

令和 2 年 7 月 2 日現在

機関番号：62615

研究種目：研究活動スタート支援

研究期間：2017～2018

課題番号：17H07324

研究課題名（和文）自由視点3Dパノラマ映像生成に基づくVRコンテンツ生成

研究課題名（英文）A Study on the Free-viewpoint Panoramic Video

研究代表者

池畑 諭（Ikehata, Satoshi）

国立情報学研究所・コンテンツ科学研究系・助教

研究者番号：70804061

交付決定額（研究期間全体）：（直接経費） 2,300,000円

研究成果の概要（和文）：VRコンテンツ拡充のための技術についての検討を行った。特に、没入感のある体験を実現するためには撮影されたシーンの正確な3次元形状を復元できる事が望ましい。その実現のために、我々は畳込み深層ネットワークに基づく複数陰影画像からの3次元形状復元手法（フォトメトリックステレオ法）を世界で初めて提案した。また、必要なデータセットを作り上げ、これを公開した。以上の成果に加えて、補助的な課題として3次元復元に役立つ深層学習に基づく外れ値除去の手法や、効率的な離散最適化アルゴリズム、撮影位置同定のためのジオタグ画像の利用方法等を提案した。

研究成果の学術的意義や社会的意義

研究成果において副次的に生成されたソースコードやデータを公開したことの学術的な意義は大きい。データベースは公開後、世界中の研究者によって利用されるに至っている。また、我々は1980年代から続くフォトメトリックステレオ法の研究に対して、新しい概念の提示による技術的進歩、データセットの共有と両方の観点から、大きな影響を与えた。現在でも我々の提案した観測地図と呼ばれる表彰に基づく手法を発展させた研究が発表され続けている。

研究成果の概要（英文）：In this project, we developed the techniques to enrich the immersive virtual reality contents. Especially, we aimed to recover the detailed 3-d geometry from images to generate the virtually synthetic scenes more physically plausible. We presented the world's first convolutional-neural-network-based photometric stereo algorithm that can accurately recover the detailed scene geometry. We further published the first large-scale training dataset of the photometric stereo problem. In addition to those achievements, we supplementarily presented the robust deep-learning algorithm to remove outliers in the observation, efficient discrete optimization algorithms and geo-localization algorithm using geo-tag information.

研究分野：コンピュータビジョン

キーワード：フォトメトリックステレオ 3次元センシング 最適化アルゴリズム

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

筋萎縮性側索硬化症や脊髄損傷由来の下半身不随のような重度な身体障害を伴う難病患者にとって、VR (バーチャルリアリティ) 技術が、仮想的に旅行に出掛ける等の現実の肉体的制約を超越した体験を可能にするのではないかと期待されている。ところが、現在公開されている写実的な VR コンテンツの殆どは Google Jump や Facebook Surround 360 に代表されるように、パノラマカメラで撮影された映像をそのまま HMD (ヘッドマウントディスプレイ) に表示したものであり、視点を变化させる事はできても、VR 空間内を自由に移動する事はできない。すなわち、「空間内を自由に歩き回れる写実的な VR コンテンツを自動で生成する手法」が必要とされていた。

2. 研究の目的

(1) 本研究では、上記の背景の通り、写実的な VR コンテンツ生成を目的とした。申請時においては、離散的に撮影された画像から実際には存在しない任意視点の画像を合成する IBR (Image-based Rendering) 法をパノラマ画像に適用して任意視点画像と近い視点のパノラマ画像を合成し、それを HMD に表示することを有用なアプローチとして提案した。パノラマカメラを利用する事により、撮影のコストやデータ量を削減し、また合成後の画像を直接 HMD に提示できる形式であるパノラマ画像にすることで高速な頭部運動にも対応する事が期待された。

(2) また、パノラマカメラ画像は通常画像とカメラ幾何も画像特徴も異なり既存手法が直接適用できないため、パノラマ画像間でのロバストな対応を得て効果的に任意視点のステレオパノラマ画像を合成するにはどのような手法が効果的かを明らかにしたり、撮影の手間と生成される画像の品質の双方の観点から、最適なカメラ配置やシーンの規模と必要画像枚数の関係を明らかにすることを研究的な挑戦として設定した。

