

カメラパラメータ推定による紙面を対象とした超解像ビデオモザイク

池谷 彰彦[†] 佐藤 智和^{†,††} 池田 聖^{††} 神原 誠之^{†,††}
中島 昇[†] 横矢 直和^{†,††}

Super-Resolved Video Mosaicing for Documents by Camera Parameter Estimation

Akihiko IKETANI[†], Tomokazu SATO^{†,††}, Sei IKEDA^{††}, Masayuki KANBARA^{†,††},
Noboru NAKAJIMA[†], and Naokazu YOKOYA^{†,††}

あらまし 本論文では、手持ちのビデオカメラで紙面を撮影した動画から、紙面とカメラの三次元的な位置関係を表すカメラ外部パラメータを推定し、これに基づいて紙面のモザイク画像を生成する手法を提案する。カメラ外部パラメータ推定に基づく本手法は、初期フレームにおいてカメラが紙面に対して傾いている場合でも、正対したモザイク画像を生成できるという特徴をもつ。本手法では、はじめに、各フレームのカメラ外部パラメータを特徴点追跡結果から推定する。次に、映像から一度消失した後、再出現した特徴点を検出し、これを用いてカメラパラメータを補正する。これにより、累積的な推定誤差を最小化する。最後に、カメラパラメータを用いて、全フレーム画像をモザイク画像平面上に投影する。その際、Iterative Back Projection アルゴリズムを適用することで、超解像モザイク画像を生成する。実験により、本手法の有効性を確認した。

キーワード ビデオモザイク、カメラ外部パラメータ復元、超解像

1. ま え が き

大容量記憶装置の低価格化やインターネットの普及により、紙面に描かれた文章や写真を個人が手軽にデジタル化し、画像データとして保存・伝送することが可能となりつつある。現在、このような目的のためには、フラッドベッドスキャナのように、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持ち運びには適していない。そこで、持ち運び可能なカメラ付き携帯電話などのカメラ付き機器を用いて、手軽に高精細な画像を取り込む技術が求められている。しかしながら、一般的なカメラ付き機器で得られる画像の解像度は、A4 ページを一度に撮影することを想定すれば 150 dpi 程度であり (200 万画素のカメラの場合)、フラット

ベッドスキャナの数百～数千 dpi に対して著しく解像度が不足している。

このような問題に対して、紙面を複数枚の画像またはビデオ映像として拡大撮影し、各画像の位置合せをすることで、モザイク画像と呼ばれる 1 枚の高解像度画像を生成するモザイク技術に関する研究が行われている。なかでも、ビデオを入力としたビデオモザイクは、フレームごとの画像の変異が小さいために画像間の変換パラメータをロバストに推定できることや、多くの画像を一度に用いることでモザイク画像の品質を改善できることなどの理由から、近年盛んに研究されている。

ビデオモザイクの分野においては、従来から様々な手法が提案されている。Szeliski らは、ホモグラフィと呼ばれる 8 自由度の射影変換行列を用いて、平面に対するビデオモザイクを行う手法を提案している [1]。この手法では、連続する 2 枚の入力画像の組に対して、輝度値の差の二乗和が最小となるように、ホモグラフィパラメータを算出している。このホモグラフィベースのビデオモザイク手法は、後に計算量の削減、推定精度の向上など、様々な改良が加えら

[†] NEC インターネットシステム研究所, 生駒市
NEC Internet System Laboratories, 8916-47 Takayama,
Ikoma-shi, 630-0101 Japan

^{††} 奈良先端科学技術大学院大学情報科学研究科, 生駒市
Graduate School of Information Science, Nara Institute of
Science and Technology, 8916-5 Takayama, Ikoma-shi, 630-
0192 Japan

れている [2] ~ [7] . 特徴点を用いる手法 [2] ~ [4] では、画像に含まれる特徴点のみを用いてホモグラフィを計算するため、計算量の劇的な削減に成功している . また、画像系列全体での最適化を行う手法 [6], [7] では、画像間の変換パラメータの誤差の累積を最小限に抑えることで、モザイク画像の生成品質を高めている .

しかしながら、これらの手法では、あるリファレンス画像 (通常、初期フレーム) に対してモザイク画像平面を張り、このモザイク画像平面上で入力画像の貼り合わせを行っている . そのため、リファレンス画像平面が紙面と平行でない場合、生成されるモザイク画像も傾いたものとなるという問題があった . この問題が発生している例を図 1 に示す . これは、A4 紙面に印刷された文書のモザイク結果であり、図中、矩形で示された領域がリファレンス画像に対応する領域である . リファレンス画像撮影時にカメラが紙面に対して傾いているため、モザイク画像全体にわたって射影ひずみが発生しているのが確認できる .

ホモグラフィベースの手法においては、変換パラメータ推定のための誤差関数がこの傾きと無関係であるため、本質的にこの問題を解決することはできない . そのため、従来手法を用いて射影ひずみのないモザイク画像を生成するには、初期フレーム撮影時に、紙面とカメラを厳密に正対させる必要がある . しかしながら、手持ちのカメラでこれを実現するのは困難であり、この問題が紙面を対象としたモザイクを実用化する上での障害の一つとなっていた .

