

LI-003

紙面を対象としたカメラパス推定による高解像ビデオモザイクング High-resolution Video Mosaicing for Documents by Estimating Camera Parameters

佐藤 智和^{†,‡}
Tomokazu Sato

池田 聖[†]
Sei Ikeda

神原 誠之^{†,‡}
Masayuki Kanbara

池谷 彰彦[‡]
Akihiko Iketani

中島 昇[‡]
Noboru Nakajima

横矢 直和^{†,‡}
Naokazu Yokoya

山田 敬嗣[‡]
Keiji Yamada

1. はじめに

近年のインターネット・カメラ付き携帯電話や PDA の普及に伴い、紙面に描かれた文章を手軽に電子化し、保存・伝送することに対する需要が高まっている。現在、このような目的のためには、画像スキャナのように、画像センサを紙面の上で機械的に等速に動かすことでスキャンを行なう装置を用いることが一般的であるが、機材が大きいために手軽とは言いがたい。

そこで、ビデオカメラを用いて紙面を近接位置からなぞるように撮影し、ビデオモザイクングを行なうことで紙面を電子化することが考えられる。平面を対象とするビデオモザイクングに関する研究は、従来から多く成されているが、なかでも、動画像上の自然特徴点を用いる手法 [1, 2] は、比較的計算コストが小さいという特長がある。このような従来研究では、画像間の変換パラメータとして、カメラの内部パラメータを含む 8 自由度の射影行列を算出し、これらを用いてモザイク画像を生成している。しかし、変換パラメータの自由度が大きいため、累積的な推定誤差が大きく、多数の画像を用いる場合には生成画像が歪むという問題がある。

これに対し、本研究ではカメラの内部パラメータをあらかじめ推定し、画像間の変換パラメータを 6 自由度のカメラ外部パラメータとして推定することで、累積的な推定誤差を抑える。加えて、画像上で再出現したと考えられる自然特徴点を自動で対応付け、カメラ外部パラメータの全体最適化処理を行なうことで累積誤差による歪みを最小化する手法を提案する。なお、本手法では、初期フレームにおいて、カメラが紙面におおむね正対しているものとする。

2. カメラパス推定によるビデオモザイクング

図 1 に、提案手法の処理の流れを示す。本手法では、まず図 2 に示すように、カメラを紙面に近接させたままなぞるように撮影を行ない、動画像中の自然特徴点を時系列順に自動追跡することで、逐次的にカメラの外部パラメータを復元する (A)。ここで、同図中に × 印で示すような位置に存在する自然特徴点は、動画像中で一度消失した後に再出現するため、次に、このような自然特徴点の検出および対応付けを行ない (B)、自然特徴点の三次元座標とカメラパラメータを動画像全体で最適化する (C)。最後に、推定されたカメラパラメータを用いてモザイク画像を生成する (D)。以下では、各処理について詳述する。

(A) カメラの外部パラメータの逐次推定

我々が従来提案した特徴点追跡によるカメラパラメータの推定手法 [3] を利用してカメラの外部パラメータを逐

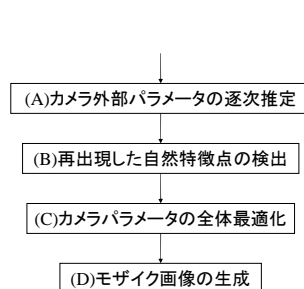


図 1: 処理の流れ

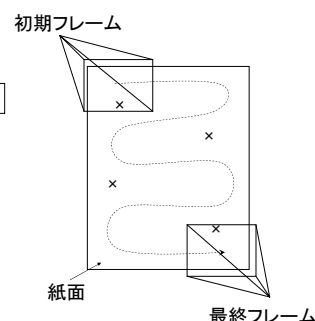


図 2: 紙面の撮影方法

次に復元する。同手法では、初期フレームにおいて三次元位置が既知のマーカを用いることで、自然特徴点の出現・消失の頻発する動画像に対してもロバストにカメラの外部パラメータを復元することができる。本研究では、初期フレームにおいてカメラが紙面におおむね正対するという制約および、対象が平面であるという制約を設けることで、三次元位置が既知のマーカを用いることなくロバストにカメラの外部パラメータを復元する。

