

カメラペンシステムにおける筆跡復元精度の向上

近野 恵[†] 黄瀬 浩一[†] 岩村 雅一[†] 内田 誠一^{††} 大町真一郎^{†††}

[†] 大阪府立大学大学院 工学研究科 〒599-8531 大阪府堺市中央区学園町 1-1

^{††} 九州大学大学院 システム情報科学研究科 〒819-0395 福岡市西区元岡 744

^{†††} 東北大学大学院 工学研究科 〒980-8579 仙台市青葉区荒巻字青葉 6-6-05

E-mail: [†]chikano@m.cs.osakafu-u.ac.jp, ^{††}{kise,masa}@cs.osakafu-u.ac.jp, ^{†††}uchida@ait.kyushu-u.ac.jp,
^{††††}machi@ecei.tohoku.ac.jp

あらまし 本稿では、我々が開発中のカメラペンシステムにおける筆跡の復元精度向上法を提案する。カメラペンシステムとは、カメラを取り付けたペンを用いて紙面に筆記した文字や図等を、筆跡そのままにデジタルデータ化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 SURF トラッキング、後者は文書画像検索法を用いる。ここで、復元した筆跡にずれや欠損が含まれる等、復元精度が低いことが問題となっている。この主な原因は、SURF トラッキングの精度が低いことである。この問題を解決するために、LK トラッキングを利用する。効果を検証するため、SURF トラッキングを用いた従来手法と、LK トラッキングを用いた提案手法の比較実験を行った結果、従来手法に比べ精度の向上が見られた。

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

1. ま え が き

タブレットやスマートフォンといった、手軽に携帯できるデジタル機器の普及が目覚ましい一方、とっさに書く備忘録や、アイデアを考えながらのメモ書き、板書の記録、会議資料への書き込み等、用紙に手書きすることも多い。しかし、手書きした情報を、編集や共有がし易いデジタルデータに変換するには、スキャナや手作業での入力が必要不可欠であるため、デジタル機器を用いて直接記録する場合に比べ、多くの手間が必要となる。用紙への手書きという簡単な記録方法を残しつつ、筆記した情報をデジタル化することが可能であれば、情報の有効な利用や、デジタル化に掛かる手間の削減につながると考えられる。

手書きとデジタル化の利点を融合した実用的なシステムの 1 つにアノトシステム [1] がある。これは、特殊なドットを印刷した用紙と、カメラを搭載したペンを使用することで、紙面上のペン先の位置を特定し、筆跡を復元するシステムである。このシステムは、用紙への手書きを自動でデジタル化できる一方、特殊な用紙が必要であるという問題がある。

この問題を解決するため、岩田らは特殊な用紙を使用せずに筆跡をデジタルデータ化するカメラペンシステム [2] を提案している。このシステムでは、撮影画像を用いて紙面の動きをトラッキングすることにより筆跡を復元すると共に、データベースから対応する範囲を検索することで、用紙に対応する電子文書上に筆跡を復元する。ただし、現状ではトラッキングの精度が低いという問題点がある。これにより、筆跡にずれや欠損が生じる等、筆跡復元精度の低下を引き起こす。また、トラッキング結果を用いて文書画像検索を行うため、復元位置が異なる、文書上へ復元できない等の問題が生じる。

本研究では、これらの問題を解決するために、安定性の高い手法として知られ、OpenCV でも利用可能な LK(Lucas-Kanade) 法 [3] を用いたトラッキング (LK トラッキング) を筆跡の復元処理に使用することで、復元精度の向上を図る。また、文書画像検索に、クエリの拡張手法 [4] を追加することにより検索の失敗を防ぐ。これらの効果を検証するため、従来手法と、上記処理を用いた手法の比較実験を行った。その結果、LK トラッキングとクエリの拡張を組み合わせた手法において、筆跡の復元精度が最も向上した。

2. 関連研究

紙媒体の情報は、扱い易さや読み易さという利点がある一方、紙媒体の情報から電子データにアクセスできない等の隔りがある。この隔りを埋めて、紙媒体と電子媒体を同様に扱うことができれば用途は広がると考えられる。

この目的のため、3 つの商用システムが開発されている。1 つ目は、電磁誘導方式のシステム [5] である。タブレット上に用紙を置き、デジタルペンから発信される微弱な電波を、タブレットで受信することでペンの位置を特定する。復元精度は高いが、用紙より大きいタブレットを使用するため携帯性が低い点が問題である。2 つ目は、赤外線と超音波の速度の違いを利用したシステム [6] である。デジタルペン以外にセンサを搭載した小型のデバイスが必要であるため、携帯性が低下するという問題がある。3 つ目のアノトシステムでは、特殊なドットが印刷された用紙と、カメラを搭載したペンを使用する。このシステムでは、筆記時に撮影された複数のドットの配置を利用して筆跡を求める。デジタルペン以外のデバイスが不要で携帯性に優れているが、専用紙が必須という問題が残されている。

