

画像群を用いた移動カメラの位置姿勢推定による 任意視点画像の生成

大江 拓也[†] 山下 淳[‡] 金子 透[†] 石原 進[†] 浅間 一[†]

[†] 静岡大学工学部 〒432-8561 静岡県浜松市中区城北 3-5-1

[‡] 東京大学大学院工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

E-mail: [†] {f0810035, tmtkane}@ipc.shizuoka.ac.jp [‡] {yamashita, asama}@robot.t.u-tokyo.ac.jp
[†] ishihara@ishilab.net

あらまし 現在、車両へのドライブレコーダや車載カメラの搭載が一般車でも普及しつつある。そのためドライブレコーダや車載カメラが撮影している画像、映像を利用したシステムが求められている。これらの要求に応じるために車両間での通信による画像、映像の共有を行う研究が行われている。しかし画像、映像を共有するだけでは利用者はシステムを使用することは難しい。そこで本論文では、複数の画像に対してカメラの位置姿勢推定を行いその結果を用いて3次元地図を生成する。そして3次元地図を元に利用者にとって必要な任意の視点画像を生成する手法を提案する。我々の手法では、画像解析ソフトであるBundlerとPach-based Multi-view Stereoを利用することでStructure from Motionを用いたカメラの位置姿勢推定と密な点群による3次元地図生成を行う。実験により室外環境での3次元地図生成と、任意視点の画像の生成を行い利用者に任意視点画像の提供可能なシステムであることを示した。

キーワード Structure from Motion, 3次元モデル, Image-based modeling

Generation of Arbitrary Viewpoint Images from Image Compilation by Estimation of Position and Pose of Mobile Camera

Takuya Oe[†] Atsushi Yamashita[‡] Toru Kaneko[†] Susumu Ishihara[†] Hajime Asama[†]

[†] Faculty of Engineering, Shizuoka University 3-5-1 Johoku, Naka-ku, Hamamatsu-shi, Shizuoka, 432-8561, Japan

[‡] School of Engineering, The University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan

E-mail: [†] {f0810035, tmtkane}@ipc.shizuoka.ac.jp [‡] {yamashita, asama}@robot.t.u-tokyo.ac.jp
[†] ishihara@ishilab.net

Abstract Currently, drive recorders and in-vehicle cameras have become popular with general automobiles. Hence, a system using images taken by such devices is required. To satisfy the demand, retrieval of the shared data in the system is important. In this paper, we propose a method for generating arbitrary viewpoint images for users based on the 3-D map which is constructed by estimating the position and pose of the camera for each viewpoint. Our method uses Bundler and Pach-based Multi-view Stereo of the image analysis software, and estimates the camera poses by Structure from Motion and generates the 3-D map by dense point groups. Experimental results showed the effectiveness of the proposed method.

Keyword Structure from Motion, 3D model, Image-based modeling

1. 序論

近年、車両へのドライブレコーダ(図1)や車載カメラの車両への搭載が一般化しつつある。これらの機器は走行中画像、映像を撮影している。車両間で通信を行いそれらを共有することができれば運転前、運転中の運転補助に使用することができる。したがって多数のカメラで取得した画像を利用して3次元地図を生成し、任意の視点から見た画像を生成するシステムが必要となる。



図1 ドライブレコーダ

しかし多数のカメラや移動カメラにより撮影された画像から任意の視点から見た画像を取得し利用者に提供する際には、各カメラの位置姿勢情報が必要となる。撮影された画像からカメラの位置姿勢を求める方法として、SfM (Structure from Motion) があげられる。SfM とは与えられた画像から、そこに映っている物体形状を復元すると同時にカメラの位置姿勢を推定する手法である。

また本研究で想定する画像は通信などで取得する画像となるため時系列順に並んでいないと考えられる。よって時系列で並んでいない画像に対しての SfM が必要となる。ここで本研究では時系列順に撮影された複数の画像を画像列、それに対し時系列順ではない複数の画像を画像群と呼ぶ。

