加工できる環境が整いつつあり、3次元モデルを生成する技術は、以前にも増して、ますます重要となっている。

2. 研究の目的

携帯型端末に搭載された(即ち、モバイル環境にある)RGBDカメラを想定し、そのカメラから得られるストリーミングデータを入力として、リアルタイムに省メモリで3次元モデルを高精度に生成する技術の開発を目指す。そして、モバイル環境における複合現実感への応用を通してその有効性を実証する。具体的には、モデル表現を平面領域の集合から区分平面領域の集合、パラメトリック曲面の集合へと拡張することで省メモリ・処理速度を維持しつつ対象領域の複雑化に対応し、あてはめ面の最適化、データ統合の高精度化を通して生成モデルの精度向上を目指す。

3. 研究の方法

(1) 3次元モデル表現

省メモリ化を図るためには、3次元モデルをいかにコンパクトに表現するかが最も重要である。既に開発した平面領域による3次元モデル表現を発展させ、より表現能力の高い3次元モデル表現を開発する。具体的には、3次元点群を領域分割し、各領域を区分平面、あるいは、Bスプライン曲面やベジェ曲面といったパラメタ表現された曲面(パラメトリック曲面)で表現する手法を開発する。パラメトリック曲面を3次元モデル表現能力は高くなるが、制御点の選択、曲面の次数の選択、処理速度という問題が絡み合って存在する。ここでは、実際のデータを扱いながら、最大公約数的に実用的なパラメトリック曲面による3次元モデル表現を検討する。

(2) モデルあてはめ

領域に分割された3次元点群を区分平面やパラメトリック曲面にあてはめる場合、どのような基準で最適化するかは大きな問題である。点群の平面や曲面からのずれを最小化するようにあてはめればよいということではないからである。あてはめた平面や曲面を離散化し3次元データを格納する際に、複数の3次元点が同一のセルに保存されてしまうことを避けなければならない。したがって、平面や曲面の解像度が与えられたとき、異なる3次元点は異なるセルに対応するような平面や曲面のあてはめが好ましい。これを離散最適化問題として定式化し、効率的なアルゴリズムを考案する。

(3) ストリーミングデータの統合

RGBDカメラはビデオレート(30fps)で距離画像を取得できるため、同じ点を繰り返し計測する。計測回数が増えれば、その点の計測データ(位置情報)の信頼度が上がる。1回の計測では精度は低いが、繰り返し計測したデータの時間窓を用いた移動平均によってデータ統合を行い、位置情報の精度を保証するのが一般的である。こうした計算は、各点に対して独立に行われており、近傍の点の計測値との関係を取り入れるという観点が欠落しているのが現状である。3

次元モデル生成のための対象は大きさを持っているので、計測データを近傍の点と比べた場合、大きな変化はないはずである。ここでは、各点に対して、その近傍の点の計測データを考慮し、繰り返し計測からの精度保証の向上を検討する。

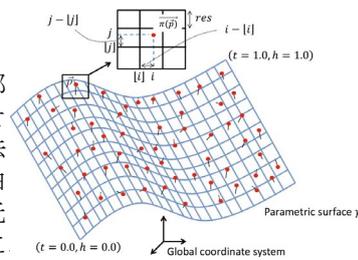
(4) モバイル環境下での複合現実感への応用

複合現実感では、実世界と3次元モデルの座標系を統一する位置合わせ、遮蔽処理、および、光源環境に即した見えの生成が重要となる。ここでは、モバイル環境下という前提で（すなわち、計算機能力、メモリが限られている状況下で）、これらの技術を開発する。また、現状のKinectでは屋外においてデータ取得が困難となっているため、屋外を想定したモバイル環境下でのRGBDデータ取得技術を開発する。

4. 研究成果

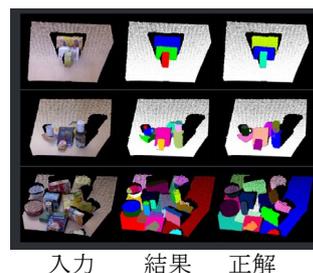
(1) パラメトリック曲面による3次元モデル表現

平面領域の集合による3次元モデル表現には、平面領域の境界部分で曲率が大きく変わるため、その表現能力には限界がある。そこで、パラメトリック曲面によって3次元モデルを表現する手法を検討した。曲面上の点を2次元のパラメタによって表現し、曲面上の各点の法線方向をそのパラメタで記述することで、3次元点をパラメトリック曲面上に投影することを可能とした。さらに、パラメトリック曲面をセルに離散化し、セル中のどの位置に投影点があるかの情報を利用して、元の3次元点の正確な情報をそのセルに格納する手法を考案した。これによって、原理的には、いかなるパラメトリック曲面であっても、それを用いた3次元モデル表現が可能で、その表現によって、もとの情報を失うことなくRGBDデータを格納することができることとなった。



