

フリーハンド撮影画像群からの低歪超広角画像の合成法

An Image Stitching Method for low-distortion super-wide-angle Pictures Synthesis form Pictures Taken by free-handed Camera

笹本 亮介†
Ryosuke Sasamoto山本 強†
Tsuayoshi Yamamoto土橋 宜典†
Yoshinori Dobashi

1. はじめに

近年、メモリの大容量化や表示システムの高度化によって、より多くの情報を含む画像が扱われるようになってきた。そのような画像の中でも1枚の画像中に広範囲の情報を含むことができる超広角画像は魅力的なコンテンツであると言える。超広角画像を作成する手法としては専用のカメラを用いて、あるいは、複数枚の画像を合成するという2つの方法が挙げられるが、本研究ではレンズによる歪みを抑えられ、尚且つ特殊な機器を必要としない後者の手法に着目した。従来の画像合成法には2種類あり、カメラの位置や撮影方向を限定し画像中の物体の3次元情報を考慮せずに画像全体を1枚の平面として扱い合成する方法と、フリーハンドもしくはカメラの位置・方向を限定して撮影された画像に対して物体の3次元位置・形状を推定し3次元空間で合成を行うものがあつた。本稿では、以降、3次元位置・形状のことを3次元情報と述べることにする。そのうち前者では撮影画像が限定されるため狭い用途でしか利用できず、後者では3次元情報の正確な推定が困難なため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。カメラ位置・方向を制限せずに撮影した画像に対して、3次元情報を考慮せずとも画像同士を目立たない部分で合成することで自然な合成が可能であると考え、本稿ではフリーハンドで撮影した画像群を用いて物体の3次元形状を考慮せずに不自然さのない合成を目指す。本研究では、違和感のない画像合成が可能なイメージスティッチングという方法を応用し自然に超広角画像を合成する手法を提案する。

2. 関連研究

画像合成の手法はフリーソフトから研究機関で開発されているものまで幅広く世に出回っている。その中の1つである Agarwala らの手法[1]は複数の視点から撮影した大量の画像から長大なパノラマ画像を生成する手法である。この手法では画像群から自然なパノラマ画像を合成できるが、撮影範囲に対して平行に移動して撮影しなければならないという制限がある。また、Furukawa らの手法[2]はフリーハンド撮影された画像群に基づいて撮影範囲の3次元空間を構築し、その3次元情報を考慮した画像の位置合わせや合成を行っている。この手法では壁や地面などの簡単な平面領域に対しては自然な合成が可能であるが、家具などの形状の細かい物体に対しては3次元形状の推定が困難であるため合成結果が不自然になるという問題がある。本研究では、撮影位置・方向を制限することや3次元情報を利用せずに、画像の輝度情報のみを用いてより不自然さのない画像合成を目的とする。

3. 提案手法

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群からの超広角画像の自然な合成を目指している。一般的に視点位置の異なる画像の厳密な合成は被写体の3次元形状を考慮しないと不可能であるが、画像同士を目立たない部分で張り合わせることで、実際の形状とは誤差が生じるものの、自然な合成を実現する。提案手法は画像間の特徴点マッチング、画像の射影変換、エネルギーマップ作成、イメージスティッチングという4つの処理からなる(図1)。まず、画像群の特徴点抽出とマッチングを行い、画像間の位置関係の情報を得る。次に得られた位置関係の情報をを用いて画像を射影変換し、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して、重なる領域を算出する。また、本稿ではこの画像が重なっている領域のことを「マッチング領域」と呼ぶことにする。エネルギーマップ作成ではマッチング領域において、繋ぎ目にしても目立たない部分を算出する。イメージスティッチングでは前述の処理によって得られたエネルギーマップから画像を合成する際に最も違和感のない境界線を推定して張り合わせる。以下、各処理についてより詳しく説明する。

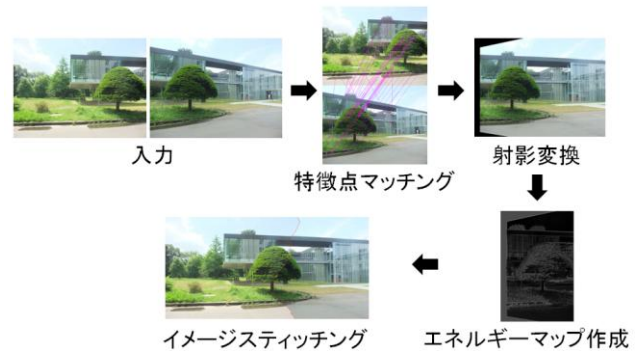


図1 提案手法の流れ

3.1 特徴点マッチング

画像からの特徴点抽出に関しては多くの研究が行われているが、本研究では SIFT 特徴量[3]を用いた。SIFT 特徴量の特長として、画像の回転、スケール変化、照明変化などに頑健であることが挙げられる。本研究では入力にフリーハンド撮影画像群を用いており、これらの画像変化が起こることは容易に想像できるので SIFT 特徴量は提案手法に最適であると考え、SIFT 特徴量を用いたマッチングを行う[3]。この処理を行うことによって画像間の対応点の位置関係が得られる。

3.2 画像の射影変換

フリーハンド撮影により得られた画像群はそれぞれの視点位置や視点方向が異なっている。そのため、撮影画

像をそのまま合成しようとしても不自然な結果となってしまうため、視点位置を統一させなければならない。本研究では射影変換を用いて視点位置の統一を実現する。射影変換とは画像内の平面を異なる視点から撮影された平面に見えるように変換する技術である。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} との位置関係を用いて以下の式で定義される。