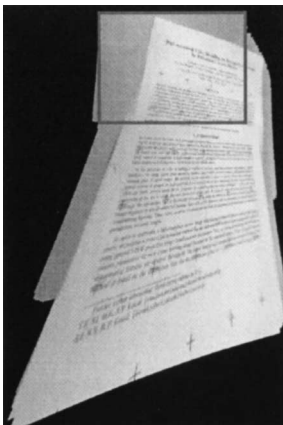


図 1 従来手法によって、モザイク画像に射影ひずみが発生した例

Fig.1 Mosaic image with perspective distortion generated by previous method.

カメラが紙面に対して傾いている場合でも、射影ひずみのないモザイク画像を生成するためには、カメラの位置、姿勢、及びモザイク画像平面との幾何学的な関係を正しく推定することが必要である . 三次元復元分野では、特徴点追跡結果から、シーンと各フレームにおけるカメラの幾何学的な関係を表すカメラ外部パラメータを推定する Structure from Motion と呼ばれる手法が提案されている [8], [9] . また、シーンを撮影した複数の入力画像を、カメラ外部パラメータ推定結果に基づいてカメラに正対した平面に射影することで、シーン中の建物のテクスチャ画像を獲得する手法も提案されている [10] .

本論文では、Structure from Motion 手法を紙面に対して適用することで、各フレームにおけるカメラ外部パラメータを推定し、これを用いることで、カメラが紙面に対して正対していない場合でも、射影ひずみのない高解像度のモザイク画像を生成する手法を提案する . 更に、紙面全体をカメラで漏れなく撮影する際、紙面上の特徴点が頻繁に再出現する点に着目し、このような再出現特徴点を利用してカメラ外部パラメータを補正する手法についても提案する . 従来研究に対する手法面での新規性はこの二点にあり、このような工夫を加えることにより、紙面を対象とした実用レベルに近いビデオモザイク方式を開発した .

提案手法の概要は、次のとおりである . はじめに、特徴点追跡結果から、各フレームのカメラ外部パラメータを推定する . 次に、映像から一度消失した後、再出現した特徴点を自動検出し、消失前と再出現後の追跡結果を同一特徴点系列に統合する . 続いて、映像全体でカメラパラメータの最適化を行い、推定誤差を最小化する . 最後に、全フレーム画像をモザイク画像平面上に投影する . その際、超解像処理を適用することで、高解像度モザイク画像を生成する . なお、本手法では、紙面が平面であり、また、カメラの内部パラメータは既知とし、撮影中は変化しないものとする .

2. カメラパラメータ推定による超解像ビデオモザイク

図 2 に、提案手法における処理の流れを示す . はじめに、紙面を近接位置からなぞるように動かしたカメラの外部パラメータを推定し (A) , 推定されたパラメータを用いて超解像モザイク画像を生成する (B) .

以下では、カメラ外部パラメータと誤差関数を定義した後、前記 (A) (B) の詳細について述べる .

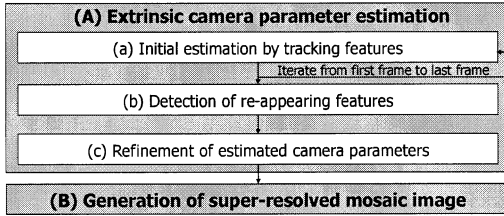


図 2 提案手法の処理の流れ
Fig. 2 Flow diagram of the proposed method.

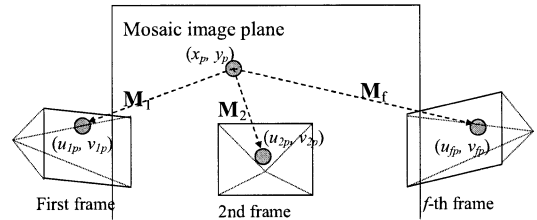


図 3 モザイク画像平面とカメラの関係
Fig. 3 Mosaic image plane and camera.

2.1 カメラ外部パラメータと誤差関数

ここでは、本手法で用いるカメラ外部パラメータと、その推定に用いる誤差関数について述べる。

本研究では、図 3 に示すように、モザイク画像平面と第 f フレームにおけるカメラの関係を表す変換行列 M_f を、カメラの位置パラメータ (t_{1f}, t_{2f}, t_{3f}) 及び姿勢パラメータ (r_{1f}, r_{2f}, r_{3f}) を用いて、以下のよう

$$(a\hat{u}_{fp}, a\hat{v}_{fp}, a)^T = M_f(x_p, y_p, 1)^T \quad (1)$$

$$M_f = \begin{pmatrix} c_1c_3 + s_1s_2s_3 & s_1c_2 & t_{1f} \\ -s_1c_3 + c_1s_2s_3 & c_1c_2 & t_{2f} \\ c_2s_3 & -s_2 & t_{3f} \end{pmatrix} \quad (2)$$