(2) 幾何特性と色分布を併用したRGBDデータの領域分割

RGBDデータに対して個々の物体領域を分割する手法を開発した。計測された3次元点群が持つ幾何情報と色情報のバランスをとるために、奥行情報からの幾何特性と色の空間分布を利用し、分割を安定化させるとともに、スーパーボクセルの再帰分割を利用して高速化を図るといった特長を備えている。本手法によって、高速で頑健なRGBDデータの分割が可能となった。



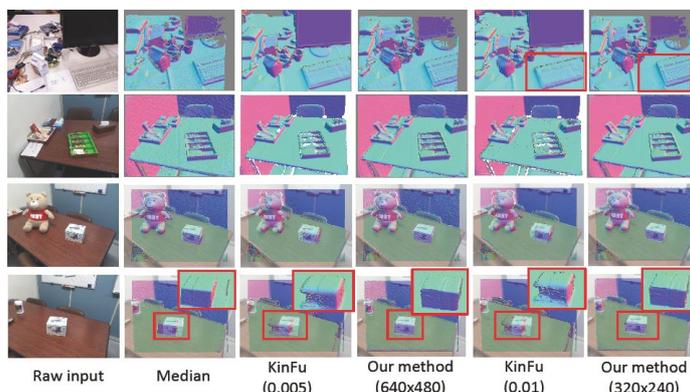
入力 結果 正解

(3) 離散最適化によるRGBDデータのモデルあてはめ

計測したRGBDデータからそのデータを説明する平面を推定する問題を混合整数計画問題として定式化し、離散化誤差を考慮した最適な平面当てはめの手法を開発した。ここでは、計測点にはアウト라이어が存在するという前提で、計測点がイン라이어か外れ値かを表す離散変数と当てはめる平面までの誤差を表す連続変数を導入して問題を定式化している。また、RGBDデータの個々の点に対して、離散化後にイン라이어とする多項式曲面を定めるパラメタが存在する領域（実行可能領域）の境界をデータ点と離散化に用いる構成要素の端点の座標を用いて記述し、イン라이어数の和を最大化するパラメタを求める問題を離散最適化問題として定式化した。そして、実行可能領域が凸となるために入力データが満たすべき条件を見出し、その条件が満たされる場合、局所最適解ながら解を効率的に求めるアルゴリズムを考案した。

(4) 変分メッセージパッシングを用いた奥行データ統合

これまでの3次元モデル生成法では、同一点を繰り返し計測して得られたストリーミングデータの奥行値に対して、時間窓を用いた移動平均をとることでデータ統合し、3次元モデルを逐次的に更新するというアプローチがとられていた。しかし、こうした計算は各点に対して独立に行われているため、その近傍において統合された奥行値との整合性が必ずしも保たれないという問題を抱えていた。また、各点でRGB値も得られているにも

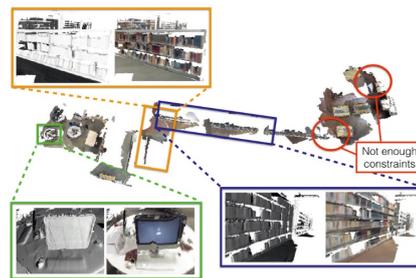


かかわらず、奥行値のみを用いて統合するというのが現状であった。そこで、指数族の共役分布を用いた離散、連続ベジアンネットワークを近似的に推論するための手法として知られている変分メッセージパッシングを用いた奥行データ統合法を開発した。計測した各点の奥行値

を確率分布で表現し、奥行データの統合を確率的に最適化する問題として定式化した。これにより、各点の近傍の点で計測された RGBD 値をその点の奥行データ統合に組み込むことができ、従来の統合手法と比較して、リアルタイム性を保持しつつ、正確で頑健なデータ統合が可能となった。

(5) 大規模な 3 次元モデル実時間生成

時系列 RGBD 画像をセグメントに分割してセグメントごとに局所モデルを生成し、統合によるひずみを最小化するようにその局所モデルを統合するという 2 ステップアプローチによる 3 次元モデル生成法を開発した。本手法は、省メモリ、リアルタイム処理、高精度の全てを実現する 3 次元モデル化手法となっている。本手法によって、60m x 5m x 2.5m 程度の図書館の 3 次元モデルを構築できることを確認した。