3. 研究の方法

(1) パノラマ画像に基づく IBR 法の問題点として、写実的な画像は合成できるものの、合成された画像は常に撮影時の光源環境下で固定されてしまうという問題が見つかった。より VR 体験を充実させるためには、単純に画像を補完するのではなく、光の反射を考慮したり、光源環境を操作できるような汎用性の高い情報を復元する必要がある事が明らかになった。そのため、パノラマ画像の研究に取り組み前に、複数の異なる光源環境下で撮影された画像群から被写体の法線や材質を復元するフォトメトリックステレオ法の研究に注力する事とした。特に深層学習をフォトメトリックステレオ法に適用するための方法を検討した。

(2) 町の中を歩き回るような VR コンテンツを生成する場合に、撮影された画像と地図との対応を見つけることは非常に重要である。一般的には GPS による位置測位が用いられるが、一方で密集した住宅地等では GPS による自己位置同定が困難である場合がある。しかし画像の特徴ベースの手法のみでは大域的な位置を抽出する事はやはり困難である。そこで、カメラによって撮影された GPS 情報の付与されていない画像と Google Street View のように位置情報が付与された画像の両方を組み合わせて地図上の撮影位置を正確に計算する技術を検討した。

(3) フォトメトリックステレオ法に限らず、VR コンテンツ生成のためには様々な最適化問題を解く必要がある。このような最適化技術は基礎的な問題ではあるが、あらゆる研究の基本要素であり、決して軽視してはいけない。例えば、IBR 法において重要な距離画像を生成するためには大規模な MRF 等の離散最適化問題を効率良く解く必要がある。本研究の予備的な課題として離散最適化問題を解くためのアプローチを検討した。

4. 研究成果

本年度は、主にディープラーニングを利用した 3 次元復元手法の検討を行った。特に異なる光源下で撮影された陰影画像から被写体の表面法線を復元するというフォトメトリックステレオ法をターゲットとして研究を行った。

(1) CNN を利用した End-to-End フォトメトリックステレオ法の提案 (個人研究)

フォトメトリックステレオ法は、画像が生成されるプロセスを物理的なモデルに基づいて定式化し、その逆問題を解くことによって画像から被写体の法線を求めるという手法である。一般的にモデルが複雑になるほどに最適化が困難になるため、鏡面反射や二次反射、影等の反射現象をすべてモデル化する事は困難であった。そのため、モデル化の部分を機械学習で代替する事ができればフォトメトリックステレオ法が様々な材質や形状に対応できるようになると期待され

る。しかし、フォトメトリックステレオ法は入力画像の枚数が不定なので、例えばディープラーニングのように学習段階とテスト段階において入力の形状が固定された機械学習のアプローチでは一般的に解くことはできない。そこで我々は、不特定枚数の画像をいったん固定長の情報に変換してそれに対してCNN (Convolutional Neural Network) を適用することによって、フォトメトリックステレオ法をディープラーニングによって解く手法を実現した。実験の結果、精度と速度の両方において(ただし学習時間は除く)、既存手法を上回る結果を確認した。また副次的な産物として、これまでは存在しなかったフォトメトリックステレオ法における大規模なベンチマークデータセットを作成した。本研究の成果はコンピュータビジョンのトップカンファレンスである ECCV2018 に採択された。

(2) Recurrent Neural Network を利用した外れ値除去

フォトメトリックステレオ法では、ランバート拡散反射モデルを仮定して、そのモデルで表現できない反射や影等を外れ値として除外するというアプローチが存在する。我々は以前この分野の研究で、Sparse Bayesian Learning 法 (SBL 法) と呼ばれる外れ値がベイズ分布に従うという仮定に基づいた L0 最適化手法を利用して高い精度を実現したが、我々は新たにディープラーニングを利用して外れ値を除去する事を試みた。具体的には、LSTM (Long short-term memory) を利用して学習した結果から、反復最適化の解の更新を最適化するという手法を提案して実際に SBL 法と比べて高い精度で法線が復元可能であることを示した。この成果は機械学習のトップカンファレンスである NIPS2017 に採択された。