$$s_i = \sin(r_{if}), \quad c_i = \cos(r_{if}) \quad (i = 1, 2, 3) \quad (3)$$

ただし、

a : 媒介変数、

(x_p, y_p) : 特徴点 p のモザイク画像平面上での座標、

$(\hat{u}_{fp}, \hat{v}_{fp})$: 第 f フレームにおける理想カメラ画像への (x_p, y_p) の投影座標、

(u_{fp}, v_{fp}) : 第 f フレームにおける実画像平面上での特徴点 p の座標。

ここで、理想カメラ画像とは、入力画像からレンズひずみを取り除いた画像であり、理想カメラ座標における特徴点 p の投影座標 $(\hat{u}_{fp}, \hat{v}_{fp})$ は、既知のカメラ内部パラメータ(焦点距離, アスペクト, 投影中心, レンズひずみ)を用いて実画像平面上の座標 (u_{fp}, v_{fp}) に変換される。

本手法においては、対象を平面としているため、モザイク画像平面を $z = 0$ と定義すれば、式 (2) によって定義されるカメラの変換パラメータ M_f は、一般的な 4×4 のカメラ外部パラメータ行列と本質的に同等である。

次に、変換行列 M_f の推定に用いる誤差関数を定

義する。一般に、撮影時の量子化や、特徴点抽出の誤差、及びカメラ外部パラメータの推定誤差により、特徴点の第 f フレームへの投影座標 (u_{fp}, v_{fp}) と、実際の画像上での検出座標である (u'_{fp}, v'_{fp}) とは一致しない。そこで、本論文では、第 f フレームの特徴点 p に関する誤差 E_{fp} を次のように定義する。

$$E_{fp} = (u_{fp} - u'_{fp})^2 + (v_{fp} - v'_{fp})^2 \quad (4)$$

本手法では、誤差 E_{fp} の和を最小化することで、カメラ外部パラメータ M_f とモザイク画像平面上での特徴点座標 (x_p, y_p) の推定を行う。

2.2 カメラ外部パラメータの推定

図 2 に示すように、カメラ外部パラメータ推定は三つの処理によって構成される。まず、特徴点を追跡することにより、初期フレームから最終フレームまでのカメラ外部パラメータ M_f を推定する (a)。次に、全フレームにおいて、一度消失した後、再出現した特徴点を検出する (b)。最後に、カメラパラメータを全フレームにわたって最適化することにより、推定誤差を最小化する (c)。以下、各処理について説明する。

2.2.1 特徴点追跡によるカメラパラメータ推定

本処理では、特徴点を追跡することにより、各フレームにおけるカメラ外部パラメータ M_f と、特徴点のモザイク画像平面上での座標 (x_p, y_p) を推定する。なお、本処理は [9] の手法の拡張である。

はじめに、初期フレームにおける処理について述べる。本研究では、初期フレームにおける画像平面とモザイク画像平面が正対していると仮定し、 M_f を単位行列に初期化する。初期フレームの画像上で検出された特徴点のモザイク画像平面上での座標 (x_p, y_p) についても、この仮定に基づいて計算する。なお、この時点で画像平面とモザイク画像平面とが正対していない場合には、誤ったカメラパラメータが推定されることになるが、これは後段の補正処理 (c) によって補正されることになる。

次に、後続フレーム ($f > 1$) での処理について述べる．各フレームのカメラ外部パラメータ M_f は、以下のステップを最終フレームまで繰り返すことによって逐次推定される．

- 特徴点の追跡

Harris のインタレストオペレータ [11] とテンプレートマッチングにより、前フレームの各特徴点について、現フレームにおける対応点を探索する．更に、RANSAC アルゴリズム [12] を用いて、特徴点の誤対応を排除する．

- カメラ外部パラメータの推定

前のステップで決定された特徴点の画像上での追跡座標 (u'_{fp}, v'_{fp}) と、対応するモザイク画像平面での位置 (x_p, y_p) を用い、PnP 問題を解くことによって、カメラの外部パラメータ M_f を算出する．ここでは、先に定義した誤差関数の和 $\sum_p E_{fp}$ を Levenberg-Marquardt 法によって最小化することで、カメラの外部パラメータ M_f を得る．なお、モザイク画像平面での座標 (x_p, y_p) は、前のフレームまでに推定されている値を用いる．

- モザイク画像平面上的特徴点座標の推定

モザイク画像平面上的特徴点 p の座標 (x_p, y_p) を、特徴点ごとに $\sum_{i=1}^f E_{ip}$ を最小化することで算出する．

- 特徴点の追加と削除

正確なカメラパラメータを得るためには、誤追跡の発生しにくい特徴点を選択する必要がある．ここでは、複数の評価尺度を用いることで、カメラパラメータ推定に用いる信頼度の高い特徴点の組をフレームごとに自動更新する [9] ．

