

カメラペンシステムにおける up・down 判定と筆跡位置修正の有効性評価

工藤 力[†] 黄瀬 浩一^{††}

[†] 大阪府立大学工学部

^{††} 大阪府立大学大学院工学研究科

〒 599-8531 大阪府堺市中区学園町 1-1

E-mail: †kudo@m.cs.osakafu-u.ac.jp, ††kise@cs.osakafu-u.ac.jp

あらまし カメラペンシステムとは、小型カメラを取り付けたペンを使用して印刷文書への筆跡をデジタルデータとして復元するシステムである。このシステムは、筆跡を求める処理と、その筆跡の文書上の位置を求める処理から成り立っており、前者には筆跡形状復元法、後者は文書画像検索法を用いる。文書画像検索を用いて、撮影画像からデータベースに登録した文書を検索し、筆跡形状復元法を用いて求めた筆跡を文書上に記録することができる。本稿では、このカメラペンシステムに対し、up・down 判定法と筆跡位置修正を行う手法を加えることを提案する。up・down 判定法を加えることにより、多画数の手書きに対して筆跡を復元することができる。また、筆跡位置修正を行うことにより、ペン先の開始点がずれることから生じる問題を解消することができる。これらの手法の有効性を調べるために、多画数の筆記を行い、筆跡位置修正の有無による誤差の違いを調べる比較実験を行った。比較実験の結果から、これらの手法は、カメラペンシステムに対して有効であることがわかった。

キーワード 文書画像検索, 文書画像処理, 手書きパターン, カメラペンシステム

1. はじめに

日々刻々とペーパーレス化が進み、デジタル化してゆく情報社会の現代でも、「紙とペン」による手書きは、日常の記録を取るための重要な手段である。手書きによる記録でも、今なお大切な情報が扱われていることが多いため、それらを保存することは重要である。本稿では、手書きによる記録をデジタルデータとして蓄積する、ライティングライフログ (writing-life log) を提唱する。ライティングライフログにより、手書きを収集することで、手書き情報の共有や検索、編集が容易になり、それらを有効活用することができると考えられる。

紙とペンを用いた手書きによる記録を簡単にデジタルデータ化する手法として、近野らのカメラペンシステム [1] がある。このシステムでは、筆記した紙の凸凹 (紙指紋) をカメラで追跡することにより筆跡を復元する。また、文書上の筆記位置を特定することができ、対応する電子文書上に筆跡を復元できる。しかし、近野らのシステムでは、2つの問題点がある。1つ目は、ペンが紙面についているか否かの判別ができないため、一筆書きの筆跡にしか対応できないという問題である。2つ目は、書き始めの際にペンを傾けると、筆跡開始位置がずれるため、復元された筆跡にずれが生じるという問題である。

この2つの問題点を解決するために、up・down 判定法と筆跡位置修正の処理を導入する。up・down 判定とは、紙面にペン先がついているかどうかを判定するものである。up・down 判定には筆圧を感知するセンサと、それをシステムに伝えるモジュールが必要となる。筆跡位置修正では、文書上に書いた筆

跡を用いて、筆記開始時にずれていたペン先位置を修正する。これらの手法により、精度の高いカメラペンシステムが実現できると考えられる。

本稿では、近野らのシステムにマイクロスイッチを導入し、多角文字に対応可能なカメラペンシステムの実現を提唱する。提案システムでは、ペンにマイクロスイッチを取り付け、スイッチにかかる圧力を検出することでペンの up・down を判定する。さらに、筆跡位置修正では、撮影画像に写った筆跡を利用することで、ペン先の位置のずれを修正することを試みる。本システムを用いて、筆跡復元の精度を評価する実験を行ったところ、一筆書き以外の筆跡にも対応でき、より精度の高い筆跡復元を行えることが確認できた。

2. 関連手法

手書きによる記録をデジタルデータに変換するデバイスは、既製品として多く世の中に出回っている。それらに使用している手法は大きく3つに分類できる。

1つ目は、電磁誘導方式である。代表的な製品として、Technote [2] や CyberTablet [3] などが挙げられる。タブレットのような磁場を発生することができる物の上に筆記を行うと、電子ペンに内蔵されたコイルに誘導電流が流れる。電磁誘導方式は、その誘導電流を検出することで、ペン先位置を求める手法である。筆跡の復元精度は高いが、電子ペンとタブレットが必要であるため、携帯性が低いという問題がある。