(3) Geo タグを利用した自己位置推定の高精度化

地図上における写真の撮影位置の同定は例えば自動運転においては非常に基本的なタスクである。一般的に多視点画像からの 3 次元復元手法は同時にカメラの位置姿勢を推定するが、これが例えば数十キロにわたるようなデータの場合、画像間の連続性が保存されていなかったり、蓄積する推定誤差によってスタートとゴールの地点がずれてしまう Scale Drift と呼ばれる現象が発生する。本研究ではこの問題に対処するため、写真の GeoTag 情報を利用する手法を提案した。GeoTag は、写真に GPS 等によって埋め込まれた位置情報の総称で、カメラ自体が保持しているのみならず、例えば Google Street View のようなサービスで利用されている画像群にも付与されている。我々は、特に撮影したカメラ自体の GPS が機能しない場合を想定して(例えば建物の密集地等)、外部で利用可能な画像から適切な GeoTag が付与された画像を抽出し、その結果を利用して位置情報を最適化するという手法を提案した。この成果は ECCV2018 における自動運転に関わる Workshop である CVRSUAD に採択された。

(4) 階層的コストボリュームフィルタリングに基づく最適化手法の提案

コストボリュームフィルタリングは、MRF (Markov-Random-Field) 等の離散最適化問題を効率良く解くためのアルゴリズムとして提案された。この手法は、最適化の尤度 (Unary term) と平滑化項 (Pairwise term) を同時に最適化するのではなく、尤度に基づいて生成されたコストボリュームを平滑化する事によってエネルギーの近似値を得ようとするアプローチである。しかしこの手法は、対象となる離散最適化問題のラベル数によっては計算量が膨大になるという欠点が挙げられた。そこで我々は、ラベル数が大きくならないように候補となるラベルから可能性の高いラベルのみを抽出して、その周辺で最適解を探すというアプローチを提案した。具体的には、あらかじめ画像の解像度を下げて(それにより計算量が下がる)その状態ですべてのラベルを利用して最適化した結果から、あり得るラベルのサブセットを作成し、解像度をそれよりも段階的に上げた問題に対してサブセットのみを利用して最適化するという過程を、解像度を徐々に上げていながら行った。この研究成果は Multimedia Tools and Applications に採択された。

5. 主な発表論文等

〔雑誌論文〕 計1件(うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件)

1. 著者名 Ryosuke Furuta, Satoshi Ikehata, Toshihiko Yamasaki, Kiyoharu Aizawa	4. 巻 -
2. 論文標題 Efficiency-enhanced cost-volume filtering featuring coarse-to-fine strategy	5. 発行年 2017年
3. 雑誌名 Multimedia Tools and Applications	6. 最初と最後の頁 1-23
掲載論文のDOI (デジタルオブジェクト識別子) https://doi.org/10.1007/s11042-017-4897-1	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

〔学会発表〕 計3件(うち招待講演 0件/うち国際学会 3件)

1. 発表者名 Hao He, Bo Xin, Satoshi Ikehata, David P. Wipf
2. 発表標題 From Bayesian Sparsity to Gated Recurrent Nets
3. 学会等名 Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017 (国際学会)
4. 発表年 2017年

1. 発表者名 Satoshi Ikehata
2. 発表標題 CNN-PS: CNN-based Photometric Stereo for General Non-Convex Surfaces
3. 学会等名 European Conference on Computer Vision (ECCV) (国際学会)
4. 発表年 2018年

1. 発表者名 Kazuya Iwami, Satoshi Ikehata, Kiyoharu Aizawa
2. 発表標題 Scale Drift Correction of Camera Geo-Localization using Geo-Tagged Images
3. 学会等名 6th Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving (ECCV2018 Workshop) (国際学会)
4. 発表年 2018年

〔図書〕 計1件

1. 著者名 米谷 竜、斎藤 英雄、池畑 諭、牛久 祥孝、内山 英昭、内海 ゆづ子、小野 峻佑、片岡 裕雄、金崎 朝子、川西 康友、齋藤 真樹、櫻田 健、高橋 康輔、松井 勇佑	4. 発行年 2018年
2. 出版社 共立出版	5. 総ページ数 272
3. 書名 コンピュータビジョン	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----