研究としては、以下の手法がある。

印刷文書と電子データを関連付ける手法として PaperLink [7] が提案されている。これは、印刷文書にマーカで印を付けると、ハイパーリンクのように電子データに関連付けるシステムである。紙媒体のデータと電子データとをつなぐという面で画期的なシステムであるが、手書きの文字は対象とされていない。

手書きを対象としたシステムとして、紙面への手書きをデジタル化する手法を3つ紹介する。まず1つは、Munich らが提案した、ペン先を検出することで手書きやスケッチを復元するシステム [8] である。2つ目は、署名照合システム [9] である。これは、2つのカメラを用い、ペン先をトラッキングすることで求めた署名と、照会用の署名を比較するシステムである。3つ目は、Seok らによる印刷文書上の筆跡を追跡するシステム [10] である。このシステムでは、ペン先のトラッキングとアップダウン判定により、実時間で筆跡を復元する。これら3手法に共通する利点は、特殊なペンや用紙が不要なことである。一方、紙面やペンの動きを撮影する固定カメラが必要なため、携帯性が低下する上に、使用する度にカメラを設定し直さなければならないという問題点もある。

これらに対し、カメラを取り付けたペンを使用するシステムが提案されている。その1つに、用紙の紙指紋を利用し、トラッキングにより筆跡を復元するシステム [11] がある。再出現点が求められるという利点があるが、筆跡の位置や大きさが求められないという問題点もある。一方、文書画像検索を用いることで、文書上に筆跡を復元するシステム [12] がある。実時間で文書上に復元できるが、白紙には対応出来ない点が問題である。これら2つのシステムの利点を生かし、統合したシステムが岩田らによるカメラペンシステム [2] である。このシステムでは、復元精度が問題となっている。詳細は4. で述べる。

3. カメラペンシステム

我々は、上記の問題を解決するため、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。このシステムでは、筆跡を復元するだけでなく、対応する電子文書上に筆記位置そのままに復元することも可能である。これにより、文書に印や訂正等を書き加えた場合、筆記情報と文書との関連が一目で分かるため、より有益な情報が得られると考えられる。また、筆跡をその形状のまま復元することで、手書きの図やイラストもデジタルデータとして残すことができる。手書き文字は、復元後に OCR(Optical Character Recognition) 等の既存の技術を用いれば、テキスト化することも可能である。

このようなシステムが実現できれば、手書き情報の活用の用途が広がると考えられる。例えば、タグ付けして保存しておくことで、必要な情報を高速に検索でき、編集や共有が容易になる。保管の点でも、紙媒体のまま保存するより場所の節約になる。また、ライフログの蓄積にも役立つと考えられる。ライフログは、日常の行動を長期間にわたりログとして残すことである。その1つのライティング・ライフログは、筆記をログとするものである。例えば、メモ書きや日記、食事の記録、スケッチ等日常的な記録を日々残すことで、過去の振り返り・生活の

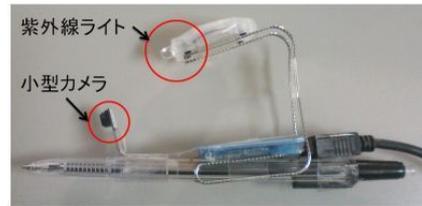


図1 カメラペン

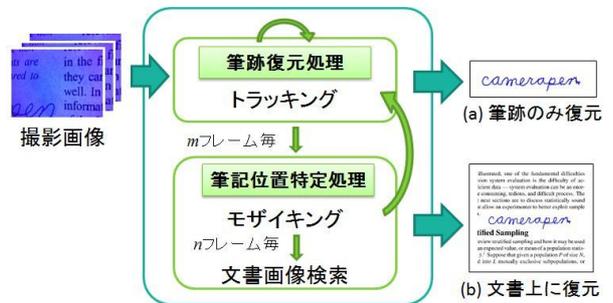


図2 筆迹復元の流れ

見直しへの利用や、行動の傾向から新たなサービスを提供することも考えられる。カメラペンシステムは、日常的に増えるこれらの膨大な情報を、デジタルデータで残す際に、デジタル化に掛かる労力を省く上で役立つと考えられる。

4. 従来手法 [2]

まず、提案手法の基礎となる岩田らのカメラペンシステム [2] について述べる。同システムでは、図1のような小型カメラを取り付けたペンを使用する。紙指紋をカメラで鮮明に撮影するため、このペンには、撮影範囲に光が当たるように紫外線ライトを取り付ける。紙指紋は、紙の繊維による紙面の凹凸のことで、筆跡を復元する際に紙面の特徴として用いる。

