

形状変形による古文書画像のシームレス合成

Use of Shape Deformation to Seamlessly Stitch Historical Document Images

● Wei Liu

● Wei Fan

● Li Chen

● Sun Jun

● 直井 聡

あらまし

中国では、古文書の保護と有効活用に向けて、デジタルデータとして保存することが進められている。大判の古文書に対しては、小型の非接触型スキャナにより複数の領域に分けてスキャンし、それらの画像をきれいにつなぎ合わせて画像合成するステッチング技術が期待されている。しかし古文書は紙質が悪く表面が平坦ではないため、スキャンした画像に局所的な歪みが発生し、画像合成が困難という課題があった。富士通研究開発中心有限公司は、この問題を解決するために形状変形を用いてシームレスに画像合成する技術を開発した。まず、一方の画像を変形してもう一方の画像に近似した後、最小限の変形で整合性を最大化する1本もしくは2本の最適な継ぎ目を推定する。この方法を用いれば、高品質な古文書画像が得られる。

本稿では、古文書画像の合成処理特有の課題を示し、それを解決するための技術とその性能評価について述べる。

Abstract

In China, efforts are being made to preserve historical documents in the form of digital data so that they can be effectively used while being protected. Stitching technology is expected to play a role in these efforts by dividing a large-sized historical document into multiple areas, scanning those areas using a compact contactless scanner, and cleanly connecting the resulting images. However, historical documents are often on paper of low quality and have an uneven surface. The resulting local distortion in the scanned images makes it difficult to stitch them together. To solve this problem, Fujitsu Research and Development Center Co., Ltd. has developed a seamless image stitching method using shape deformation. This method first deforms one of two images to approximate the other image and then estimates single or dual optimal seams to minimize deformation and maximize consistency. With this method, high-quality images of historical documents can be obtained. This paper describes the problem associated with the stitching of historical document images, proposes a method for solving it, and evaluates its performance.

ま え が き

中国では、古文書の保護と有効活用に向けて、デジタルデータとして保存することが進められている。大判の古い新聞紙や書道作品などのデジタルアーカイブ化では、大型の専用スキャン装置が使われているが、非常にコストがかかる。そこで、安価な小型のスキャナで分割して撮影し、画像合成技術により1枚の画像として保存する方法が注目されている。富士通研究開発中心有限公司は、小型の非接触型スキャナを使って、大判の古文書を複数の領域に分けてスキャンし、それらの画像をシームレスに合成する技術を開発した。この技術は、オーバーラップした複数の画像から1枚のシームレスで高解像度な画像を合成できることが特徴である。パノラマ写真、道路景観などの写真のほか、大判古文書のデジタルアーカイブ化などに利用されている。

古文書画像の合成に必要な要件は、従来の画像合成に対するものよりも厳しい。その理由には、画像構造および画像の位置ずれが目立つということがある。文書の表面が平坦ではないことで合成画像の内容が違ったものになってしまうことと、古文書の紙の品質が低いために画像に歪みが発生するという2点がその主な理由である。したがって、直線などの画像構造が形状的な連続性を保つように画像を合成することが課題である。

従来の技術とその課題

従来の合成法は、

- (1) 最適な継ぎ目を1本用いる方法
- (2) 2枚の画像の中に対応する継ぎ目の位置を特定して、その継ぎ目に沿って画像の内容の変形と合成を行う方法

の二つに分類される。最初の分類に含まれる方法の一般的な手順は、2枚の画像の間の一括変換結果⁽¹⁾を推定して大まかに位置合せした後に、最適な継ぎ目に沿ってこの二つの部分を合成して画像を作成するというものである。一例として、Kwatraらが提案した方法⁽²⁾では、オーバーラップした領域の輝度やテクスチャーの差異を最小化することにより、最適な継ぎ目の位置を特定する。Veenaが提案した方法⁽³⁾では、継ぎ目に沿った虚像を最小