2つ目は、超音波方式である。代表的な製品として、Airpen [4] や TAXAN のデジタルインクペン [5] などが挙げられる。この

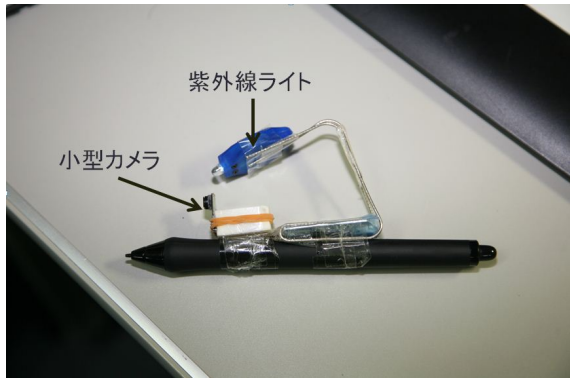


図 1 カメラペン

手法では、超音波と赤外線を発信するデジタルペンと、小型の受信機を用いる。小型の受信機を固定し、ペンから送信された超音波や赤外線の受信時間の差を利用して、ペン先位置を特定する手法である。利点として、使用するデバイスが小型なので携帯性に優れていることが挙げられる。その一方で、どの文書を扱っているか特定できないため、文書上に筆跡を復元することができない。また、筆記位置を求める際、送受信を使用しているためノイズに弱いという問題がある。

3つ目は、ドットパターン認識方式である。代表的な製品として、アノトペン [6] が挙げられる。この手法では、ドットパターン埋め込んだ専用紙とカメラを搭載したペンを用いる。ペンの先端についているカメラでドットパターンを読み取り、ペン先の絶対座標を求めることができる。紙とペンしか使用しないため、携帯性に優れており、筆跡の復元精度も高い。しかし、普通紙では利用できず、特殊な用紙が必要不可欠であることが問題がある。

電磁誘導方式ではタブレットのような磁場を発生できるデバイス、超音波方式では小型の受信機、そしてドットパターン方式はドットを埋め込んだ専用紙が必要である。これらの制限を撤廃するため、我々は、デジタルペン以外のデバイスや専用紙が不要なシステムの開発を目指している。それが、近野らのカメラペンシステムである。このシステムの利点は、使用するものはカメラペンのみなので、携帯性に優れており、安価な部品のみで構成されているため、コストパフォーマンスが良い。また、あらかじめ文書をデータベースに登録しておくことにより、撮影画像から文書を特定できるため、文書上に筆跡を復元することができる。

しかし、このシステムでは、多画数の筆跡に対して、一筆書きで筆跡を復元してしまう問題がある。また、書き始めにペンが傾くことで、ペン先の座標の位置が推定位置よりずれてしまう問題がある。詳細は 3. で述べる。

3. 既存手法

既存手法で使用するカメラペンは、一般的に売られているペンに小型カメラと紫外線ライトを取り付けたものである。小型カメラと紫外線ライトの位置関係を図 1 に示す。既存システムでは、筆跡の形状を復元する処理と、筆記位置を特定し、復元

した筆跡を文書上に配置する処理から成り立っている。

3.1 筆跡形状復元

筆跡の形状復元は、撮影画像から得られた特徴点を追跡することで実現する。紙から特徴点を抽出するために紙指紋を利用する。紙指紋とは、紙の繊維から生じる凹凸の模様のことである。紫外線ライトを照射することにより、より鮮明な紙指紋を得ることができる。紙指紋に対し、コーナー検出を行うことで、特徴点を抽出する。また、特徴点の追跡には、局所的な領域の特徴を利用する LK トラッキング [7] を用いる。この手法では、連続したフレーム間の特徴点からオプティカルフローを求める。オプティカルフローは、画像の輝度情報を基に求めた物体の動きベクトルのことである。オプティカルフローにより、連続する 2 フレーム間の特徴点の対応がとれるため、各 2 フレーム間の射影変換行列を推定することができる。ここで、 F_i をフレーム i と $i+1$ 間で計算された射影変換行列であるとする。最新のフレームのペン先位置 (x_p, y_p) は F_i を用いて初期フレーム上の座標 (x'_p, y'_p) に以下のように変換される。