4.1 処理の流れ

処理の流れを図2に示す。このシステムは、筆迹復元処理と、文書上の筆記位置の特定処理から成り立っている。まず、筆記時に紙面を撮影した動画から SURF(Speeded Up Robust Features) を抽出する。次に抽出した SURF 特徴点を用いてトラッキングを行うことで、ペン先の移動量を求め筆迹を復元する。この SURF を用いたトラッキング手法のことを、以下では SURF トラッキングと呼ぶ。この処理で図2(a)のように筆迹の形状を求めることができる。また、 m フレームに1回、筆記位置特定処理を行う。位置の特定には、文書画像検索法を使用し、データベースから印刷文書に対応する文書と、文書上のペン先位置を検索する。これを基に、図2(b)のように電子文書上に筆迹を復元する。以下に各処理の詳細を示す。

4.2 筆迹復元処理

まず、任意のフレームを基準フレームとする。次に、撮影画像の紙指紋や印刷文字から SURF を抽出する。SURF を使用するのには、回転やスケール変化に対してロバストなためである。また、特徴点の抽出に再現性があるため、1度フレームアウトした部分が再びフレーム内に写った場合、再出現点を求めることもできる。

処理は以下の通りである。抽出した SURF 特徴点を用い、

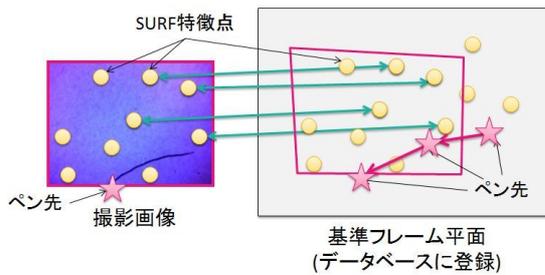


図 3 SURF 特徴点の対応関係

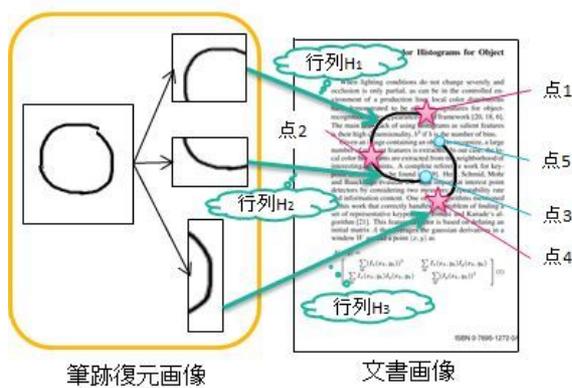


図 4 筆記位置特定

図 3 のように、ハッシュに保存された過去のフレームの SURF 特徴点との対応関係を求める。得られた複数の点对応関係を基に、撮影画像上の SURF 特徴点とペン先の位置を基準フレーム平面上に射影する。これにより、基準フレーム平面における現在のペン先位置を求めることができる。このようにして求めたペン先の位置を、時系列順につなぎ合わせるることにより筆跡を復元する。また、基準平面上に射影後の SURF 特徴点は、ハッシュに登録しトラッキングに使用する。

従来手法ではハッシュを 2 つ使用する。1 つは広範囲の SURF 特徴点を保存し、再出現点を求めるためのものである。もう 1 つには、過去数フレーム分の SURF 特徴点だけを登録することで、近傍の SURF 特徴点のみ保存し、現在の撮影画像から得られる SURF 特徴点と対応付き易くするものである。

4.3 筆記位置特定処理

筆記位置特定に用いる文書画像検索とは、文書画像をクエリとし、データベースの中からクエリに対応する文書画像とその範囲を検索する手法である。カメラペンシステムでは、その一手法である LLAH (Locally Likely Arrangement Hashing) [13] を用いる。LLAH は実時間検索が可能な手法であり、印刷文字から抽出した複数の特徴点 (LLAH 特徴点) の配置を用いて検索する。複数の LLAH 特徴点を組み合わせるため、撮影画像に含まれる特徴点数が少ないと、十分な数の組み合わせが作成できず、検索精度が低下する。

この問題に対応するため、検索の前に画像モザイクングを行う。画像モザイクングは、複数の画像を用いて画像を合成する処理である。このシステムでは、複数の撮影画像を基準フレーム平面上に変換し張り合わせ、その結果得られる画像をクエリとすることで、1 枚の画像から得られる特徴点数を増やす。

次に、筆記位置の特定方法について示す。まず、データベースと撮影画像から求めた LLAH 特徴点を比較し、対応関係を求める。その対応関係を基に、撮影画像におけるペン先の座標を文書上に変換する。