化することで、2枚の画像間の光度の不整合と形状的な位置ずれの双方を抑制する。

しかし、非剛体の画像合成においては、概して局所的な歪みがいくつかに生じることから、継ぎ目を1本用いただけでは画像構造（文字や図など）の位置ずれが起こる可能性がある。この問題を解決するため、合成対象の2枚の画像の中に対応する継ぎ目を2本探し出す方法が開発されてきた。例えば、Jiaらは、輝度の一貫性と構造の連続性を基に2本の対応する継ぎ目を推定する方法⁽⁴⁾を開発した。この継ぎ目は、それぞれの画像の中で変形を伝搬させて完全な画像を形成するために用いられる。Fangらは、オーバーラップした領域の中の変形歪みを最小限に抑えながら、色のマッチングを最大化する変形機能⁽⁵⁾を開発した。Fanらは、動的プログラミング⁽⁶⁾を用い、前面にある文字への交差が最小限となるような最適な継ぎ目を探し出し、特徴マッチングを用いて二番目の継ぎ目を推定する方法⁽⁷⁾を開発した。これらの方法を用いると、継ぎ目の反対側にある内容は変形後に合成されることから、画像の内容は良好に保持される。しかし、直線など形状的に連続な要素は、変換後には直線ではなくなることもあり得る。

提案方法の概要

筆者らは、前章で述べた問題を克服するために、形状変形を用いるシームレス画像合成法を開発した。最初に、画像変形を最小限に抑えた上で、最適な継ぎ目を1本用いて2枚の画像を合成する。次に、図形的な不整合の問題を克服するために、変形による整合性がもっとも高い継ぎ目を2本用いる。この方法によれば、それぞれの継ぎ目に沿った画像内容の変形には整合性があり、直線など形状的に連続な要素は良好に保持される。

シームレス画像合成法は図-1に示す手順で行う。

- (1) 2台のカメラを用いて、文書の上半分と下半分の画像を取り込む。
- (2) 大まかな位置合せにより、オーバーラップ領域を求める。
- (3) 安定した制御点を抽出する。
- (4) 1枚の画像を変形し、画像変形を用いて最適な継ぎ目を推定する。
- (5) 最適な継ぎ目を用いて画像を合成する。