(B) 再出現した自然特徴点の検出

逐次処理で追跡された自然特徴点に対して、再出現の検出を行なう。ここでは、初期フレームの画像面に平行な投影面上において、全ての自然特徴点のテンプレートを作成し、自然特徴点の空間的な距離が閾値以下の自然特徴点の組に関してのみ、正規化相関を尺度とする評価値を算出する。算出された評価値が閾値以上の自然特徴点の組は、カメラの動きによって画像上から消失し、再出現した自然特徴点であると考えられるため、これらを同一の自然特徴点とみなし三次元位置を統合する。このような投影面上でのマッチングを行なうことで、カメラの移動によるテンプレートの変形に対応することができる。

(C) カメラパラメータの全体最適化

カメラパラメータと自然特徴点の三次元位置を動画像全体で最適化することで、累積的な推定誤差を最小化する。最適化の誤差関数には、自然特徴点の画像上での追跡座標と、各フレームの画像面上への自然特徴点の三次元位置の投影座標の自乗距離の和を用いる。これにより、初期フレームの画像面に対する投影平面の傾きも同時に補正される。

(D) モザイク画像の生成

全てのフレームの画像を (B), (C) で用いた投影平面上に投影することでモザイク画像を生成する。ここでは、各フレームの投影中心から各画素への自乗距離の逆数を重みとする重み付きのブレンディング処理を行なう。これにより、各フレームでの画像の色の違いや幾何学的な位置ずれによる見た目の違和感を減少させる。

[†] 奈良先端科学技術大学院大学 情報科学研究科

[‡] NEC インターネットシステム研究所

3. 実験

手持ちのビデオカメラを紙面に接近させて撮影した実際の動画像を用いて、モザイク画像の生成実験を行なった。撮影には、ワイドレンズ (Sony VCL-HG0758) を取り付けた一台のビデオカメラ (Sony DSR-DP-150) を用い、図 3 を含む 202 枚の画像から成る 13 秒の動画像 (解像度: 720×480 画素, プログレッシブスキャン, 15fps) を得た。ただし、カメラの内部パラメータは、あらかじめ Tsai の手法 [4] により算出した。

まず、得られた動画像に対して前章で述べた (A), (B), (C) の順にオフラインで手法を適用することにより、図 4 に示すカメラの外部パラメータを得た。同図中の点群は、自然特徴点の三次元位置を、曲線は推定されたカメラパスを、鍾台は 10 フレーム毎のカメラの姿勢を表している。処理 (A) においては、各フレームで平均 74 個の自然特徴点が平均 20 フレーム間自動で追跡され、また処理 (B) においては、151 点の自然特徴点が統合された。本実験では、計算機に PC (Pentium 3 GHz, Memory 1GB) を用い、(A), (B), (C) の処理に要した時間は、202 フレームに対して、それぞれ 38 秒, 16 秒, 150 秒であった。

次に、推定されたカメラパラメータを元に、前章 (D) で述べた手法によって、図 5 に示す解像度 1600 × 2180 のモザイク画像を生成した。同図から、推定誤差の累積による画像歪みの影響は小さく、良好な画像が得られていることが確認できる。しかし、モザイク画像の一部で文字のブレが見られた。これは、再出現した自然特徴点が正しく対応付けられなかったためであると考えられるため、今後、処理 (B) に関してよりロバストな手法の検討が必要である。

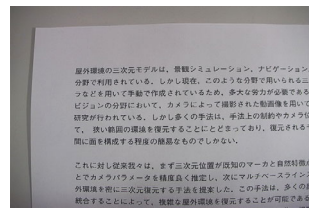
図 6 に、入力動画像の初期フレームの一部を拡大した画像と、モザイク画像の対応する箇所を拡大した画像を示す。同図から、モザイク画像の一部において、入力画像よりも高解像な画像が得られていることが確認できる。今後、超解像の手法を適用することで、より鮮明で高解像なモザイク画像の生成が可能であると考えられる。