カメラの位置姿勢推定手法である SfM や 3 次元の環境復元手法は多く存在する。例えば、SfM を用いたカメラの位置姿勢推定の方法として全方位画像列を入力とし、初期値を因子分解法で得て、これをバンドル調整法で精度向上させ形状および位置姿勢を同時に精度良く推定する松久らの方法[1]がある。しかし、この方法ではカメラの位置姿勢を推定する場合、画像列を使用する必要がある。

本研究ではドライブレコーダや車載カメラを想定しており、通信によって取得した画像は画像列でないと考えられる。そのため画像列に対応した手法ではなく画像群に対応した SfM を行いカメラの位置姿勢推定を行う必要がある。

3次元環境復元手法としては車載カメラによって撮影されたビデオを元に形状復元を行う手法が David らによって研究されている[2]。この手法は容易に運動推定が可能な車両を用いることで、走行経路の周辺環境をリアルタイムで 3次元構築することを可能としている。しかし広域な地図を構築する場合、1台の車両で全ての道を走行し画像を取得することは困難である。

上記問題点を解決するために、本研究では画像群に対応した SfM によりカメラの位置姿勢推定を行う。その結果を用いて 3次元地図を復元し、任意視点画像として適切な画像を生成する事で、利用者に提供する任意視点画像の生成を行うシステムを構築する。

2. 処理手順

本研究での処理は 5 段階に分けて行われる (図 2)。

まず対象とする環境において単眼カメラで画像を複数枚撮影する (2.1 節)。

次に取得した画像群の各画像に対して特徴点の抽出を行う。その後取得した特徴点をそれぞれ対応付ける (2.2 節)。

そして得られた特徴点を元に SfM によりカメラの位

置姿勢推定を行う (2.3 節)。

その後カメラの位置姿勢を利用して 3次元地図生成を行う (2.4 節)。

最後に利用者が利用したい任意視点の位置を決定し、その位置からの任意視点画像を 3次元地図から生成する (2.5 節)。

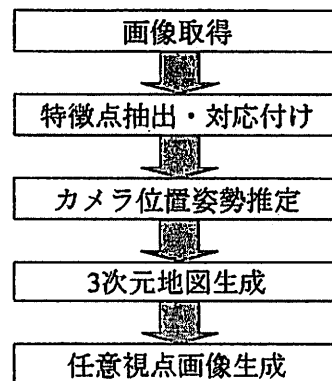


図 2 処理手順

2.1. 画像取得

撮影は単眼カメラを利用し対象とする環境で複数枚撮影を行う。通信によって画像を取得することを想定しているため、画像は十分な数の枚数を取得できるとする。またカメラの撮影方向や撮影位置は環境内であれば任意の方向であるとし、撮影画像は時系列順に並んだ画像列である必要はないとする。

2.2. 特徴点抽出・対応付け

撮影画像から特徴点を抽出・対応付けを行う。特徴点の抽出・対応付けには SIFT (Scale Invariant Feature Transformation) [3]を用いる。SIFT とは画像中の特徴点の位置と、その点を記述する 128次元の特徴ベクトルを抽出し、これによって画像の回転やスケール変化にロバストな特徴点抽出、対応付けを行う手法である。

2.3. カメラ位置姿勢推定

これまでの手順で得られた特徴点の対応を元に、カメラの位置姿勢推定を行う。カメラの位置姿勢推定には本研究では画像解析ソフトである Bundler [4]を用いる。カメラ位置姿勢推定手法は SfM である。

また入力画像は未校正であり、内部パラメータが未知であるため焦点距離を同時に推定する。

本手法で用いる SfM はカメラをピンホールモデルとして仮定し、カメラごとに推定するパラメータは焦点距離(f)、カメラ中心位置(c)、半径歪みパラメータ(k_1, k_2)、回転行列(R)、並進ベクトル(t)である。この時画像平面の原点を画像中心にとると、3次元点 $m = (m_x, m_y, m_z)$ は、画像点 $p = (p_x, p_y)$ に以下の式で変換される。

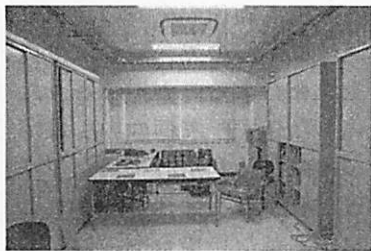
$$\mathbf{m}' = \mathbf{R}(\mathbf{m} - \mathbf{c}) \quad (1)$$

$$\mathbf{p} = (-fm'_x/m'_z, -fm'_y/m'_z) \quad (2)$$

$$\mathbf{p} = (1 + k_1|\mathbf{p}'|^2 + k_2|\mathbf{p}'|^4)\mathbf{p}' \quad (3)$$