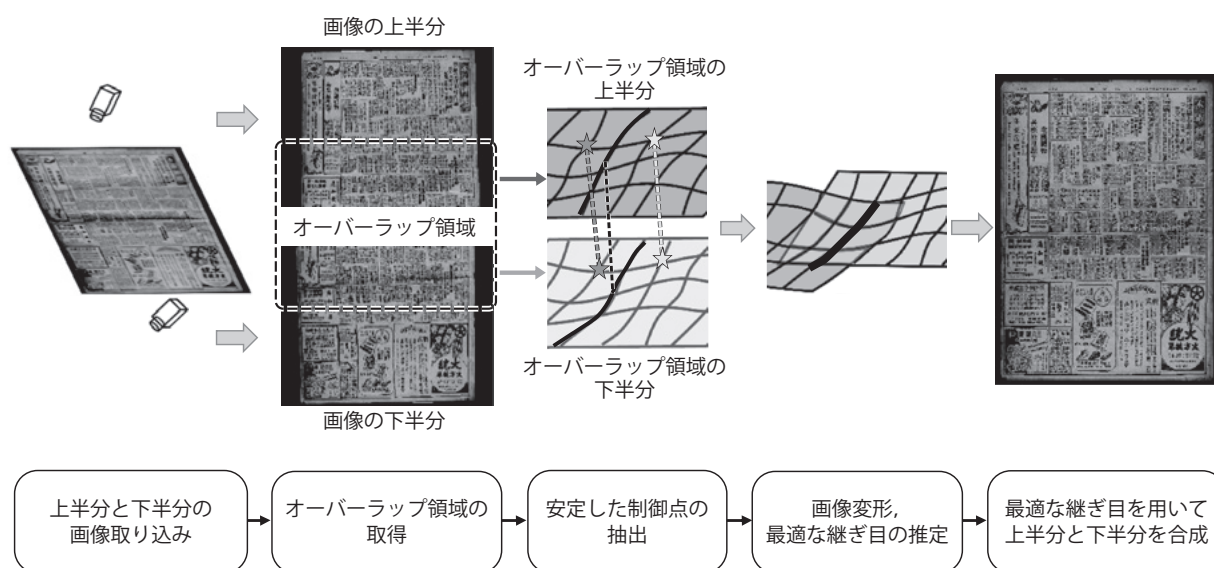


図-1 提案手法のフローチャート

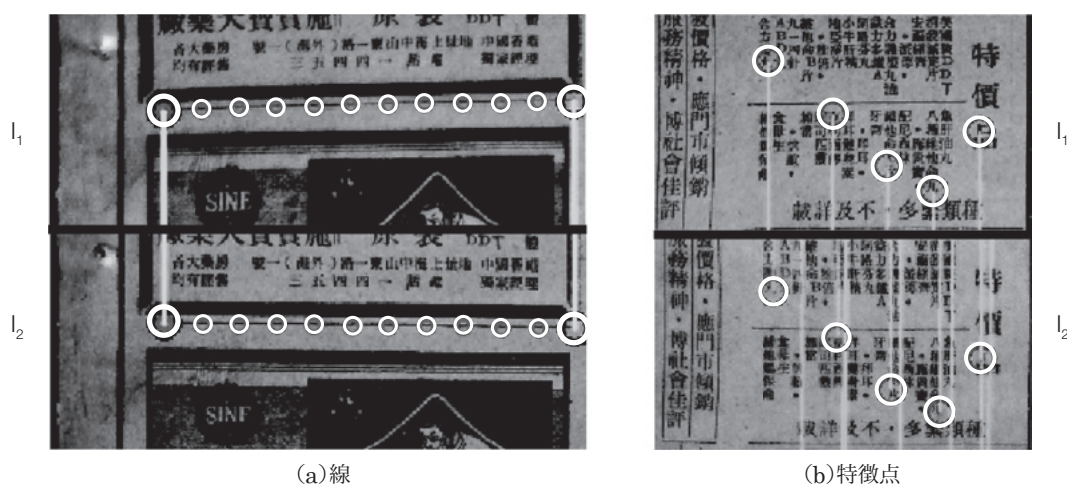


図-2 抽出された制御点の例

次章以降で、提案方法のポイントである画像変形と最適な継ぎ目の推定について述べる。

画像変形

● 安定した制御点の抽出

文書の表面が平坦ではないことから2枚の画像の間に差異が生じる可能性があるものの、それらの画像には安定した特徴が存在する。古文書であっても狭い範囲内では文書の表面は局所的に連続しているため、2枚の画像の局所的な領域の間には対応関係が存在する。対応する点を抽出すると、それらの点は安定した制御点の集合の一部を成す。

古文書の画像の中には、直線などの幾何学的要素が含まれている場合がある。要素間の相関性は画像変形により変化しないことから、これらの要素は安定した特徴となる。2枚の画像の中から対応する線をペアで抽出すると、その線上の点が安定した制御点の集合のほかの部分を成す。抽出された制御点の二つの例を図-2に示す。

● 移動最小二乗法による画像変形

2枚の画像が大まかに位置合せされており、オーバーラップ領域が取得されているものとする。また、局所的な歪みのため、この場合には2枚の画像間の一括変換処理は適していないものとする。こ

のため、変形場を用いて局所的な変形を表現する。変形場のモデル化には、移動最小二乗法 (MSL)⁽⁸⁾ が用いられる。MSLを用いる目的は、アフィン空間を近似し、制御点の近似誤差を最小化するためである。

2枚の画像の対応する制御点の位置を、 p_i および q_i とする。画像内にある任意のピクセルの座標が v であるとする、MLSによって次式の値を最小とする最適なアフィン変換^(註) $I_v(x)$ が導出される。

$$\sum_i w_i |I_v(p_i) - q_i|^2$$

ここで、 w_i は現在のピクセルを基準とする制御点の重みを示す。

M として表記されるアフィン空間⁽⁹⁾ の解は、上記の式を直接微分することで求められる。アフィン空間を求めた後には、それぞれのピクセルを変換することにより画像の変形が行われる。