$$(x'_p, y'_p) = (x_p, y_p) \prod_{k=1}^i F_k \quad (1)$$

このようにして、各フレームのペン先位置をそれぞれの射影変換行列を用いて変換することで、初期フレーム平面上に置くことができる。最後に初期フレーム上のペン先の座標群を繋ぎあわせることで、筆跡の形状を復元する。

3.2 筆記位置特定

文書上に筆跡復元するために、文書画像検索を利用する。文書画像検索を用いることで、カメラの撮影画像から文書とペン先の推定位置を求めることができる。

既存手法では、文書画像検索の手法の一つである LLAH (Locally Likely Arrangement Hashing) [8] を用いる。LLAH は、文字の重心を特徴点とし、特徴点を表現する特徴量に基づいて検索する手法である。登録文書画像と撮影画像との特徴点の対応関係を利用することで、文書上の撮影範囲を特定することができる。そのことによって、撮影画像から文書画像への射影変換行列を求めることができる。この射影変換行列を用いることにより、筆跡を文書上に配置することができる。ペンが大きく傾くことで、撮影画像に強い射影歪みが含まれると、データベースに登録された画像と大きく異なってしまうため、検索失敗の原因となる。そこで、撮影画像に対して、幾何学的変換を行うことで、傾きの小さい画像を生成し、データベースとのマッチングを取りやすくするクエリ拡張を行う。

3.3 クエリ拡張

ペンが大きく傾くことで、撮影画像に強い射影歪みが含まれると、データベースに登録された画像と大きく異なってしまうため、検索失敗の原因となる。そこで、データベースとのマッチングを取りやすくするため、撮影画像に対して幾何学的変換を行い、傾きの小さい画像を生成し、クエリ拡張を行う。

クエリ拡張を用いた文書画像検索では、撮影画像を用いてデータベース画像との射影変換行列 M を求める。次に、行列 M を用いて撮影画像を文書平面上に変換する。これを拡張ク

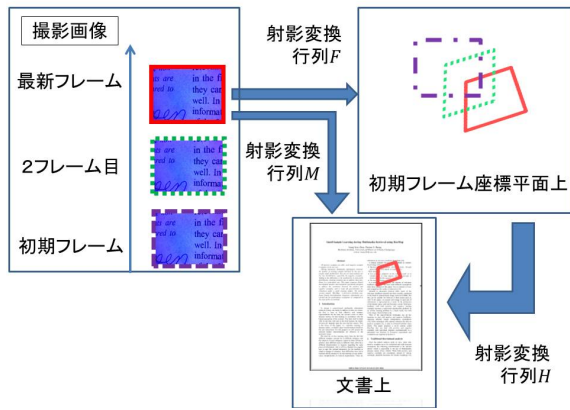


図 2 射影変換行列の推定

エリ 1 とする．拡張クエリ 1 は，行列 M に含まれている誤差のために，射影歪みを含んでいる可能性がある．拡張クエリ 1 を 3 つの異なる方向に回転させ，新たに拡張クエリを 3 つ生成する．生成方法は以下に示す通りである．まず，回転軸を設定する．回転軸は，拡張クエリ 1 の画像の底辺に対し，平行な軸とその軸に対して垂直な軸を設定する．両軸の交点は撮影画像の中心画像を基に設定する．画像の底辺に対して垂直方向に 10 度左回転したクエリ，また 10 度右回転したクエリを生成する．また平行な軸に対して手前方向に 10 度回転したクエリも生成する．これらの 3 つの回転は，ユーザに右利きの人が多いことから設定している．

撮影画像から文書画像検索をするたびに，拡張クエリを適用し，データベース画像との対応点が最も多いクエリを特定する．そのクエリとデータベース画像との対応点をもとに，誤差の少ない行列 M を求める．

