

手持ちカメラによるビデオモザイクおよび超解像

Video Mosaicing and Superresolution with Handheld Camera

池谷 彰彦[†]
Akihiko Iketani

中島 昇[†]
Noboru Nakajima

佐藤 智和^{†,‡}
Tomokazu Sato

池田 聖[‡]
Sei Ikeda

神原 誠之^{†,‡}
Masayuki Kanbara

横矢 直和^{†,‡}
Naokazu Yokoya

山田 敬嗣[†]
Keiji Yamada

1. はじめに

カメラ付き携帯電話は、人物、風景のみならず、雑誌紙面や時刻表といった文書の撮影にも利用されるようになった。しかし、A4 文書をファクシミリ画質で撮影するには、400 万画素程度の解像度が必要であり、OCR を適用する場合には、さらに 4 倍以上の解像度が必要となる。

この問題に対し、ビデオカメラを用いて文書を近接位置からなぞるように撮影し、撮影された部分画像を逐次貼り合わせるビデオモザイクが提案されている[1][2]。

我々は、カメラパス推定によるビデオモザイク[2]と、超解像を組み合わせることで、高解像なモザイク画像を生成する手法を開発し、評価を行なった。本稿では、その結果について報告する。

2. 提案手法

本手法の処理の流れを図 1 に示す。本手法は、カメラパラメータ推定部と、超解像画像生成部とから構成される。カメラパラメータ推定部では、画像中の特徴を追跡することによって、各フレームにおけるカメラの運動パラメータを推定する[2]。超解像部では、カメラパラメータに基づいて、各フレームを初期フレームの画像面に平行な投影平面上に投影し、さらに複数フレーム間で重複して撮影されている領域に対して、Iterative Back Projection[3]による超解像処理を適用する。以上の処理により、高解像なモザイク画像を生成する。以下では、カメラパラメータ推定部、超解像部の処理について述べる。

2.1. カメラパラメータの推定

はじめに、ユーザは図 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う。この場合、カメラ運動は、紙面に平行な移動だけでなく、垂直方向への平行移動、および回転を伴ったものとなる。そこで、Harris のインタレストオペレータ[4]を用いて画像中の特徴点を検出し、時系列順に追跡することで、逐次的にカメラパラメータを推定する。しかし、この方法では、カメラパラメータの推定誤差が累積するという問題がある。そこで、本手法では、次の処理によって、推定されたカメラパラメータを補正する。

特徴点のうち、図 1 中に×印で示すような位置に存在するものは、動画の中で一度消失した後に再出現する。このような特徴点の検出および対応付けを行い、この対応付けの結果に基づいて、特徴点の三次元座標とカメラパラメータを補正する。

[†] NEC インターネットシステム研究所

[‡] 奈良先端科学技術大学院大学 情報科学研究科

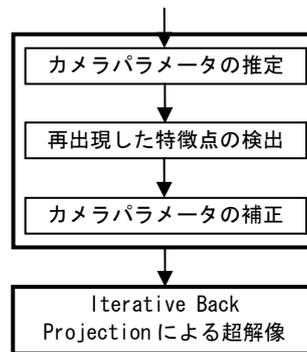


図 1 処理の流れ

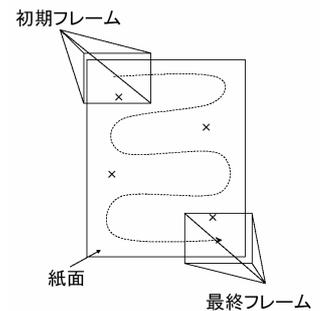


図 2 紙面の撮影方法

手持ちのカメラで紙面を撮影した動画を図 3 に示す。本手法によって推定されたカメラパラメータを図 4 に示す。図中の点群は自然特徴点の三次元位置、曲線は推定されたカメラパス、錐台は 10 フレーム毎のカメラ姿勢を表す。