2.4. 3次元地図生成

これまでの手順の SfM によって復元された 3 次元点群は、すべての画像で整合性の取れていた特徴点に対応する点のみが復元されるため、一般に非常に疎である (図 3)。そこで SfM で推定されたカメラパラメータを利用して、多視点ステレオ法で密な点群による 3 次元地図を復元する。本研究では、多視点ステレオ法の一つである Patch-based Multi-view Stereo [5] を利用する。実環境の例を図 3 (a) に示し、復元された 3 次元点群による 3 次元地図を図 3 (b) に示す。



(a) 実環境



(b) 復元結果

図 3 密な 3 次元点群による 3 次元地図

2.5. 任意視点画像生成

3 次元地図内の 3 次元座標と視線方向を利用者は任意で指定する。視線方向は 3 次元地図上において x 軸, y 軸, z 軸に対して利用者が見たい方向の角度を指定することで任意視点位置と方向を決定する。

任意視点の位置・姿勢を決定し、その位置からの画像を出力する。出力する画像は人間が周囲の状況を把握するのに最低必要な視野角は 110 度であることから視線方向を中心に左右へ 60 度ずつ広げたときに見える視野を画像として出力する (図 4)。

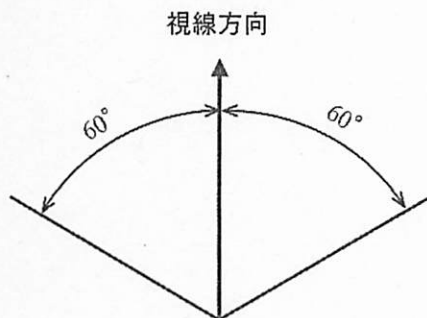


図 4 視野角

3. 実験

本研究の有効性を確認するための実験を行った。

3.1. 実験条件

図 5 に示す室外環境で実験を行った。カメラの位置姿勢推定を行うために使用した画像サイズは 2496 ×

1664pixel であり、使用画像枚数は 202 枚である (図 6)。



図 5 実験環境

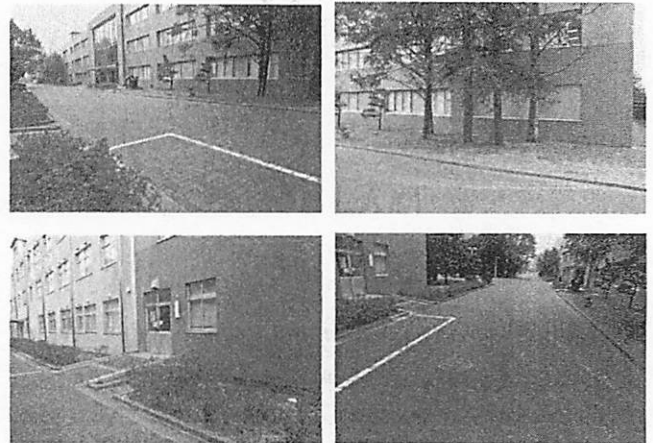


図 6 取得画像例

3.2. 実験結果

撮影された画像から Structure from Motion によってそれぞれのカメラの位置姿勢推定を行った。その後得られたカメラ位置姿勢の推定結果の情報と撮影画像を入力と情報として、Patch-based Multi-view Stereo より密な 3 次元点群による 3 次元地図生成を行った。図 5 に示す実験環境において生成された 3 次元地図を図 7 に示す。

その後 3 次元地図を生成してから任意視点画像の生成を行った。

任意視点として $(x, y, z) = (0.0, 1.0, -6.0)$ の位置, $x = 0.0^\circ$, $y = 0.0^\circ$, $z = 0.0^\circ$ の姿勢を与えた時の実画像を図 8(a) に示し, $(x, y, z) = (1.0, 1.0, 1.0)$ の位置, $x = 0.0^\circ$, $y = 270.0^\circ$, $z = 0.0^\circ$ の姿勢を与えた時の実画像を図 8(b) に示す。なお図 8(a), (b) に示す画像は 3 次元地図生成に使用した画像に含まれず検証用に新たに撮影した画像である。