3.4 文書上への射影変換行列の推定

筆記処理が終わったのちに，初期フレーム座標平面上に筆跡の形状を復元し，初期フレーム座標平面から文書画像平面に変換する射影変換行列 H を用いることで，文書上に筆跡を復元する．この筆跡を復元する際に用いる行列 H をロバストに推定するために 2 つの射影変換行列を用いる．1 つ目の射影変換行列は，クエリ拡張を用いることで求まる射影変換行列 M である．2 つ目の射影変換行列は，LK トラッキングを用いることで求まる最新フレームから初期フレームにおける射影変換行列 F である．図 2 より，この 3 つの行列について以下の式 (2) が成り立つ．

$$M = FH \quad (2)$$

処理するフレーム数が増えるたびに，式 (2) を使用し，行列 H を求めることができる．各フレームに対する，行列 H を保存しておき，筆記の最後に保存した複数の行列 H に基づき，もっとも適切な行列を推定する．理論的には，複数の行列 H は同一の行列である．ただし，この保持した複数の行列 H には，ノイズが含まれる．そこでノイズを取り除き，残りの行列の平均値を求めることとする．

ノイズを取り除く方法として，以下の方法を用いる．推定さ



図 3 既存手法による筆跡

れた行列を $H_i (1 \leq i \leq n)$ とし，差を $ij = \|H_i - H_j\|_F$ で表す． H_i に対しての ij が閾値以下の行列 H_j の集合を G_i とし， G_i の中で要素数最大のもを G_a とする．これらを使用し $H = 1/|G_a| \sum_{g \in G_a} g$ を推定行列とする．

3.5 問題点

既存手法には 2 つの問題がある．1 つ目は，撮影画像の情報のみで筆跡を復元しているため，ペン先が紙面についているかどうかの判定ができないという問題である．従って，図 3 のように，ペン先の移動は全て筆跡として復元されるため，多画数の筆跡には対応できない．2 つ目は，筆記開始時に，ペンが傾くことにより復元された筆跡にずれが生じる問題である．本来なら，射影変換行列 H からペンの傾斜角度が求まるので，3D モデルを作ることによって撮影範囲とペン先の位置関係を求めることができる．しかし，既存システムでは，簡単化のために 3D モデルを使わず，射影変換による歪みは無視できるとしている．そのためずれが生じている．

4. 提案手法

多角数の文字に対応した復元精度のより高いカメラペンシステムを実現させるため，ペンの up・down 判定法と筆跡位置修正を行う処理を加える．ペンの up・down 判定法とは，マイクロスイッチの情報からペンが紙面についているかどうかを判定する方法である．

また，実際に書いた筆跡と復元した筆跡の間には，ずれが生じる．それらは，平行移動，回転，スケール変換などの幾何学的なずれである．その中で，影響の大きい平行移動のずれを筆跡位置修正で対処する．以下各節で，up・down 判定法と筆跡位置修正の処理について詳しく説明する．

4.1 up・down 判定法

提案手法では，ペンの up・down を判定するため，マイクロスイッチとモジュールを導入する．マイクロスイッチは，図 4 に示すように，ペンのインクチューブの端に取り付け，ペン先が紙面に接触することで押される．ペンの完成形を図 5 とする．

up・down 判定法はカメラペンシステムが扱っている撮影画像の各フレームについて判定する．各フレームごとのペン先座標に up か down かの情報を付加することで，そのフレームでのペン先が紙面に接触したかどうか分かる．また 2 フレーム間の座標を調べ，両フレームの座標がともに down だった際に，ペン先座標間を線で結ぶことで筆跡を復元する．