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

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

特徴点マッチングにより得られた対応点の関係を(1)式に代入し、近似的に視点位置を統一する。しかし、実際には射影変換は主に平面に対して有効であり、本研究が想定している奥行きのある3次元空間に対しては完全に視点変換を行うことができない。3次元空間に対する完全な視点変換は3次元位置・構造を把握していないと不可能であるが、次節で述べる処理によってこの問題を補う。

3.3 エネルギーマップ作成

本研究では画像コンテンツの「重要度」を表すエネルギーマップを用いる。エネルギーは画素毎に定義され、値が高いほど人間の目には違和感を感じやすい画素となるので張り合わせの境界線には適していない。エネルギーの低い画素を連結して境界線とすることで自然な画像合成を可能とする。また、エネルギーは2つの画像の重複領域であるマッチング領域内のみ定義する。マッチング領域内の2つの画像の輝度値から各画素のエネルギーを算出し、エネルギーマップの作成を行う。入力画像を I_1, I_2 , x, y をそれぞれ画像中の x 座標, y 座標, エネルギーを E とすると、以下の式で表すことができる。

$$E_1 = \left| \frac{\partial}{\partial x} I_1 \right| + \left| \frac{\partial}{\partial y} I_1 \right|, \quad E_2 = \left| \frac{\partial}{\partial x} I_2 \right| + \left| \frac{\partial}{\partial y} I_2 \right| \quad (2)$$

$$E_M = |I_1 - I_2| \quad (3)$$

$$E = \alpha E_1 + \beta E_2 + \gamma E_M \quad (4)$$

E_1, E_2 ではそれぞれの画像内で注目画素の近傍の輝度値を比較し、輝度値の差が大きい部分が高エネルギー、つまり重要になるように定義されている。また、 E_M ではマッチング領域内において、各画素における2つの画像の輝度値の差分を計算してエネルギーマップとしている。

(2)式では各々の画像中で輝度値の変化が激しい部分が人間の目が注目し易い部分だと考え、そのような画素を境界線に選択しないようにしている。(3)式では2つの画像の輝度値を比較し、輝度の近い部分が境界線として選択されるようにしている。(4)式ではこれらのエネルギーを係数 α, β, γ で重み付けして加算しているが、係数はユーザが任意に指定することができ、画像に応じて適切な値を設定することが可能である。これにより境界線に適した画素を定義することが可能である。本節で定義されたエネルギーマップは、画像の種類やシーンなどを考慮し改良することによって更なる精度の向上を図ることができる。

3.4 イメージスティッチング

本研究では画像を境界線で張り合わせるイメージスティッチングという手法を用いる。境界線の選択方法には

多くの研究がなされているが提案手法では3.3節で述べたエネルギーマップにおいて、境界線上にある画素に割り当てられているエネルギーを累積していき、その累積値が最も低くなるように境界線を選択した [4]。

4. 実験結果

提案手法により得られた結果を図2に示す。図2(a), (b)はフリーハンドで撮影された入力画像, (c)は射影変換を行った画像である。結果画像(図2(d))では照明条件の変化によるエッジは見られるものの境界線に違和感がなく自然に合成できていることがわかる。また、図2(e), (f)はそれぞれ E_M, E をエネルギーマップとして用いてイメージスティッチングを行った際の部分拡大画像である。赤く表示されているラインが境界線である。図2(f)の方が自然に合成されていると言える。

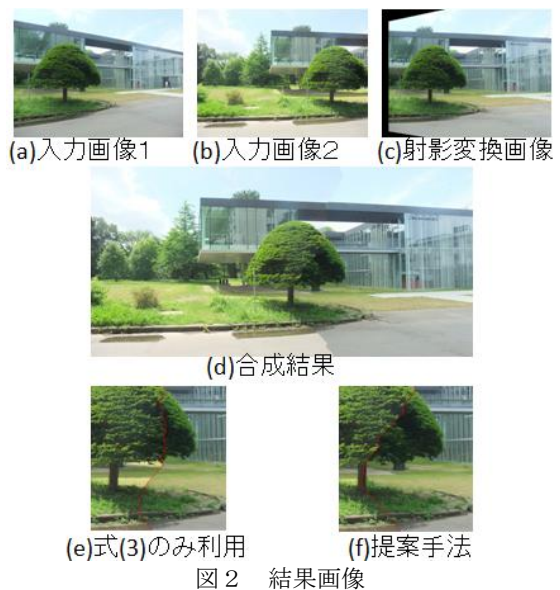


図2 結果画像

5. まとめと今後の課題

本稿ではフリーハンドで撮影した画像群を用いて物体の3次元形状を考慮せずに超広角画像を合成する手法について提案した。提案法を用いることによりフリーハンドで撮影された画像群に対しても違和感なく合成が可能となった。今後の課題として、照明条件を考慮することや画像のシーンに合わせたエネルギーマップの改良が挙げられる。

参考文献

- [1] 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.
- [2] Yasutaka Furukawa, Brian Curless, Steven M. Seitz, Richard Szeliski, "Reconstructing Building Interiors from Images", ICCV 2009 (oral).
- [3] David G. Lowe, "Distinctive image features from scale-invariant keypoints", *Journal of Computer Vision*, 60, 2, pp. 91-110, 2004.
- [4] Avidan, S., and Shamir, A. 2007. Seam carving for content-aware image resizing. *ACM Trans. Graph.* 26, 3, 10.