(6) 実空間への仮想物体の忠実な重畳

RGBD カメラと全方位カメラを併用し、RGBD カメラによる密な狭範囲 3 次元モデルと全方位カメラ映像による疎な広範囲 3 次元モデルを統合し、プローブなしで、実環境の照明条件と 3 次元形状を推定し、仮想物体をレンダリングする実時間システムを構築した。これによって、生成した 3 次元モデルの複合現実感への応用の有効性を示すことができた。また、本手法によって、鏡面反射する仮想物体の重畳が可能になった。

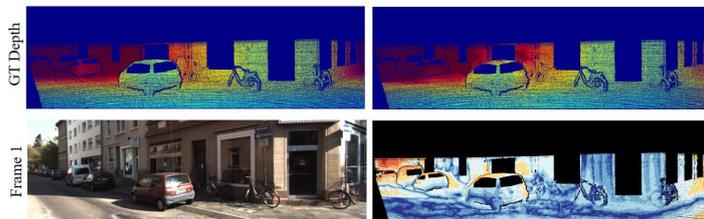


(7) 屋外環境における複合現実感のための RGBD 計測システム

屋外におけるモバイルプラットフォームによる RGBD 計測システムを構築した。屋外環境においては Kinect のような RGBD カメラを用いることは一般に難しい。そこで、LiDAR とカメラを搭載したセンサシステムを開発した。複数センサによって計測システムを構成する際には、センサ間の校正が重要な課題となる。そこで二つのセンサを同時に移動させ、移動量とカメラ画像 - 奥行画像フュージョンによってセンサ間の相対位置姿勢を高精度に推定する手法を開発した。

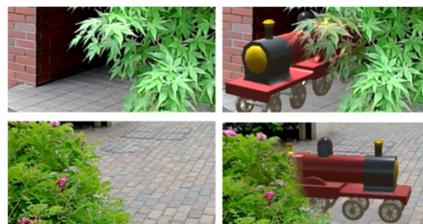
(8) 屋外環境における複合現実感のための奥行推定

カメラ画像のみから RGBD データを取得するために、2 枚の連続画像を入力とするオプティカルフロー、奥行画像の実時間同時推定手法を開発した。モーションステレオでは、カメラの位置姿勢に誤差が含まれる場合は正しい奥行が得られない。そこで、画像空間ではなく 3 次元空間で形状フィルタを適用することによって位置姿勢推定誤差に頑健な奥行画像推定手法を開発した。



(9) 複合現実感における RGBD 画像を用いた遮蔽処理

通常の RGBD センサの奥行はカラー画像に比べて疎であり精度も不十分である。そこで、CNN を利用してカラー画像の領域の意味的ラベル付けを行い、奥行値から得られた前景・背景確率分布と組み合わせることで適切に実世界と仮想世界の重畳する手法を開発した。この描画の際には、境界及び枝葉などの半透明な領域では視認性ブレンディングを用いることで自然な遮蔽処理を実現した。



5. 主な発表論文等

[雑誌論文] (計 18 件)

- (1) T. -N. Le and A. Sugimoto: Video Salient Object Detection Using Spatiotemporal Deep Features, IEEE Trans. on Image Processing, Vol. 27, No. 10, pp. 5002-5015, 2018. DOI: 10.1109/TIP.2018.2849860
- (2) F. Sekiya and A. Sugimoto: On Properties of Analytical Approximation of Morphological

Discretization for 2D curve and 3D surface, *Mathematical Morphology: Theory and Applications*, Vol. 2, 25-34, 2017. DOI: 10.1007/s10472-014-9425-7

- (3) D. Thomas and A. Sugimoto: Modeling Large-scale Indoor Scenes with Rigid Fragments using RGB-D Cameras, *Computer Vision and Image Understanding*, Vol. 157, pp. 103-116, 2017. DOI: 10.1016/j.cviu.2016.11.008
- (4) D. Thomas and A. Sugimoto: Parametric Surface Representation with Bump Image for Dense 3D Modeling Using an RGB-D Camera, *International Journal of Computer Vision*, Vol. 123, pp. 206-225, 2017. DOI: 10.1007/s11263-016-0969-3

[学会発表] (計 25 件)

