

フリーハンド撮影画像のカメラ運動を考慮したモザイクキング Mosaicing Freehand Images with Consideration of Camera Path

佐藤 靖之[†] 山本 強[†] 土橋 宜典[†]

Yasuyuki Sato Tsuyoshi Yamamoto Yoshinori Dobashi

1. はじめに

屋内もしくは屋外のシーンを様々な方向から撮影した画像を、それぞれの画像が連続性を保つように合成することで臨場感のあるパノラマ画像を生成する技術がある。このような技術はモザイク合成と呼ばれ、今までに多くの研究が行われている。これらの研究の多くは撮影されたシーンを近似的に平面として捉え、射影変換を施すことで自然な合成を実現している。

一般に、射影変換に用いられるホモグラフィ行列は8自由度を持っており、画像中の4点の対応関係を与えることで一意に決定することが可能である。しかし、この方法では画像の重複領域が少ない場合に計算が不安定になり、結果として合成した画像が不自然になるという問題が知られている。

本研究では画像間のホモグラフィ行列をカメラパラメータから導出する方法を採用する。ホモグラフィ行列の導出時に用いるカメラの外部パラメータとして、適切な深度と回転を与えることによって、一点透視投影に基づくパノラマ画像合成を行う手法を提案する。本手法により、多くの画像をモザイク合成に有効に利用することが出来るようになり、より広範囲にわたる画像の作成が可能となる。

2. 射影変換とホモグラフィ行列

三次元空間中のある平面を2つのカメラで観察した時、各画像中の対応点のカメラ座標値は射影変換によって互いに変換が可能である。射影変換に用いるホモグラフィ行列 H は、投影元の画像平面上の点 p_{src} と、投影先の画像平面上の点 p_{dst} との位置関係を以下の式で定義する。

$$p_{dst} = H p_{src}, p_{src} = H^{-1} p_{dst}$$

$$p_{src} = (x_{src}, y_{src}, 1)^T, p_{dst} = (x_{dst}, y_{dst}, 1)^T \quad (1)$$

ホモグラフィ行列 H の各要素を h_{ij} とすると、式(1)は式(2)のように表すことが出来る。

$$x_{dst} = \frac{h_{11} \cdot x_{src} + h_{12} \cdot y_{src} + h_{13}}{h_{31} \cdot x_{src} + h_{32} \cdot y_{src} + h_{33}}$$

$$y_{dst} = \frac{h_{21} \cdot x_{src} + h_{22} \cdot y_{src} + h_{23}}{h_{31} \cdot x_{src} + h_{32} \cdot y_{src} + h_{33}} \quad (2)$$

ホモグラフィ行列は平面の方程式、カメラの内部パラメータ及び外部パラメータを暗黙的に含んだ行列である。一方で、これらのパラメータが既知の場合、式(1)は式(3)のように表すことが出来る。

$$p_{dst} = K_{dst} \cdot \left(R + \frac{t \cdot n^T}{z} \right) \cdot K_{src}^{-1} \cdot p_{src} \quad (3)$$

よって、ホモグラフィ行列 H は式(4)で定義される。

$$H = K_{dst} \cdot \left(R + \frac{t \cdot n^T}{z} \right) \cdot K_{src}^{-1} \quad (4)$$

ここで K_{src}, K_{dst} は投影元及び投影先のカメラ内部パラメータであり、 R 及び t は投影元カメラを基準とした投影先カメラの外部パラメータ、 n は平面の単位法線ベクトル、 z は投影元カメラから平面までの距離である。

3. 提案手法の概要

本研究では1台のカメラで撮影されたフリーハンド撮影画像からのパノラマ画像の合成を目標としており、提案手法ではカメラキャリブレーション、画像の射影変換、イメージモザイクキングの3つの工程でこれを実現していく。カメラキャリブレーションでは事前のキャリブレーションと Structure from Motion と呼ばれる手法[2]によってカメラの内部パラメータ及び外部パラメータを取得する。画像の射影変換ではそれぞれの画像を基準となる画像のカメラ位置における画像へと変換する。これにより得られた画像をイメージモザイクキングすることによりパノラマ画像を得ることが出来る。

以下の章では、各工程についてより詳しく説明する。

4. カメラキャリブレーション

コンピュータビジョンの分野では、未校正カメラを用いた三次元復元について多くの研究が行われている。これらの研究は、未知の形状の物体を複数の方向から撮影することでカメラの内部パラメータ及び外部パラメータを求め、同時に対象の三次元形状の復元も可能とするものである。

本研究では、カメラの外部パラメータの取得にはこれらのセルフキャリブレーション手法を利用することを想定しており、本研究では Photo Tourism [3][4]の Structure from Motion の手法を使用して外部パラメータを取得している。カメラの内部パラメータについても、セルフキャリブレーションは可能だが、本研究では1台のカメラを使用して入力画像を取得しているため、カメラの内部パラメータは常に一定と見なすことが出来る。そこで、使用するカメラの内部パラメータは Zhang [5]の手法により事前にキャリブレーションし取得しておくことで精度を高めている。