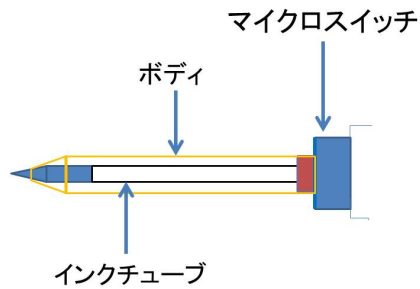


図 4 スイッチとインクチューブの関係

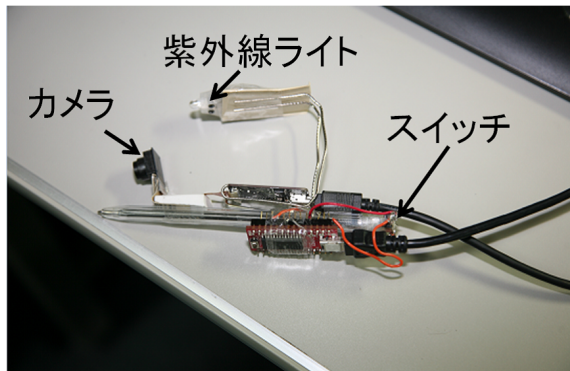


図 5 ペンの完成形

上記手法を繰り返すことにより 2 筆目以降の筆跡も復元することができる。

4.2 筆跡位置修正の処理

筆跡復元に使用するペン先の座標は各フレームから 1 つずつ得ることができる。この得られた連続したペン先座標を撮影画像上にプロットする。プロットした座標の集合を以下、ペン先座標群 X とする。このペン先座標群と文書上に書いた筆跡を使用して、筆記開始時に生じるずれの修正を行う。この時に使用するペン先座標数は多い方が正確な修正が行える。そのためまず、撮影範囲内に一番ペン先の座標が多かった時の撮影画像を特定する。特定した画像とペン先座標群から撮影範囲内にある筆跡位置を求める。

具体的な筆記位置の求め方として以下の方法を用いる。特定したカラーの撮影画像を、固定閾値で二値化し、二値画像 $I(x, y)$ を得る。文書上に書いた筆跡は、ペン先の座標を結んだ形状と同じ形をしているため、文書上に対してペン先座標群を平行移動させることで、筆跡の形状と重なり合うと考えられる。そこで、二値化画像上でペン先の座標群を N pixels 分 x, y 軸ともに平行移動させ、ペン先座標群が筆跡と重なる箇所を求め、平行移動した pixel 数をずれとして修正する。具体的に二値化画像上で筆跡を見つける方法として、二値化画像上でペン先の座標群が一番黒画素と重なり合う位置に筆跡があるとする。また、移動範囲 N は、復元する文書サイズに対して、最もずれが大きかった時の経験則に基づいて設定する。

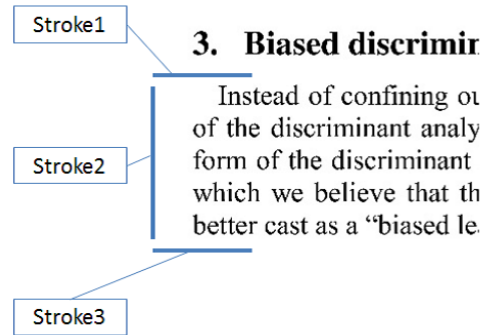


図 6 大括弧

この手法で問題となるのが、ペン先の座標群が文書上の印刷文字と重なってしまう可能性があることである。しかし、一般的に文書に筆記をする際、印刷文字の上を書くことはないため、筆跡の方が多くペン先座標と対応が取れると考えられる。また、特定した 1 枚の画像を使用するのではなく、全てのフレームを繋ぎ合わせることで得られる 1 枚のモザイク画像を使用する方法も考えられる。この方法では、特定した 1 枚の撮影画像に入りきらなかった筆跡情報を得ることができるという利点が考えられる。しかし、モザイク画像の生成は、処理時間がかかる上にモザイク画像上での位置ずれも発生することが問題となる。そこで、本手法では、簡単化のために特定の 1 枚の画像だけを使用する。

5. 実 験

提案手法の有効性を検証するために、多画数の筆記に対する文書上に復元した筆跡の精度実験を行った。筆跡復元に使用した文書のサイズ 1700×2200 である。カメラは 30fps のものを使用し、カメラのフレームサイズは 640×480 である。また、筆跡位置修正の処理としてペン先座標群を平行移動させた距離は、上下左右に 40pixels とした。