以上のステップを繰り返すことにより、各フレームにおけるカメラ外部パラメータ M_f とモザイク画像平面上的特徴点座標 (x_p, y_p) を逐次的に推定する．

2.2.2 再出現特徴点の検出

カメラを動かしながら紙面全体を漏れなく撮影しようとする、カメラは必然的に、図 4(a) に示すように、複数回の折り返しを伴って移動することになる．その結果、画像上の特徴点の中には、数フレームにわたって追跡されてから、一度フレームアウトした後、再度フレームインしてくるものが存在する．本研究ではこのような再出現特徴点を検出し、同一の特徴点として対応づけることで、後段のカメラパラメータの推定誤差の最小化に利用する．

再出現特徴点の検出は、時間的に離れたフレーム間

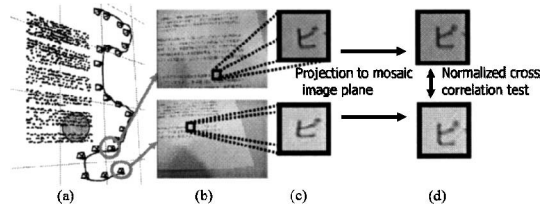


図 4 再出現特徴点の検出, (a) カメラの移動経路, モザイク画像平面上的特徴点, (b) 入力映像中の時間的に離れた 2 フレーム, (c) 各フレームにおける同一特徴点のパターン, (d) モザイク画像平面上に投影された各パターン

Fig. 4 Detection of re-appearing features.

(a) camera path, posture and feature position on mosaic image plane, (b) sampled frames of an input image sequence, (c) templates of a feature in different images, (d) templates projected to a mosaic image plane.

で、類似のパターンをもつ特徴点が存在するかを検証することによって行う．その際、図 4(c) に示すように、紙面上の同じ特徴点であっても、カメラ運動の影響による射影ひずみにより、パターンが変化してしまうことが問題となる．そこで、本手法では、各フレームにおけるパターンをモザイク画像平面上に投影することで、射影ひずみの影響を排除する．次に、モザイク画像平面上的空間的距離が一定のしきい値以下となる特徴点の組を抽出し、両者の類似度を正規化相互相関の尺度によって評価する．その結果、相関の高い特徴点の組を再出現特徴点とする．

2.2.3 カメラパラメータの補正

ここでは、バンドル調整の枠組み [13] を用いることで、カメラ外部パラメータの補正を行う．

カメラ外部パラメータ推定処理は逐次的に行うため、各フレームにおける処理は短時間でできる反面、推定誤差が蓄積するという問題がある．バンドル調整は、この推定誤差が動画像全体で最小となるように、各フレームのカメラ外部パラメータ及び特徴点座標を同時に最適化する処理である．

本研究では、バンドル調整において最小化する累積推定誤差 E を次のように定義する．

$$E = \sum_f \sum_p E_{fp} \quad (5)$$

これは、特徴点座標をカメラ画像平面上に投影したものと、実際のカメラ画像上での特徴点座標との二乗誤差の総和である． E は、逐次処理において推定誤差が蓄積するに従って大きな値をとる．また、初期フ

レームではカメラは紙面に対して正対していると仮定したが、実際にはカメラが紙面に対して傾いていた場合には、初期フレームにおいて誤ったカメラ外部パラメータが推定されるため、結果的に E は大きな値をとることになる。そこで、 E をカメラパラメータ M_f とモザイク画像平面上の特徴点座標 (x_p, y_p) について最小化する。これにより、推定誤差は解消され、初期フレームにおいてカメラが紙面に対して傾いていた場合には、正しいカメラ外部パラメータが推定されることになる。

なお、再出現特徴点については、最初の追跡結果と再出現後の追跡結果とを同一特徴点系列とみなし、誤差 E_{fp} を計算する。これにより、累積推定誤差 E を最小化する過程で、図 4 (b) に示すような時間的に離れたフレーム間でも整合性を保ちながら、カメラ外部パラメータが最適化されることになる。その結果、再出現前後の特徴点系列を統合せずに E を最小化する従来のバンドル調整に比べて、より高精度のカメラ外部パラメータを得ることができる。

2.3 超解像モザイク画像の生成

最後に、最適化されたカメラパラメータを用い、超解像処理を適用することで、超解像モザイク画像を生成する。超解像とは、サブピクセルレベルでの重なりをもつ複数の低解像度画像を統合することで、対象の高解像度画像を推定する手法である。ここでは、Iraniらによって提案された Iterative Back Projection アルゴリズム [14] を利用する。処理の流れは、次のとおりである。

今、 m 枚の入力画像 $\{I_f\} = I_1, I_2, \dots, I_m$ から、入力画像に対して n 倍の解像度をもつモザイク画像 S を生成するとする。はじめに、式 (1) 及びカメラ外部パラメータ M_f を用いて、すべての入力画像をモザイク画像平面上に投影し、入力画像に対して n 倍の細かさをもつグリッドでリサンプリングした画像を生成する。ここで、投影画像間での重なり部分に対しては、ブレンディング処理を適用する。以上の処理によって生成された画像を、モザイク画像の初期画像 $S^{(0)}$ とする。次に、初期モザイク画像 $S^{(0)}$ に対して、幾何変換と、点広がり関数 (Point Spread Function: PSF) によるぼかし処理とからなる画像撮像過程をシミュレートし、入力画像 $\{I_f\}$ に対応する低解像度画像 $\{I_f^{(0)}\}$ を生成する。ここで、 $S^{(0)}$ が正しく推定されていれば、すべてのフレームにおいて $\{I_f\}$ と $\{I_f^{(0)}\}$ は一致することになる。逆に、 $S^{(0)}$ の推定誤差が大きくなれば、