形状変形による継ぎ目の推定

シームレス画像合成法を実現するために、1本の最適な継ぎ目、あるいは2本のマッチングできる継ぎ目は、両側の画像の位置合わせする必要がある。本章では、変形量を最小とする単一継ぎ目法 (SSMD: Single Seam with Minimum Deformation) と、変形整合性を最大とする二重継ぎ目法 (DSMDC: Dual Seam with Maximum Deformation Consistency) の二つの方法を紹介する。双方とも、筆者らが提案するシームレス画像合成法に活用されている方法である。

(注) 線型変換と平行移動を組み合わせた図形や形状の変形方式。元の図形で直線上に並ぶ点に変換後も直線上に並び、平行線は変換後も平行線であるなど画像の幾何学的性質を保ったままの変換結果が得られる。

● SSMD

上述のように、図-2 (a) に対するアフィン空間を求める。シームレスに合成するには、継ぎ目付近の位置ずれを最小にする必要がある。従来の方法では、最適な継ぎ目を探し出すためにグラフカット法⁽²⁾ を用いていた。この方法では、階調度の平滑性、またはテクスチャーの類似性として定義される損失を隣接ピクセル間で最大化してしまう。しかし、階調度の平滑性およびテクスチャーの類似性は、明確に定量化できる基準ではない。このため、形状変形による最適な継ぎ目を推定するという別の方法を採用した。

図-2に示した I_1 と I_2 は合成しようとする二つの画像である。 I_1 と I_2 が垂直に重なり合っており、第1の画素座標が v であり、変形画像 V は以下の式で定義されるものとする。

$$V = \|vM - v\|$$

ここで、 M はアフィン空間である。

変形量を最小とする最適配置は、次の式から計算される。

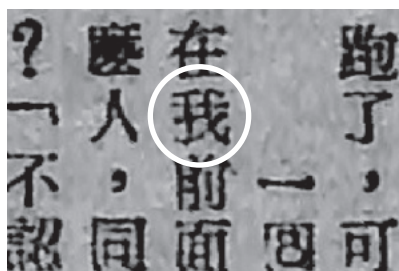
$$i^* = \arg \min_{1 \leq i \leq h} \sum_{j=1}^w V(i, j)$$

ここで、 w と h は I_1 の幅と高さである。おおよその継ぎ目は、最適配置上の点群とみなされる。

動的プログラミング⁽⁶⁾ を用いると、最適な継ぎ目は最適配置のそばを横切る経路となる。

● DSMDC

筆者らの実験において、SSMD法は従来の動的プログラミング法よりも良好な結果を得ていたが、実際には存在しない線が生じる場合 (ゴーストアーチファクト) もあった。そうしたゴーストアーチファクトの例を図-3に示す。継ぎ目の通っている



(a) 元画像



(b) 画像合成後

図-3 合成によって生じるゴーストアーチファクトの例

文字を丸印で示している。合成画像には、ゴーストアーチファクトの特徴が明らかに表れている。

筆者らは、この問題を克服するために二重継ぎ目法を開発した。マッチング可能な2本の継ぎ目を推定して結合し、2本の継ぎ目の反対側の画像を変形して最終結果を生成する。二重継ぎ目法を用いた画像合成の例を図-4に示す。右側の拡大写真が示すように、ゴーストアーチファクト問題は解消された。ただし、左側の写真のように内容は良好に保持されているものの、歪みのため直線は真っ直ぐではなくなっている。

この問題を克服するため、筆者らは二重継ぎ目法を改良したDSMDC法を開発した。最適化の目的は変形を最小化するためだけではなく、継ぎ目に沿った部分の変形の整合性を最大化するためでもある。

最初に v を片方の画像の画素座標とし、変形画像 D は次のベクトルに相当するそれぞれの画素から構築されるものとする。

$$D = vM - v$$

それに沿った変形ベクトルがほぼ同じとなる最適な継ぎ目を探し出す必要がある。最適な継ぎ目の座標 (i, j) を得るために用いた目的関数を以下に示す。

$$\min \sum_{j=1}^w \|D(i, j)\| + \lambda_1 \|X_{i,j+1} - X_{i,j}\| + \lambda_2 \|y_{j+1} - y_j\|$$