5.1 実験方法

筆記を行う前にあらかじめ、ペン先が紙面に着く箇所と離れる箇所を決めて、扱う文書にドットを打つ。そのドットに添いながら、文書中にペン先を合わせる形で筆記を行う。正解画像のドットの座標と、筆跡復元に使用したペン先の座標を比較し、座標間の距離を誤差とする。筆跡位置修正を行う前後で、誤差の平均値の変化を調べる。上記の方法で、筆記位置修正の有効性を検証する。

手書きは、2 種類の筆跡を用いた。1 種類目は A から E までのアルファベットの内、C を除いた筆跡とする。C は、1 ストロークの文字であるため up・down 判定の確認には不適であるため除いた。筆記は、文書に対して 2 行から 4 行分に高さを変えて行った。また 4 行目に関しては、筆跡が印刷文字にかかっているが、筆跡の大部分が余白にあるため、筆跡位置修正の精度を測るのに影響はないと考えられる。2 種類目は印刷文字の行を囲むようにして、図 6 に示す様に大括弧を筆記した。大括弧の 2 画目は、文書の 2 行目から 8 行目まで長さを変えて筆

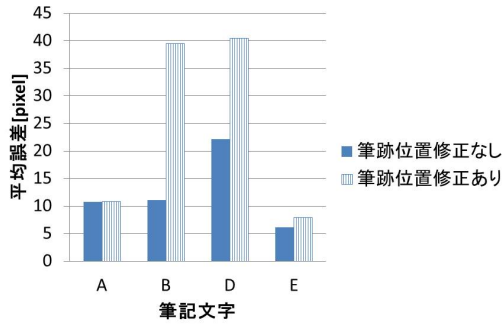


図 7 2 行分の大きさのアルファベットを書いた平均誤差

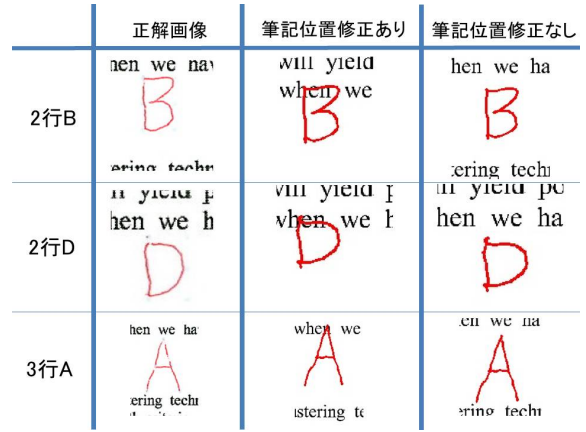


図 10 文書上の筆跡と筆跡復元の一例

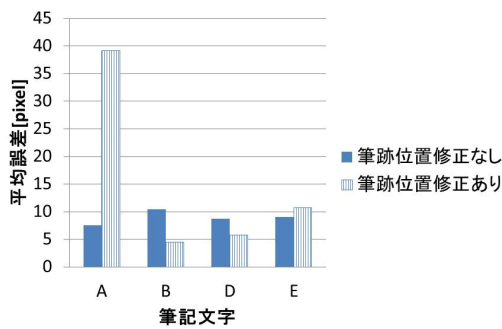


図 8 3 行分の大きさのアルファベットを書いた平均誤差

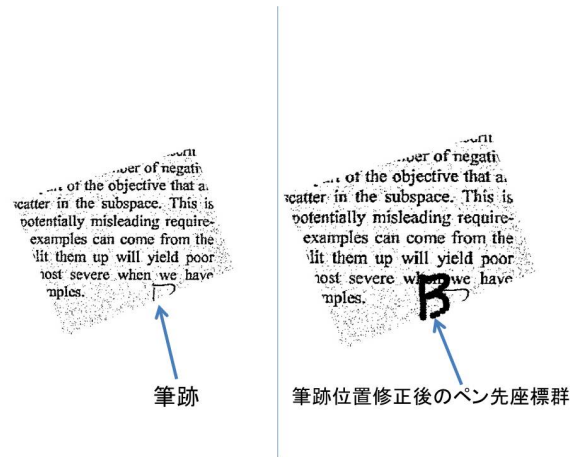


図 11 文書画像平面に射影変換した撮影画像

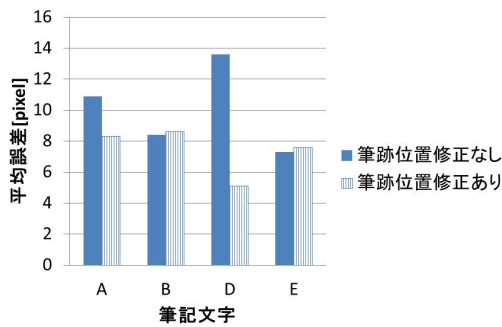


図 9 4 行分の大きさのアルファベットを書いた平均誤差

記を行った。大括弧の筆記は、1 画目は左から右方向に、2 画目を上から下方向に、3 画目は左から右方向に行う。よって 1 画目と 2 画目を書く際、ペンをあげる必要がある。そのため、up・down 判定が正しく機能していれば、1 画目の終了点から 2 画目の開始点へ、ペン先を浮かして運ぶ軌跡が復元されない。よって文書上に復元した筆跡を調べることで、up・down 判定の有効性が検証する。