同様に $\{I_f\}$ と $\{I_f^{(0)}\}$ との差分も大きくなる。そこで、入力画像と撮像過程のシミュレーションにより生成された低解像度画像との誤差画像 $\{I_f - I_f^{(0)}\}$ を計算し、モザイク画像中の対応する個所に逆投影 (Back Projection) し、加算することで、更新されたモザイク画像 $S^{(1)}$ を得る。この低解像度画像の生成と誤差画像の逆投影を、モザイク画像が収束するまで繰り返すことにより、超解像モザイク画像 S を生成する。

なお、入力画像の枚数が増えれば増えるほど、高い倍率の超解像モザイク画像が得られることが予想されるが、倍率には限界があることが経験的に知られている。本研究では、実験的にモザイク画像の倍率を $n = 2.0$ に決定した。また、PSF については、Capelらがカメラの PSF がガウス関数によって近似できることを実験的に示している [15]。そこで、本研究では、PSF として $\sigma = 0.7$ のガウス関数を採用した。

3. 実 験

提案手法に基づくビデオモザイクシステムを試作した。本システムは、デスクトップ PC (CPU: Pentium-4 3.2 GHz, メモリ: 2 GByte) と IEEE1394 接続の CCD カメラ (Aplux C104T) とから構成される。カメラの内部パラメータは、あらかじめ Tsai の手法 [16] によって校正済みであり、これらのパラメータは固定とした。

本システムの有効性を検証するために、実画像 (文章、写真の 2 種類) とシミュレーションデータそれぞれに対して実験を行った。

3.1 実画像を使ったモザイク実験

試作システムを使い、文章と写真の 2 種類の紙面に対して実験を行った。いずれの対象も A4 サイズであり、定量評価 (後述) のために、紙面上に + 印のグリッドを 40 mm 間隔で配置したものをを用いた。

3.1.1 文章を対象としたモザイク実験

本実験における入力映像及び特徴点の検出位置をそれぞれ図 5 に示す。これは手持ちのカメラによって 15 fps で撮影された 150 フレームからなる映像であり、画像サイズは 640×480 ピクセルである。図中の \times 印は、検出された特徴点の画像上の位置を表す。なお、本実験では、手法の有効性を検証するために、入力画像のすべてのフレームにおいて、カメラを紙面に対して傾けた状態で撮影を行った。

本手法によって得られたカメラ外部パラメータ、及び各特徴点のモザイク画像平面上での位置を図 6 に示

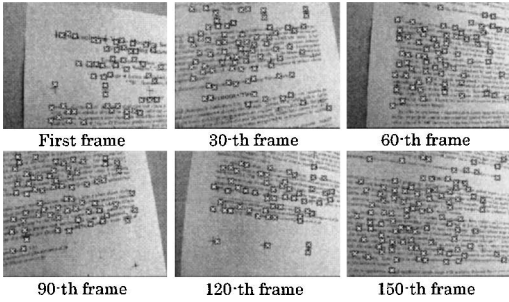


図 5 入力映像 (文章)

Fig. 5 Sampled frames of input image sequences (document).

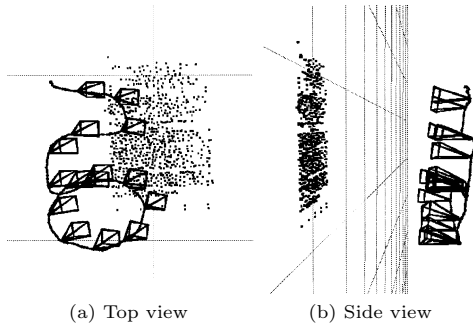


図 6 カメラ外部パラメータと特徴点座標の推定結果 (文章)

Fig. 6 Estimated extrinsic camera parameters and feature positions (document).

す。図中の曲線は、推定されたカメラの移動経路を表し、錐台は 10 フレームおきのカメラの姿勢を表す。

推定されたカメラ外部パラメータを用い、3 回の逆投影処理によって生成された超解像モザイク画像を図 7 に示す。超解像モザイク画像のサイズは、 2452×3002 ピクセルであった。図 7 より、いずれの入力画像も紙面に対して正対していなかったにもかかわらず、生成された超解像モザイク画像に射影ひずみが生じていないのが確認できる。また、入力画像の一部と、超解像モザイク画像中の対応箇所を拡大したものを図 8 に示す。入力画像では接触していた文字同士が、超解像モザイク画像では分離されており、解像度が向上しているのが確認できる。

なお、本実験における処理時間は、以下のとおりである。

- [画像入力及びカメラ外部パラメータの初期推定]
- 1 フレーム当り 67 ミリ秒 (15 fps の実時間処理)
- [再出現特徴点の検出] 1 秒 (150 フレーム全体で)
- [カメラパラメータの補正] 20 秒