ここで、 λ_1 と λ_2 は正規化係数、 $X_{i,j} = (i, j) + D(i, j)$ は変形後の座標、 y_j は $X_{i,j}$ の y 座標とする。次の三つの目的に対応している。

- (1) 形状変形は最小化する。
- (2) 隣接ピクセル間の変形度の差を最小化する。
これは、隣接ピクセル間の変形の整合性を最大化することに等しい。
- (3) 最終的な継ぎ目の滑らかさを最大化する。

この目的関数が最小化されると、隣接ピクセル

の変換は類似のものとなる傾向がある。したがって、形状的に連続な要素の整合性は良好に保持される。目的関数の最適化には、ビタビアルゴリズム⁽⁹⁾が用いられている。

性能評価

新聞(10枚)、図面(10枚)、地図(1枚)の3種類の古文書を使用して性能を評価した。判断基準は内容の保持度とした。従来の動的プログラミング法、SSMD法、およびDSMDC法の3種類の方法を評価した。

最初に、1本の継ぎ目を用いた画像合成を評価した(図-5)。グラフカット法による最適な継ぎ目では内容の欠落が生じることがある。例えば、図-5(b)では黒線が消えてしまっている。一方、SSMD法を用いた場合は、良好な合成結果を取得できる(図-5(c))。

上記のように、従来の方法では通常、最小変形(図-4)または最小内容交差(画像の交差を最小にする方法){図-6(b)}などの基準を用いて1本の継ぎ目を推定する。図-6(c)に示すように、筆者らが提案するDSMDC法による画像合成が図-6(a)に示す元画像に最も近い結果を得た。

合成画像と元画像の差分(画像の減算により生成)を図-7に示す。白い点が少ないほど、元画像との差分が小さいことを表す。筆者らが提案するDSMDC法を用いて生成した変換画像が元画像に最も近い。再現率を表-1に示す。この比率は、図-7の対応する画像の中で差分があったピクセルの個数を合計し、合計を総ピクセル数で割り、その結果を1から減じることによって算出した。これらの結果から、筆者らが提案するSSMD法、およびDSMDC法を用いて生成した変換画像の方が元画像に近いことが分かる。

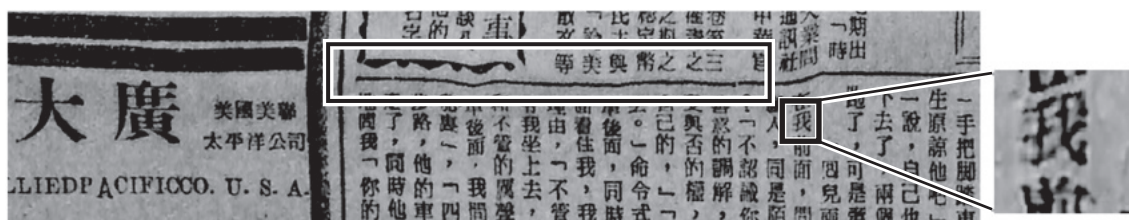
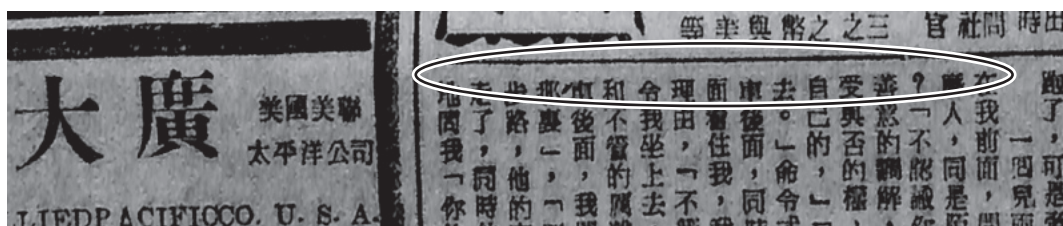
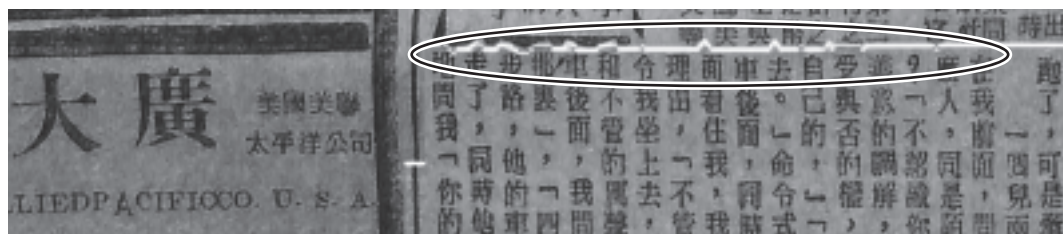