- (1) T.-N. Le and A. Sugimoto: Semantic Instance Meets Salient Object: Study on Video Semantic Salient Instance Segmentation, *Proc. of IEEE Winter Conference on Applications of Computer Vision (WACV2019)*, pp. 1779-1788, 2019. DOI: 10.1109/WACV.2019.00194
- (2) Z. Kukelova, C. Albl, A. Sugimoto, and T. Pajdla: Linear Solution to the Minimal Absolute Pose Rolling Shutter Problem, *Proc. of Asian Conference on Computer Vision (ACCV2018)*, LNCS 11363, pp. 265-280, 2018. DOI: 10.1007/978-3-030-20893-6_17
- (3) Y. Shih-Hsuan, D. Thomas, A. Sugimoto, S.-H. Lai, and R. Taniguchi: SegmentedFusion: 3D Human Body Reconstruction using Stitched Bounding Boxes, *Proc. of International Conference on 3D Vision (3DV2018)*, pp. 190-198, 2018. DOI: 10.1109/3DV.2018.00031
- (4) M. Roxas, T. Hori, T. Fukiage, Y. Okamoto, and T. Oishi: Occlusion Handling using Semantic Segmentation and Visibility-based Rendering for Mixed Reality, *Proc. of 24th ACM Symposium on Virtual Reality Software and Technology*, 8pages, 2018. DOI: 10.1145/3281505.3281546
- (5) R. Ishikawa, T. Oishi, and K. Ikeuchi: LiDAR and Camera Calibration using Motion Estimated by Sensor Fusion Odometry, *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2018)*, pp. 7342-7349, 2018. DOI: 10.1109/IROS.2018.8593360
- (6) T. Oishi: 3D Vision for Mobility, 2nd International MIS Workshop on 3D Vision and Robotics, 2018. (招待講演)
- (7) M. Roxas and T. Oishi: Real-Time Simultaneous 3D Reconstruction and Optical Flow Estimation, *Proc. of IEEE Winter Conference on Applications of Computer Vision (WACV2018)*, pp. 885-893, 2018.
- (8) D. R. Walton, D. Thomas, A. Steed, and A. Sugimoto: Synthesis of Environment Maps for Mixed Reality, *Proc. of IEEE International Symposium on Mixed and Augmented Reality (ISMAR2017)*, pp. 72-81, 2017. DOI: 10.1109/ISMAR.2017.24
- (9) F. Sekiya and A. Sugimoto: Efficiently Updating Feasible Regions for Fitting Discrete Polynomial Curve, *Proc. of 20th International Conference on Discrete Geometry for Computer Imagery (DGCI2017)*, LNCS10502, pp. 254-266, 2017. DOI: 10.1007/978-3-319-66272-5_21
- (10) T.-N. Le, and A. Sugimoto: Deeply Supervised 3D Recurrent FCN for Salient Object Detection in Videos, *Proc. of British Machine Vision Conference (BMVC2017)*, 13 pages, 2017.
- (11) F. Verdoja, D. Thomas, and A. Sugimoto: Fast 3D Point Cloud Segmentation using Supervoxels with Geometry and Color for 3D Scene Understanding, *Proc. of IEEE International Conference on Multimedia and Expo (ICME2017)*, pp. 1285-1290, 2017.
- (12) K. Fukano, Y. Mochizuki, S. Iizuka, E. Simo-Serra, A. Sugimoto, and H. Ishikawa: Room Reconstruction from a Single Spherical Image by Higher-order Energy Minimization, *Proc. of 23rd International Conference on Pattern Recognition (ICPR2016)*, pp. 1768-1773, 2016.
- (13) C. Morales, S. Ono, Y. Okamoto, M. Roxas, T. Oishi, and K. Ikeuchi: Outdoor Omnidirectional Video Completion via Depth Estimation by Motion Analysis, *Proc. of International Conference on Pattern Recognition (ICPR2016)*, pp. 3934-3939, 2016.
- (14) R. Ishikawa, M. Roxas, Y. Sato, T. Oishi, T. Masuda, and K. Ikeuchi: A 3D Reconstruction with High Density and Accuracy using Laser Profiler and Camera Fusion System on a Rover, *Proc. of International Conference on 3D Vision (3DV2016)*, pp. 620-628, 2016. DOI: 10.1109/3dv.2016.70
- (15) C. Albl, A. Sugimoto, and T. Pajdla: Degeneracies in Rolling Shutter SfM, *Proc. of European Conference on Computer Vision (ECCV2016)*, LNCS 9909, pp. 35-51, 2016. DOI: 10.1007/978-3-319-46454-1_3

6. 研究組織

(1) 研究分担者

研究分担者氏名：大石 岳史

ローマ字氏名：(OISHI, takeshi)

所属研究機関名：東京大学

部局名：生産技術研究所

職名：准教授

研究者番号（8桁）：80569509

(2) 研究協力者

研究協力者氏名：パジエドラ トマス

ローマ字氏名：(PAJDLA, tomas)

研究所力者氏名：トマ ディエゴ

ローマ字氏名：(THOMAS, diego)

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。