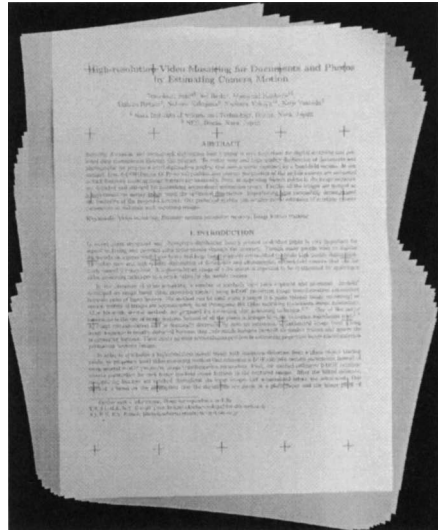


図 7 生成された超解像モザイク画像 (文章)

Fig. 7 Generated super-resolved mosaic image (document).

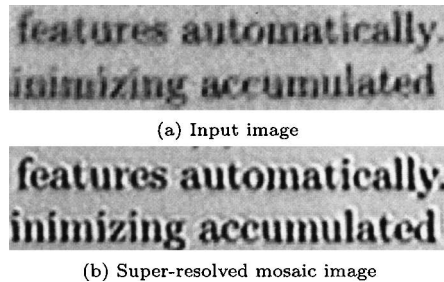


図 8 入力と超解像モザイク画像の比較 (文章)

Fig. 8 Comparison between input image and super-resolved mosaic image (document).

[モザイク画像の生成 (ブレンディング)] 32 秒

[モザイク画像の生成 (超解像反復 3 回)] 122 秒

3.1.2 写真を対象としたモザイク実験

本実験における入力映像及び特徴点の検出位置をそれぞれ図 9 に示す。文章に対する実験同様、手持ちのカメラによって 15 fps で撮影された 150 フレームからなる映像であり、画像サイズは 640×480 ピクセルである。ここでは、初期フレームにおいてカメラと紙面がおおむね正対するように撮影を行っている。

本手法によって得られたカメラ外部パラメータ、及び各特徴点のモザイク画像平面上での位置を図 10 に示す。また、推定されたカメラ外部パラメータを用い、3 回の逆投影処理によって生成された超解像モザイク画像を図 11 に示す。超解像モザイク画像のサイズは、

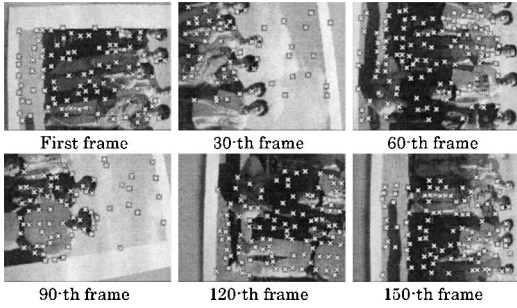


図 9 入力映像 (写真)

Fig. 9 Sampled frames of input image sequences (photograph).

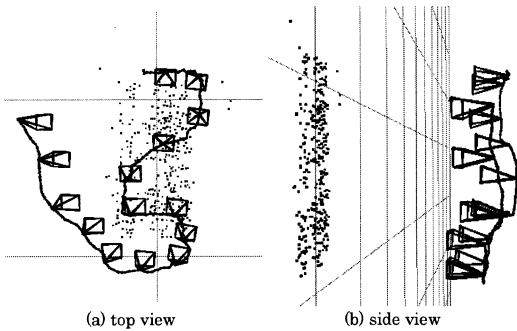


図 10 カメラ外部パラメータと特徴点座標の推定結果 (写真)

Fig. 10 Estimated extrinsic camera parameters and feature positions (photograph).

2169 × 2719 ピクセルであった。入力画像の一部と、超解像モザイク画像中の対応箇所を拡大したものを図 12 に示す。入力画像では見えにくかった眼鏡のフレームや、シャツのストライプが、超解像モザイク画像では復元されているのが確認できる。

なお、本実験における処理時間は、以下のとおりである。

- [画像入力及びカメラ外部パラメータの初期推定] 1 フレーム当り 67 ミリ秒 (15 fps の実時間処理)
- [再出現特徴点の検出] 1 秒 (150 フレーム全体で)
- [カメラパラメータの補正] 9 秒
- [モザイク画像の生成 (ブレンディング)] 31 秒
- [モザイク画像の生成 (超解像反復 3 回)] 105 秒

文章を対象とした実験より、カメラパラメータの補正処理時間が短縮されている。これは、初期フレームの撮影時にカメラと紙面を正対させたことにより、カメラパラメータの初期推定の精度が良くなり、最適化処理の収束が早くなったためと考えられる。

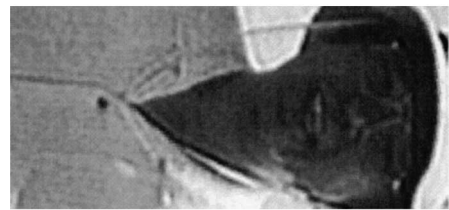


図 11 生成された超解像モザイク画像 (写真)

Fig. 11 Generated super-resolved mosaic image (photograph).