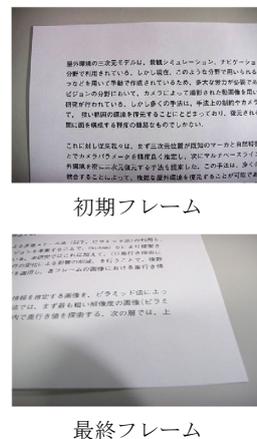


図 3 入力画像

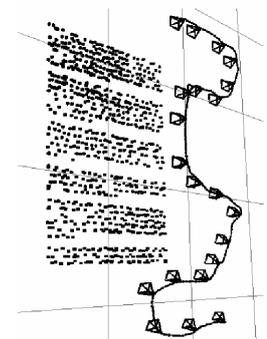


図 4 推定されたカメラパラメータ

2.2. 超解像

超解像は、重なりをもつ複数の低解像度画像から高解像度画像を生成する手法である。Irani らは、反復計算により、低解像度画像から高解像度画像を生成する Iterative Back Projection という手法を提案している[3]。

いま、 m 枚の入力画像 $\{g_k\} = g_1, g_2, \dots, g_m$ が存在するとする。生成したい高解像度画像を f とし、 f と各入力画像 $\{g_k\}$ 間の幾何変換は既知とする。

はじめに、高解像度画像の初期画像 $f^{(0)}$ を推定し、幾何変換、ぼかしから成る低解像度画像生成過程をシミュレートすることにより、各入力画像 $\{g_k\}$ に対応する低解像

度画像 $\{g_k^{(0)}\}$ を推定する. ここで, もし $f^{(0)}$ が正しい高解像度画像であれば, 推定画像 $\{g_k^{(0)}\}$ は実際に入力画像 $\{g_k\}$ と一致することになる. 逆に, $f^{(0)}$ の推定誤差が大きければ大きいほど, $\{g_k^{(0)}\}$ と $\{g_k\}$ のずれは大きくなる. そこで, 推定画像と入力画像間の差分画像 $\{g_k - g_k^{(0)}\}$ を計算し, 差分画像の画素値を $f^{(0)}$ 中の対応する箇所に逆投影 (Back Projection) し, 加算することで, 更新された高解像度画像 $f^{(1)}$ を得る. この処理を, 次式で表される誤差関数が十分に小さくなるまで繰り返す.

$$e^{(n)} = \sqrt{\sum_k \sum_{(x,y)} (g_k(x,y) - g_k^{(n)}(x,y))^2}$$

ここで, n は, 反復計算の回数を表す.

本研究では, 上記手法を用いて高解像なモザイク画像を生成する. ここで, 高解像度画像 f は, 初期フレームの画像と平行な平面上にとる. 高解像度画像の初期画像 $f^{(0)}$ については, 全ての画素において画素値 0 が与えられた画像を用いる. 高解像度画像と各入力画像 $\{g_k\}$ 間の幾何変換は, 各フレームにおけるカメラパラメータから求める.

3. 実験

本手法の有効性を検証するために, 実画像と, シミュレーション画像の両方について, 実験を行った.

まず, 図 3 に示す実画像について実験を行った. 解像度は 720×480 画素, フレーム数は 202 枚であり, フレームレートは 15fps であった. 本手法によって生成されたモザイク画像を図 5 に示す (紙面の都合上, 上半分のみ掲載). また, 入力画像の一部を拡大した画像を図 6 に, モザイク画像中の対応箇所を拡大した画像を図 7 に示す. モザイク画像では, 若干のぼけがあるものの, 入力画像中で潰れてしまっていた文字の穴やストロークが, 復元できているのが確認できる. なお, 本実験に要した処理時間は, 約 63 秒であった (CPU: Pentium4 2.4GHz, メモリ: 1GB 搭載 PC で計測).