5. 画像の射影変換

フリーハンド撮影により得られた入力画像集合は、それぞれの視点位置や視線方向が異なっている。そのため、入力画像集合をそのまま接合しようとしても、複数の視点位置からの画像が混在するため、不自然な画像となってしまう。この問題を解決し、自然な画像を合成するためには異

[†] 北海道大学大学院情報科学研究科 Graduate School of Information Science and Technology, Hokkaido University

なる視点位置の画像を基準となる視点位置から撮影した画像へと変換するプロセスが必要となってくる。

この変換のプロセスにおいて射影変換を利用する。しかし、射影変換はあくまで平面を観察している際に正しい結果が得られる変換であり、入力画像集合に撮影されているシーンは必ずしも平面とは限らない。そのため、ここではシーンがカメラから十分離れていると仮定する。このとき、シーンを近似的に平面と見なすことが出来る。これにより、射影変換によって視点位置の変換を近似的に行うことが可能となる。



図1 基準視点位置への射影変換

基準となる視点位置は任意に選択することが可能であるが、便宜上入力画像集合の中から基準となる画像を選択する。以降、この視点位置を基準視点位置、画像を基準画像と呼ぶ。基準画像を選択した後、他の入力画像と基準画像とのホモグラフィ行列を式(4)より個別に計算する。ここで、 n は基準画像平面と平行な面に対する単位法線ベクトルとして設定する。深度 z はシーンが無限遠に存在している平面だという仮定より、十分に大きな値を与えることで、深度の影響を減らすことが出来る。図1に基準視点位置への射影変換の結果を示す。図1(a)は選択した基準画像、(b)は射影変換前の画像、(c)が射影変換後の画像である。

また、基準画像と入力画像の位置関係などによっては、射影変換後の画像領域が極端に狭くなっていたり、映っていないかったりということが起こる可能性がある。このような場合、画像間の相対回転を表す行列 R を適切に設定しなおすことによって、モザイク合成に有効な画像を広範囲に渡り取得することが可能となる。

6. イメージモザイク

提案手法では画像の合成にはイメージスティッチングの手法[6]を用いた。この手法は、重複領域内において、画素値の差分をエネルギーとして定義し、重複領域内でエネルギーが最小になるラインを探索することで、違和感の少ない合成を行うことが出来る。提案手法では、射影変換により近似的な視点移動を行っているが、フリーハンド撮影画像では近似しきれずに画像間で微妙な違いが現れる可能性がある。こうした入力画像集合の性質から、より違和感の少ない合成結果が得られると考え、この手法を採用した。



図2 出力結果

7. 結果

図2に提案手法の出力結果を示す。図2は図1(a)を基準画像として、図1(b)を基準視点位置から視線方向を 45° 反時計回りに動かした視点へと射影変換を行った画像をモザイク合成した結果である。外部パラメータの推定のために、モザイク合成に使用した画像を含め23枚の入力画像を使用した。図2の前景にある建物及び木は違和感の無い合成が実現出来ているが、背景部分にアーティファクトが現れている。

8. まとめと今後の課題

本稿では、カメラパラメータからホモグラフィ行列を導出し、画像に射影変換を行うことでパノラマ画像を生成する手法を提案した。入力画像全体からStructure from Motionにより各カメラの外部パラメータを取得したことで、各画像のホモグラフィ行列を安定して取得することが可能となった。また、点対応よりホモグラフィ行列を導出する手法とは異なり、基準視点位置を指定することで入力画像集合には存在しない視点への射影変換も可能となっている。

今後の課題として、入力画像集合の視点位置のずれが大きな場合や、撮影するシーンがカメラから十分に離れていない場合への対応が考えられる。これには、撮影対象の三次元形状及び位置を考慮する必要がある。Structure from Motion手法により得られた点群の深度を利用するといった方法が考えられる。他にも、外部パラメータの推定精度の向上や、モザイクに用いる手法の工夫などが課題として考えられる。

参考文献

- [1] 蚊野 浩,金出 武雄.”任意のカメラ配置におけるステレオ視とステレオカメラ校正”,電子通信情報学会論文誌 D-II,J79-D-II-11(1996),1810-1818
- [2] R. Hartley, R. Gupta and T. Chang: “Stereo from Uncalibrated Cameras”, Proceedings of the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp761-764, 1992.
- [3] Zhengyou Zhang, “A Flexible New Technique for Camera Calibration”, Technical Report MSR-TR-98-71, Microsoft Research, 1998.
- [4] Noah Snavely, Steven M. Seitz, Richard Szeliski. ”Photo Tourism: Exploring image collections in 3D.” *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*, 2006
- [5] Noah Snavely, Steven M. Seitz, Richard Szeliski. “Modeling the World from Internet Photo Collections.” *International Journal of Computer Vision*, 2007.
- [6] Aseem Agarwala, Maneesh Agrawala, Michael Cohen, Avid Salesin, Richard Szeliski. ”Photographing long scenes with multi-viewpoint panoramas”. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006)*, 2006