5.2 実験結果

アルファベット及び大括弧の筆跡における実験結果を示す。

5.3 アルファベットの結果

図 7 から図 9 にそれぞれ、2 行から 4 行分と大きさを変えてアルファベットを書いた際の平均誤差と文字列を示す。文字が大きくなることによって、筆跡位置修正を行った場合の誤差が減少している。よって文字が大きいに筆跡位置修正の有効性が確認できる。また、図 10 に示すように、復元した筆跡は 2 画以上であるため、up・down 判定が正しく行われていることがわかる。よって up・down 判定の有効性が確認できる。筆跡位置修正を行うことで、平均誤差が大きくなった理由を考察すると、大きく分けて 2 つのパターンが考えられる。

1 つ目のパターンは、印刷文字にペン先座標群が誤対応した場合である。記述の大部分は余白に書かれるために、ペン先座標群は、筆跡の方に多く対応が取れると考えた。しかし、図 11 の様に、筆跡が撮影画像上に少し写っていない場合、ペン先座標群は筆跡よりも印刷文字に対応が取れてしまうことがわかった。この現象は、図 7 の筆跡 B、D と図 8 の筆跡 A に見られる。

2 つ目のパターンは、LK トラッキングの誤差により筆跡の形状が変化した場合である。この現象が見られた筆跡は図 8 の筆跡 E と図 9 の筆跡 B と E である。図 12 より、復元画像の方が正解画像に対して、筆跡の高さが短いことがわかる。さらに修正なし画像は、正解画像の下部で一致しており、上部でず

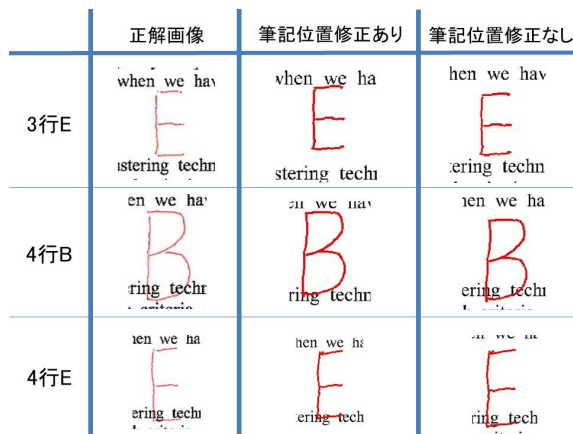


図 12 文書上の筆跡と筆跡復元の一例

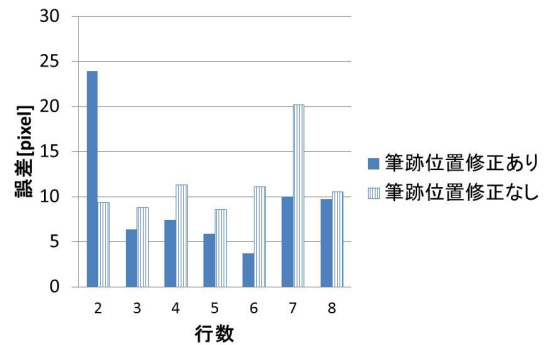


図 13 大括弧の各行に対する平均誤差

表 1 2 画目の長さ

	正解画像 [pixel]	復元画像 [pixel]	誤差 [pixel]
2 行	64	55	-9
3 行	96	102	6
4 行	128	128	0
5 行	160	153	-7
6 行	192	184	-8
7 行	224	206	-18
8 行	256	239	-17