4. まとめ

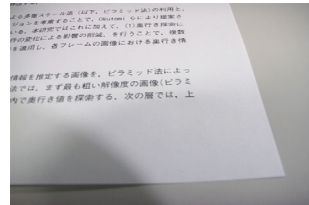
本稿では、ビデオカメラを用いて紙面を撮影することで、高解像モザイク画像として紙面を電子化する手法を提案した。実験では、実際に手持ちのカメラで紙面を撮影し、歪みの小さい画像を生成できることを確認した。今後は、処理のリアルタイム化、再出現した自然特徴点の検出手法の改良、超解像手法の適用が課題である。

参考文献

- [1] 千葉, 蚊野, 美濃, 安田: “画像特徴に基づくイメージモザイク”, 電子情報通信学会論文誌 (D-II), Vol. J82-D-II, No. 10, pp. 1581-1589, 1999.
- [2] 杉岡, 山地, 石川: “ドキュメントを対象としたビデオモザイクの検討”, 情報科学技術フォーラム (FIT) 一般講演論文集, Vol. 3, No. J-17, 2002.
- [3] T. Sato, M. Kanbara, N. Yokoya and H. Takemura: “Dense 3-D Reconstruction of an Outdoor Scene by Hundreds-baseline Stereo Using a Hand-held Video Camera,” Int. Journal of Computer Vision, Vol. 47, No. 1-3, pp. 119-129, 2002.
- [4] R. Y. Tsai: “An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 364-374, 1986.



初期フレーム



最終フレーム

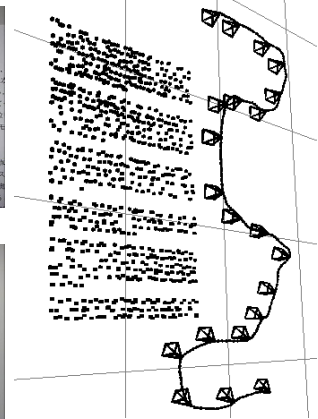


図 3: 入力動画像の一部 図 4: 推定されたカメラの位置と姿勢

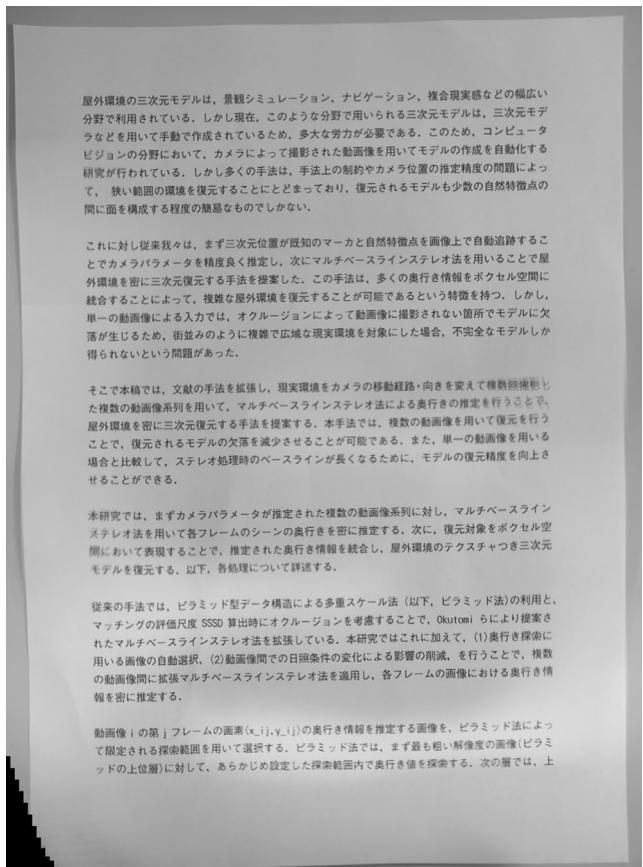
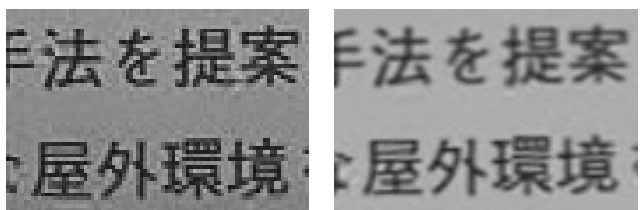


図 5: 生成されたモザイク画像



初期フレームの一部 モザイク画像の一部

図 6: 入出力画像上での解像度の比較