それぞれの位置で作成された任意視点画像を図 9(a), (b) に示す。

次に今回作成した任意視点画像と実際の画像の比較を行う。図 8(a), (b) に図 9(a), (b) と同じ方向の実際

の画像をそれぞれ示す。

視線方向に壁などの対象物がある場合、つまり図 9 (b)の場合は任意視点画像として生成が比較的できている。しかし視線方向に壁などの対象物がなく、遠距離まで対象物がない場合、また木などの同じ特徴を持つ環境の場合、つまり図 9(a)の場合は点群として生成されなかった。

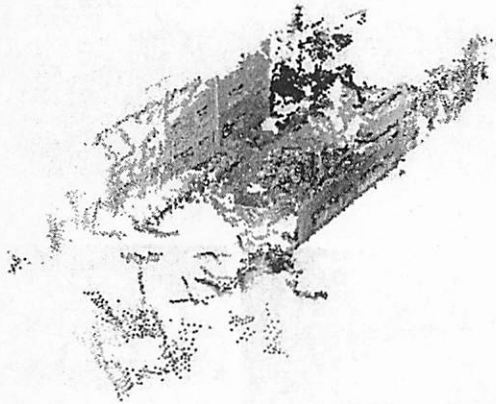


図 7 3次元地図生成結果



(a)視線方向 1

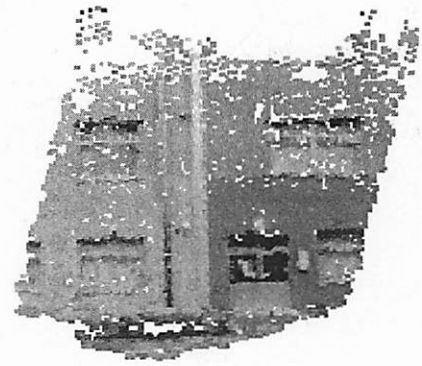


(b)視線方向 2

図 9 任意視点画像の方向の実際の画像



(a) 任意視点画像 1



(b) 任意視点画像 2

図 9 任意視点画像生成結果

4. 結論

本研究では、画像群からのカメラの位置姿勢推定、3次元地図生成、及び任意視点画像の生成の手法を提案した。実験結果より画像群からのカメラの位置姿勢を推定し、3次元地図生成を行えることが示された。また、3次元地図を元にした任意視点画像の生成が行えたことが確認できた。そのため利用者に任意視点の画像を提供するシステムとしての有効性が示された。

今後の課題としては壁などが無い場合への対応、点と点の間の補完を行うことが考えられる。

謝辞

本研究は科学研究費補助金基盤研究 B (23300024) 「リアルタイム画像カーナビのための効率的車々間データ配信技術」の助成による。ここに記して謝意を表す。

文 献

- [1] 松久亮太, 川崎洋, 小野晋太郎, 阪野貴彦, 池内克史: “因子分解法とバンドル調整による全方位画像列からの形状および位置姿勢の同時推定手法”, 画像の認識・理解シンポジウム予稿集, IS5-36, pp. 1610-1617(July 2008)
- [2] Davide Scaramuzza, Friedrich Fraundorfer, Rowland Siegwart: “Real-time Monocular Visual Odometry for Onroad Vehicles with 1-point Ransac”, Proceedings of the 2009 IEEE International Conference on Robotics and Automation, pp. 4293-4299(August 2009)
- [3] Davide G. Lowe: “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, 60, 2, pp. 91-110(January 2004)
- [4] Noah Snavely, Steven M. Seitz, Richard Szeliski: “Photo Tourism: Exploring Photo Collections in 3D”, Proceedings of the SIGGRAPH 2006 ACM Transactions on Graphics, 25, 3, pp. 835-846(July 2006)
- [5] Yasutaka Furukawa, Brian Curless, Steven M. Seitz: “Towards Internet-Scale Multi-view Stereo”, Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1434-1441(August 2010)