(a) Input image



(b) Super-resolved mosaic image

図 12 入力と超解像モザイク画像の比較 (写真)

Fig. 12 Comparison between input image and super-resolved mosaic image (photograph).

3.1.3 モザイク画像上のひずみの定量的評価

続いて、モザイク画像に生じたひずみの定量評価を行った。先の実験に用いた文章、写真上には、あらかじめ 40 mm 間隔のグリッド上に + 印を印刷しておいた。ここでは、モザイク画像上での + 印の座標を手作業で計測し、隣接するグリッド間の距離を算出することで、ひずみの評価を行った。

各実験におけるグリッド間距離 (ピクセル) の平均、

表 1 モザイク画像上でのグリッド間距離 [単位はピクセル, 括弧内は平均からのパーセンテージ]

Table 1 Distances of adjacent grid points in the mosaic image.

Target	Average	Maximum	Minimum	Standard Deviation
Doc.	351.0(100.0)	357.9(101.9)	346.4(98.6)	2.53(0.72)
Photo	332.9(100.0)	337.6(101.4)	329.4(98.9)	1.72(0.52)

最大値, 最小値及び標準偏差を表 1 に示す. また, 平均値に対する最大値, 最小値, 標準偏差のパーセンテージを括弧内に示す. この表において, 標準偏差はひずみ量の平均を表しているが, いずれの実験においても, これらの値は十分小さく, 画像のデジタル保存やインターネット伝送といった用途には問題のないレベルであるといえる.

3.2 シミュレーションによるカメラパラメータ推定精度の評価

最後に, シミュレーションデータを用いた実験により, カメラパラメータの推定精度を定量的に評価した. はじめに, 先の実験で用いたのと同じ解像度, 視野角をもつ仮想カメラを用いて, 仮想空間内のモザイク画像平面を撮影した. 次に, 提案手法を用いてカメラパラメータを推定し, 真値と比較することで精度評価を行った. 以上の実験を複数の移動速度について行い, 提案手法におけるカメラ移動速度の限界値を求めた.

以下, 本実験の詳細について述べる. 本実験では, カメラとモザイク画像平面との距離を 200 mm に固定し, カメラをモザイク画像平面と平行に一定速度で移動させた. モザイク画像平面上には多数の特徴点を配置し, これらの特徴点を各フレームにおける仮想カメラの画像上に投影したものを, 特徴点の入力画像上の座標とした. ここで, モザイク画像平面上の特徴点の位置はランダムとしたが, 個数については各フレームで平均 90 個の特徴点が検出されるように調整した. また, 特徴点の入力画像上の座標については, 特徴点検出の誤差を考慮して, 標準偏差 0.8 のガウスノイズを加え, 更に画素単位での量子化を行った. 以上の実験条件は, 本システムを用いたモザイクングにおける一般的な状況を想定し, 設定したものである.

このような条件のもと, 複数のカメラの移動速度について実験を行い, カメラパラメータの推定精度を評価した. 図 13 に得られたカメラ外部パラメータの推定精度を示す. 同図から, 移動速度が速まるにつれ, 推定精度が悪化していることが分かる. 特に, カメラ

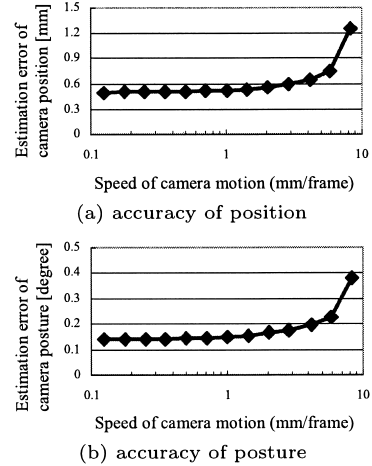


図 13 カメラ位置・姿勢の推定精度の評価 (シミュレーション)

Fig. 13 Evaluation of camera parameter estimation (simulation).

の移動速度が 4 mm/フレームを超えると, 推定精度が著しく悪化している. 本実験より, 提案手法におけるカメラの移動速度の限界は, 4 mm/フレームであるといえる. なお, この移動速度において, 各特徴点は画像の右端にフレームインしてから, 左端にフレームアウトするまでの間, 平均 10 フレームにわたって追跡された. すなわち, 特徴点の画像上での移動速度は 64 ピクセル/フレームであり, これは連続するフレーム間で 90%の重なりを保ちながら, カメラを移動させたことに相当する.

4. む す び

カメラパラメータ推定による超解像ビデオモザイクング手法を提案した. 本手法により, 初期フレームにおいてカメラが紙面に対して傾いている場合においても, 正対した高解像度のモザイク画像を生成することができる. プロトタイプシステムを開発し, シミュレーションと実際の紙面を使った実験においてその有効性を確認した. 今後は, 処理の高速化, 軽量化を図ることにより, 実際のカメラ付き携帯機器を使った超解像モザイクングシステムの開発を目指す.

