

## 物体追跡画像群からの軌跡付き 3次元モデルの生成法 A generation method for 3D models with trajectories from object tracking images

佐藤 大輔\* 岩切 宗利† 藤原 匠†  
Daisuke Sato Munetoshi Iwakiri Takumi Fujiwara

### 1 まえがき

近年、安全意識の向上のため、車両にドライブレコーダーを搭載する人が増えている。ドライブレコーダーで撮影された映像は交通事故等の際に、当時の状況を把握することに使われている。ドライブレコーダーを搭載している車両同士の交通事故であれば、両者の事故直前の見え方を確認することができる。しかし、ドライブレコーダーの視点に限定される。つまり、非搭載車両や歩行者側の視点を確認することはできない。事故の状況を様々な視点から確認できて、車両、歩行者等の移動物体の推移を把握することができれば、より詳細な検証の実施が期待できる。専門的知識のない人でも直感的、視覚的に理解しやすい媒体として 3次元モデルがある。しかし、単なる 3次元モデルは移動物体や、その軌跡情報を含んでいない。ここで、背景等静止物体の 3次元モデル、移動物体の 3次元モデル及びその軌跡を含んだものを統合 3次元モデルと呼ぶこととする。統合 3次元モデルがあれば、誰でも理解しやすい認識共有ができ、時間推移による変化を把握することができる媒体となる。そこで本研究では、動画から取得した画像群を用いて、統合 3次元モデルを生成することを目標とする。

### 2 関連研究

3次元モデル生成のための 3次元点群データの取得手法は、能動的手法と受動的手法の大きく 2つに分けられる。能動的手法の代表としては、LiDAR (Light Detection And Ranging) 及び ToF (Time of Flight) があり、深度センサーにより 3次元点群データを取得するというものである。受動的手法として、目標物の画像群から 3次元モデルを生成する SfM-MVS (Structure from Motion / Multi view Stereo) がある。SfM[1] は、画像間の目標物が持つ特徴点を抽出し、その変化から撮影された位置を推定し、点群を構築する。SfMに加えて、MVS[2] の処理を行うことで密な点群を生成することができる。上記の能動的手法及び受動的手法の異なる点の 1つは、取得コストである。能動的手法は、点群データの取得に高価な機材を必要とする場合が多いが、受動的な SfM-MVS は比較的安価かつ身近なカメラにより取得された画像群から 3次元点群データを生成することができる。その多くは、静止物体を 3次元モデル化することを目的に行っている。また、SfM-MVS 処理を実施する際に用いる画像群は 3次元モデル生成精度に大きな影響を及ぼす。画像内における物体の分類、深度の推定、移動物体の除去 [3] 等の処置により、汎用性並びに精度の向上を図る研究もされている。このように移動物体を除去する研究も

\* 防衛大学校理工学研究科情報数理, Mathematics and Computer Science, Graduate School of Science and Engineering, National Defense Academy of Japan

† 防衛大学校情報工学科, Department of Computer Science, National Defense Academy of Japan

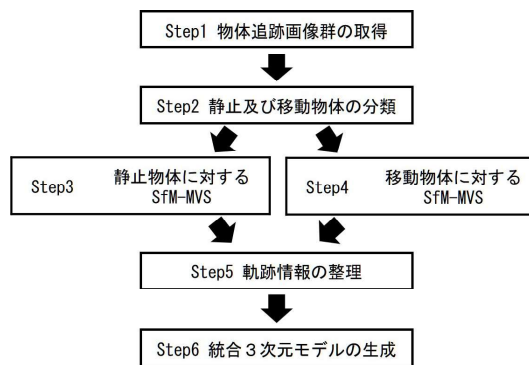


図 1 提案手法の流れ

ある一方で、移動物体とその軌跡を把握することは、時間経過により位置を変えするという特性から先の動きを推定できる等の価値がある。そのため、最近では移動物体に着目した研究も出始めている。Tao[4]らは、SfM 処理を移動物体の追跡に用いて自動運転技術の発展に寄与している。しかし、統合 3次元モデル生成に関する研究はなされていない。そこで本報告では、移動物体の 3次元モデル生成について示す。

### 3 提案手法

図 1 は、本研究の目的達成のための流れである。図 2 は、図 1 における従来手法と提案手法を比較したイラストである。最初に、映像等から Step1 を実施する。Step1 により得られた画像を Step2 及び Step3 により、静止物体の 3次元モデルを生成及びカメラの軌跡を取得する。次に Step2 及び Step4 により、移動物体の 3次元モデルと移動物体の軌跡を取得する。そして Step5 では、Step3 及び Step4 で得た情報を整理する。最後に、各 3次元モデル及び軌跡を統合することにより、3次元空間で表現することができる。ここで、図 1 における Step1、Step2 及び Step3 に関する研究は多く存在する。しかし、Step4 については移動物体の追跡に関する研究 [4] はされているものの 3次元モデル生成は検討されていない。そのため本研究では、Step4 の実施要領を検討した。通常の SfM-MVS 処理では静止物体を仮定しているため、移動物体の画像群を用いた場合でも、カメラ位置が動いた形でモデルが出力される。そこで提案手法では、カメラ位置と物体のベクトルを逆計算し、カメラ位置を固定した場合の物体の軌跡を取得する。

### 4 実験

統合 3次元モデル生成の一助とするため、移動物体の SfM-MVS によるカメラ固定時における軌跡付き 3次元モデル生成を確認した。映像の取得、フレーム抽出、移動物体の抽出、移動物体の 3次元モデル生成の順に実施し、目視により評価した。