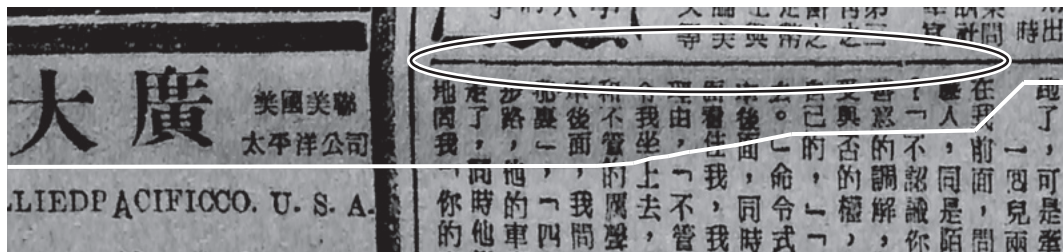
図-4 2本の継ぎ目を用いた画像合成



(a)元画像

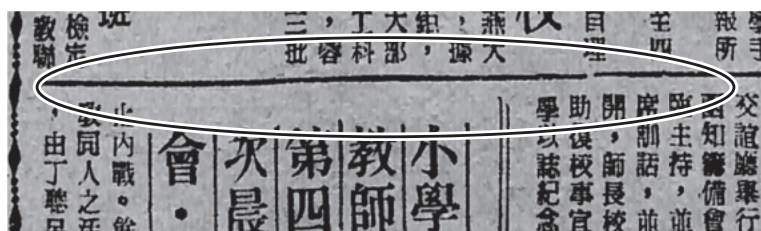


(b)グラフカット法による合成結果

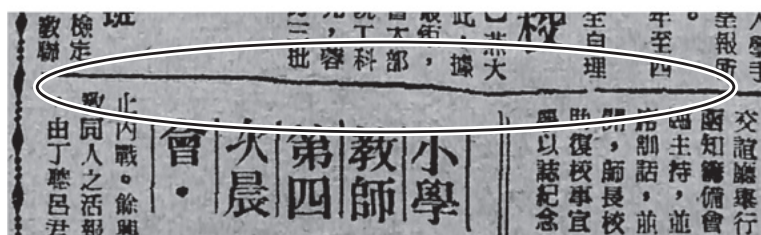


(c)SSMD法による合成結果

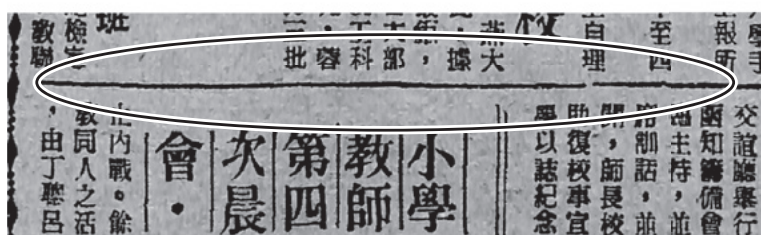
図-5 1本の継ぎ目を用いた画像合成の評価



(a)元画像



(b)最小内容交差による合成結果



(c)DSMDC法による合成結果

図-6 2本の継ぎ目を用いた画像合成の評価

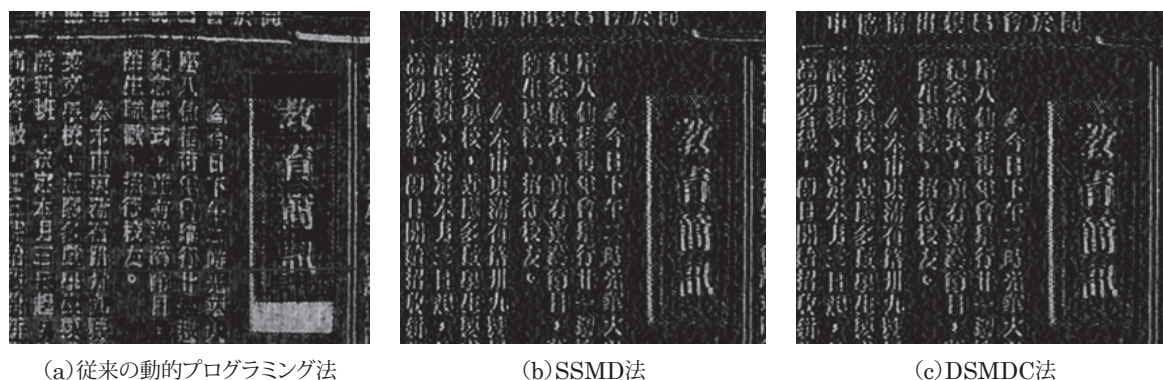


図-7 合成画像と元画像の差分

表-1 再現率

	動的 プログラミング	SSMD	DSMDC
新聞	78.58%	80.17%	80.45%
図面	82.08%	83.33%	87.89%
地図	80.86%	81.10%	80.90%

む す び

本稿で提案したシームレス画像合成法は形状変形を用いて、直線が真っ直ぐではなくなるといった形状的な不整合の問題を克服する。この方法では、合成する画像の一方を変形してもう一方の画像に近似した後、最小限の変形で整合性を最大化する1本もしくは2本の継ぎ目を推定する。この方法を用いれば、継ぎ目に沿った部分の画像内容の変形度は一定となり、直線など形状的に連続する要素は良好に保持される。富士通研究開発中心有限公司は、今後もシームレス画像合成法の開発を継続し、中国の古文書の保護と有効活用のために貢献していきたい。

参考文献

- (1) L. G. B. Mirisola et al. : Exploiting inertial sensing in mosaicing and visual navigation. Intelligent Autonomous Vehicles, Vol.6, Part 1, p.306-311 (2007).
- (2) V. Kwatra et al. : Graphcut textures : image and

- video synthesis using graph cuts. ACM Transactions on Graphics, Vol.22, No.3, p.277-286 (2003).
- (3) C. V. Veena : Minimizing seam artifacts in image stitching. Asian Journal of Information Technology, Vol.6, No.2, p.209-214 (2007).