続いて, シミュレーション画像について実験を行った. 本実験では, スキャナで取り込まれた文書画像を仮想の三次元空間に配置し, それを左から右へ移動するカメラで描写した場合の動画像を人工的に生成し, この動画像からモザイク画像を生成した. 本実験の場合, 生成されたモザイク画像と, オリジナル画像を比較することで, モザイク画像の画質を定量的に評価することができる. そこで, 実験に用いる入力画像のフレーム数を 30, 20, 10 と減らしていった場合の画質の変化について検証を行なった. 画質の評価は, オリジナル画像とモザイク画像との画素値の平均自乗誤差で行なった. また, 参考までに, オリジナル画像と入力画像との平均自乗誤差についても計算した. 表 1 に, その結果を示す. 入力枚数が減少するにつれ, 平均自乗誤差が増加している, すなわち画質が悪化しているものの, 10 枚の入力画像でモザイク画像を生成した場合でも, 平均自乗誤差が入力画像の 1/2.7 程度に抑えられているのが確認できる.

モザイク生成に 使用した枚数	30	20	10	(参考) 入力画像
平均自乗誤差	368.7	379.7	396.6	1081.4

表 1 オリジナル画像とモザイク画像の平均自乗誤差

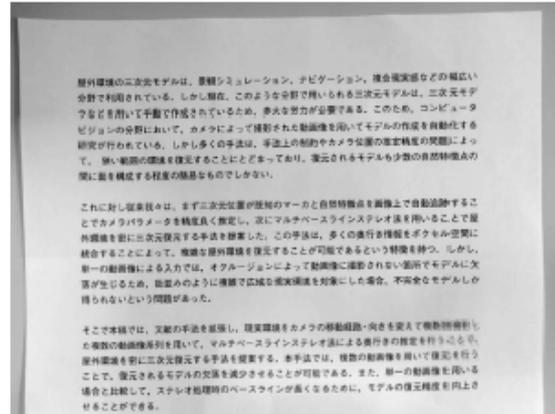


図 5 生成されたモザイク画像 (上半分のみ)

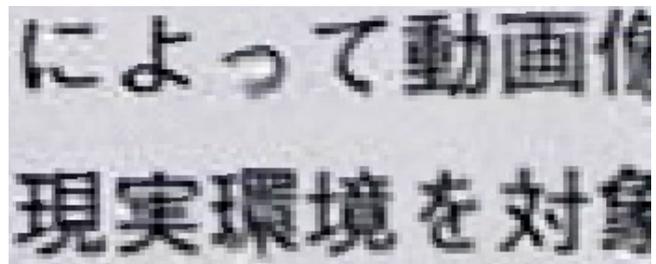


図 6 入力画像の一部

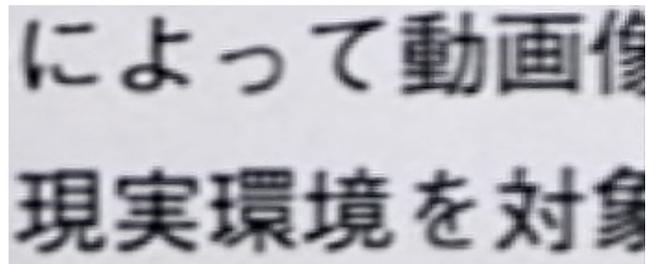


図 7 モザイク画像の一部

4. まとめ

本稿では, カメラパラメータ推定によるモザイクと, 超解像を組み合わせて, 手持ちカメラで撮影された動画像から, 高解像モザイク画像を生成する手法を提案した. 今後は, 処理の高速化, およびぼけを生じることなくモザイク画像を生成する超解像手法の検討が課題である.

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] 佐藤, 池田, 神原ほか, “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”, 情報科学技術フォーラム (FIT) 論文集掲載予定, 2003.
- [3] Irani, Peleg, “Improving Resolution by Image Registration”, CVGIP : Graphical Models and Image Processing, Vol.53, pp.231-239, 1991.
- [4] Harris, Stephens, “A Combined Corner and Edge Detector”, Proc. Alvey Vision Conf., pp.147-151, 1988.