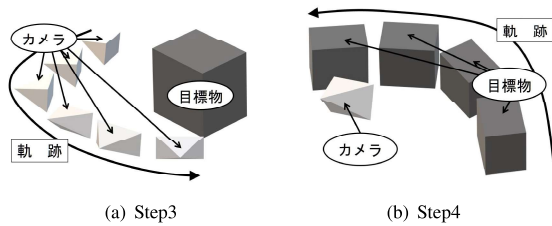


図 2 Step3 と Step4 の比較

表 1 実験用の機材及びソフトウェア

項目	機材及びソフトウェア名	諸元等
カメラ	GoProHERO4	1080p,30fps,45Mbps, 1920 × 1080
移動物体	logicool Z533 梱包箱	37cm × 27cm × 22cm
フレーム抽出ソフト	FFMpeg	ver4.2.7
画像編集ソフト	MicroSoft ペイント 3 D	ver6.2209.1037.0
SfM-MVS ソフト	Pix4DMapper	verPro3.2.23
3次元データ処理ソフト	Blender MeshLab	ver2.9.2 ver2022.2

表 2 Pix4DMapper 実行時の主な設定

パラメータ	設定値
Image Scale	Multiscale,860 × 540
Point Density	Optimal
Minimum Number of Matches	3

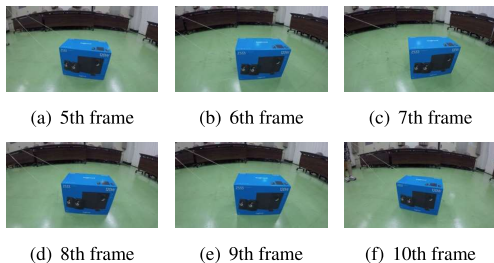


図 3 画像編集ソフトで抽出した1フレームの一例

#### 4.1 実験手法

まず、カメラを用いて移動物体を約 2 m 移動させる様子を手持ちカメラで追うようにして動画を撮影した。次に、物体追跡画像群の取得のため、撮影した映像からフレーム抽出ソフト (FFMpeg) により解像度の高いものを利用するため I フレーム及び P フレームの画像を抽出した。本実験では移動物体の 3 次元モデル生成を重視し、高い精度で背景を除去するため、画像編集ソフト (ペイント 3 D) で目視により移動物体の抽出を行った。最後に、SfM-MVS ソフト (Pix4DMapper) により、移動物体の 3 次元モデル生成し、カメラを固定した際の 3 次元モデルの位置を 3 次元データ処理ソフト (Blender 及び MeshLab) で表した。



図 4 データセットの一例

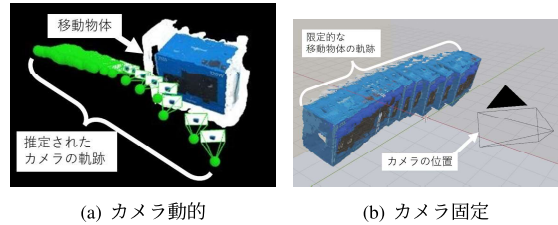


図 5 生成した 3 次元モデル

表 1 は使用した機材及びソフトウェアであり、表 2 は SfM-MVS ソフトのパラメータ設定である。

#### 4.2 実験結果と考察

図 3 は、I フレーム 13 枚、P フレーム 63 枚の計 76 枚を抽出したフレーム抽出結果である。図 4 は、移動物体の抽出結果である。図 5(a) は、生成された 3 次元モデルと各フレームのカメラ位置である。図 5(b) は、カメラの軌跡から逆算した位置に、3 次元モデルを配置したものである。SfM-MVS 処理により移動物体の周囲に生成された点群は、アーティファクトとして除去した。取り除いたアーティファクトは、移動物体と白地背景の境界に生じた特徴点によるものと考えられる。本実験の結果、これまでの SfM-MVS 処理でノイズ成分として取り除かれていた移動物体を 3 次元化し、その運動を復元できることが分かった。

#### 5 むすび

本研究では、物体追跡画像群から軌跡付き 3 次元モデルを生成することを目標とし、従来、研究されてきていない移動物体の 3 次元モデル生成について実施した。その結果、物体追跡画像群から移動物体の 3 次元モデルを生成することができた。今後は、提案手法の各処理を応用した全体の統合 3 次元モデル生成を目指し、研究を継続していきたい。

#### 参考文献

- [1] Johannes L. Schonberger, Jan-Michael Frahm, "Structure-from-Motion Revisited", IEEE Conference on Computer Vision and Pattern Recognition, DOI:10.1109/CVPR.2016.445(2016)
- [2] Pierre Moulon, Pascal Monasse, Renaud Marlet, "Global Fusion of Relative Motions for Robust, Accurate and Scalable Structure from Motion", IEEE International Conference on Computer Vision, DOI:10.1109/ICCV.2013.403(2013)
- [3] Sudheendra Vijayanarasimhan, Susanna Ricco, Cordelia Schmid, Rahul Sukthankar, Katerina Fragkiadaki, "SfM-Net: Learning of Structure", Computer Vision and Pattern Recognition, arXiv:1704.07804(cs)(2017).
- [4] Tao Liu, Yong Liu, "Moving Camera-Based Object Tracking Using Adaptive Ground Plane Estimation and Constrained Multiple Kernels", Journal of Advanced Transportation, Article ID 8153474, 15 pages (2021)