- (4) J. Jia et al. : Image stitching using structure deformation. Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol.30, No.4, p.617-631 (2008).
- (5) H. Fang et al. : Textureshop : texture synthesis as a photograph editing tool. ACM Transactions on Graphics, Vol.23, No.3, p.354-359 (2004).
- (6) D. P. Bertsekas : Dynamic programming and optimal control, volume 1. Athena Scientific, Belmont, MA, 1995.
- (7) W. Fan et al. : Paper stitching using maximum tolerant seam under local distortions. The 2014 ACM symposium on Document engineering, p.35-44 (2014).
- (8) D. Levin : The approximation power of moving least-squares. Mathematics of Computation of the American Mathematical Society, Vol.67, No.224, p.1517-1531 (1998).
- (9) S. Scott et al. : Image deformation using moving least squares. ACM Transactions on Graphics, Vol.25, No.3, p.533-540 (2006).
- (10) G. D. Forney Jr : The viterbi algorithm. Proceedings of the IEEE, Vol.61, No.3, p.268-278 (1973).

著者紹介



Wei Liu

富士通研究開発中心有限公司 所属
現在、画像処理および画像認識に関する研究に従事。



Sun Jun

富士通研究開発中心有限公司 所属
現在、画像処理およびディープラーニングに関する研究に従事。



Wei Fan

富士通研究開発中心有限公司 所属
現在、画像処理および画像認識に関する研究に従事。



直井 聡 (なおい さとし)

富士通研究所 フェロー 兼
富士通研究開発中心有限公司 董事兼総
経理



Li Chen

富士通研究開発中心有限公司 所属
現在、画像処理および画像認識に関する研究に従事。