れている．このずれを修正したため，修正あり画像の下部での差が大きくなり，わずかに平均誤差が大きくなったと考えられる．この問題は，LK トラッキングの精度を上げることで解決できる．

5.4 大括弧の結果

大括弧で囲んだ各行に対して，6つの座標の平均誤差を，図13に示す．括弧の長さが3行より長い場合，筆跡位置修正が有効であることがわかる．また7行から8行を，大括弧で囲んだ時，筆跡位置修正を行っても，平均誤差が10pixelsほどずれていることがわかる．

ここで，表1に正解画像の2画目の長さとして筆跡復元画像の2画目の長さの差を示す．誤差の値は，正解画像の方が長い時を正，短い時を負とする．表1から，括弧の長さは，7,8行間にわたると，急に短く復元されることがわかった．このような結果になった理由として，LK トラッキングの際にでる誤差と考えられる．また，表1と図7により括弧の長さが2行分以下の場合，撮影画像に筆跡が小さいため，誤対応を起こしてしまうことが判明した．よって，今後の課題として，常にペン先がカメラに映る構造を考案する必要がある．

以上2つの実験結果から，提案手法によるペンの up・down 判定法は有効であり，筆跡位置修正は，LK トラッキングによる誤差を除けば，撮影範囲内に十分な筆跡が映っている際に，有効であると考えられる．

6. まとめ

既存手法によるカメラペンシステムには，2つの問題点があ

る．1つ目は，ペン先の移動が全て筆跡として復元されるため，多画数の筆跡には対応できないという問題である．2つ目は，筆記開始時にペンが傾くことにより復元された筆跡にずれが生じる問題である．本稿では，これらの問題を解決するために，2つの手法を提案した．1つ目の問題に対して up・down 判定法を用いた．up・down 判定は，マイクロスイッチを加えたカメラペンを用いて行う．ペンの up・down 判定を行うことで，各フレームにおけるペン先の紙面への接触を判定し，多画数の筆跡に対応することを提案した．2つ目の問題点に対して，撮影画像に映った筆跡を用いて，筆跡位置を修正する手法を用いた．撮影画像を文書画像平面上に射影変換し，ペン先座標と筆跡のずれを求めることで，筆記開始時に生じるの修正を行う．

この2つの手法を既存手法に加えることで，多画数に対応し，LK トラッキングの誤差がなく筆跡が撮影範囲内に十分に入っている場合は，ペン先位置のずれを修正することができた．

今後の課題として，撮影画像に必ずペン先が映るような構造を考える必要がある．より精度の高い筆跡位置修正を行う方法の考案や，LK トラッキングの精度向上などが挙げられる．

謝辞 本研究の一部は日本学術振興会科研費基盤研究 (B) (22300062)，挑戦的萌芽研究 (21650026) の補助による．

文 献

- [1] M. Chikano, K. Kise, M. Iwamura, S. Uchida, and S. Omachi, "Recovery and localization of handwritings by a camera-pen based tracking and document image retrieval," Pattern Recognition Letters, 2012, inpress.
- [2] <http://pothos.to/technote/technote.html>.
- [3] <http://www.adesso.com/en/component/content/article/58-graphic-tablet/528-cybertabletz8.html>.
- [4] <http://www.airpen.jp/>.
- [5] <http://www.taxanprojector.jp/pastproduct/mousepen/>.
- [6] <http://www.anoto.com/>.
- [7] B.D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," Proc. of IJCAI, pp.674-679, 1981.
- [8] 竹田一貴, 黄瀬浩一, 岩村雅一, "メモリ効率と安定性を向上させた llah によるカメラベース大規模実時間文書画像検索" 画像の認識・理解シンポジウム (MIRU2011), pp.350-357, jul 2011.