文 献

- [1] R. Szeliski, "Image mosaicing for tele-reality applications," Proc. IEEE Workshop on Applications of Computer Vision, pp.230-236, 1994.
- [2] 千葉直樹, 蚊野 浩, 美濃導彦, 安田昌司, "画像特徴に基づくイメージモザイクング," 信学論 (D-II), vol.J82-D-II,

no.10, pp.1581-1589, Oct. 1999.

- [3] S. Takeuchi, D. Shibuichi, N. Terashima, and H. Tominaga, "Adaptive resolution image acquisition using image mosaicing technique from video sequence," Proc. IEEE Int. Conf. on Image Processing, vol.I, pp.220-223, 2000.
- [4] C.T. Hsu, T.H. Cheng, R.A. Beuker, and J.K. Hong, "Feature-based video mosaicing," Proc. IEEE Int. Conf. on Image Processing, vol.II, pp.887-890, 2000.
- [5] M. Lhuillier, L. Quan, H. Shum, and H.T. Tsui, "Relief mosaics by joint view triangulation," Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, vol.I, pp.785-790, 2001.
- [6] P.F. McLauchlan and A. Jaenicke, "Image mosaicing using sequential bundle adjustment," Image and Vision Computing, vol.20, pp.751-759, 2002.
- [7] D.W. Kim and K.S. Hong, "Fast global registration for image mosaicing," Proc. IEEE Int. Conf. on Image Processing, vol.II, pp.295-298, 2003.
- [8] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," Int. J. Comput. Vis., vol.9, no.2, pp.137-154, 1992.
- [9] 佐藤智和, 神原誠之, 横矢直和, 竹村治雄, "マーカと自然特徴点の追跡による動画像からのカメラ移動パラメータの復元," 信学論 (D-II), vol.J86-D-II, no.10, pp.1431-1440, Oct. 2003.
- [10] H. Kawasaki, A. Miyamoto, Y. Ohsawa, S. Ono, and K. Ikeuchi, "Multiple video camera calibration using EPI for city modeling," Proc. Asian Conf. on Computer Vision, vol.1, pp.569-574, 2004.
- [11] C. Harris and M. Stephens, "A combined corner and edge detector," Proc. Alvey Vision Conf., pp.147-151, 1988.
- [12] M.A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," Commun. ACM, vol.24, no.6, pp.381-395, 1981.
- [13] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, "Bundle adjustment — A modern synthesis," in Vision Algorithms: Theory and Practice, pp.298-375, 2000.
- [14] M. Irani and S. Peleg, "Improving resolution by image registration," CVGIP: Graphical Models and Image Processing, vol.53, no.3, pp.231-239, 1991.
- [15] D. Capel and A. Zisserman, "Super-resolution enhancement of text image sequences," Proc. Int. Conf. on Pattern Recognition, vol.1, pp.600-605, 2000.
- [16] 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.

(平成 16 年 10 月 12 日受付, 17 年 2 月 10 日再受付)



池谷 彰彦 (正員)

1997 阪大・工・電子制御機械卒。1999 同大大学院博士前期課程了。同年 NEC 入社。以来, コンピュータビジョンの研究に従事。2004 本会学術奨励賞, PRMU 研究奨励賞, FIT ヤングリサーチャー賞受賞。



佐藤 智和 (正員)

1999 阪大・工・情報工卒。2003 奈良先端科学技術大学院大学情報科学研究科博士後期課程了。現在, 同大情報科学研究科助手。コンピュータビジョンの研究に従事。2001 本会学術奨励賞受賞。博士(工)。情報処理学会, IEEE 各会員。



池田 聖 (学生員)

2001 広島大・理・物理卒。2003 奈良先端科学技術大学院大学情報科学研究科博士前期課程了。現在, 同大博士後期課程に在学中。IEEE 会員。



神原 誠之 (正員)

1997 岡山大・工・情報工卒。2002 奈良先端科学技術大学院大学情報科学研究科博士後期課程了。現在, 同大情報科学研究科助手。コンピュータビジョン, 拡張現実感の研究に従事。1999 本会学術奨励賞受賞。博士(工)。情報処理学会, 日本バーチャルリアリティ学会, IEEE 各会員。



中島 昇 (正員)

1990 慶大・理工・電気卒。1992 同大大学院修士課程了。同年 NEC 入社。以来, 文書画像処理の研究に従事。1996 本会学術奨励賞受賞。



横矢 直和 (正員)

1974 阪大・基礎工・情報卒。1979 同大大学院博士後期課程了。同年電子技術総合技術研究所入所。以来, 画像処理ソフトウェア, 画像データベース, コンピュータビジョンの研究に従事。1986-1987 マツギル大・知能機械研究センター客員教授。1992 奈良先端科学技術大学院大学・情報科学センター教授。現在, 同大情報科学研究科教授。1990 情報処理学会論文賞受賞。工博。情報処理学会, 人工知能学会, 日本認知科学会, 映像情報メディア学会, IEEE 各会員。