上記の検索処理は数フレームに 1 回行われるため、検索と検索の間は、文書上の位置を特定することができない。そこで、次の検索までの筆跡は、4.2 で求めた筆跡で補う。具体的な処理は次のようになる。まず、トラッキングにおける基準フレーム平面と文書画像平面の射影変換行列 H_i を求める。これは、検索処理が行われる度に、検索結果とトラッキング結果から計算される。次に、 H_i を用いて、4.2 で求めた筆跡を文書画像上に変換する。例えば、図 4 の点 1 から点 2 までは、点 1 で求めた H_1 を用いて筆跡を変換し、点 2 から点 3 までを、点 2 で求めた H_2 を用いて筆跡を変換する。このようにして、文書画像上に筆跡を復元する。

4.4 従来手法の問題点

従来手法の主な問題点として、SURF トラッキングの精度が低いことが挙げられる。精度の低下は、筆跡復元処理、筆記位置特定処理の両方に影響を与える。前者では、SURF トラッキングを利用してペン先の動きを求めるため、精度が悪化すれば筆跡の形状を正しく求めることができず、筆跡にずれや欠損部分が生じる等の問題が起こる。後者では、SURF トラッキング時に求める SURF 特徴点の対応関係を基にモザイクング画像を作成するため、点对応が正確でないと、モザイクング画像にずれやノイズが生じる。この画像をクエリとして検索した場合、画像から得られる LLAH 特徴点がデータベースに登録された特徴点と大きく異なってしまうため、対応する文書画像が検索できない。これにより、電子文書上に復元できない場合が生じる。

SURF トラッキングの精度低下の原因は、SURF 特徴点の対応関係を求める時に、誤対応点や未対応点が多く生じることである。これらが生じる理由は、SURF 特徴点を安定して取得できていないことである。特に、撮影画像に印刷部分と余白部分が写っている場合、印刷部分から多くの SURF 特徴点が抽出され、紙指紋からはほとんど抽出できず十分な数が得られない。もう 1 つの理由として、SURF 特徴点を保存するためにハッシュを 2 つ用いることがある。2 つのハッシュで用いる基準平面の傾きが違うことによって誤差が生じ、ハッシュに保存された値にずれが生じて、誤対応を招く原因となっている。

精度低下以外の問題として、複数の射影変換行列を用いることにより、筆跡にずれが生じることが挙げられる。例えば、図 4 のように、復元処理において一続きの滑らかな筆記を復元したにもかかわらず、文書上に変換した時にずれが生じることがある。このシステムでは、図 4 に示すように、筆跡の一部分を射影変換行列 H_i を用いて文書上に変換するという処理を繰り返すことで、文書上に筆跡を復元する。この変換の時に、毎回異なる行列を使用しているため、上記のようなずれが生じる。

もう 1 つの問題として、メモリ使用量の増加がある。筆跡復元処理では、再出現に備えるため毎フレーム抽出した SURF 特徴点をハッシュに保存する必要がある。これにより、メモリ使用量が增大するという問題が生じる。



図5 クライアントサーバモデル

5. 提案手法

提案手法は、従来手法を基にしている。その処理の構成は、図2から画像モザイクングを省いたものである。従来手法と同様、処理時間削減のためクライアントサーバモデルを採用している。このモデルについては、提案手法での変更点を含めて5.1で説明する。また、上記の問題を解決するために、従来手法に3つの改良を加える。詳細は、5.2から5.4で述べる。

5.1 クライアントサーバモデル

図5に提案手法で使用するモデルを示す。クライアント側では、 i フレーム目 ($i > 1$) の撮影画像の取得、LLAH特徴点の抽出を行い、検索処理のためLLAH特徴点をサーバ側へ送る。次にサーバ側から $i-1$ フレーム目の文書画像検索結果を受け取り、検索結果とトラッキングにより得られた筆跡を基に、文書上の位置を特定する。続いて、 i フレーム目のトラッキングを行い筆跡を復元する。その後、再び撮影画像 ($i+1$ フレーム目) を取得し、同様の処理を繰り返す。

サーバ側では、受け取ったLLAH特徴点を用いて検索処理、射影変換行列の推定を行う。この時、クエリの拡張を行い安定性を高める。これにより、最終的に得られた検索結果をクライアント側へ送る。このように並列処理を行うことで、処理時間を短縮する。このシステムでは、サーバ側での処理時間のほうが短い場合が多く、この場合は送受信のタイミングを合わせるために、待機時間が生じる。

5.2 LKトラッキング

4.4で述べた、SURFトラッキングの問題点を解決するため、局所的な小領域の特徴を利用するトラッキングの一手法を用いることで、安定してトラッキングし、ずれや欠損を低減する。本稿では、LKトラッキング[3]を用いる。この手法では、連続する2フレーム間のオプティカルフローを計算し、これを基に2フレーム間の対応関係を求める。そして、SURFトラッキングと同様に、基準フレーム平面上にペン先位置を射影し、時系列順につなぎ合わせることで筆跡を求める。

LKトラッキングは、トラッキングの精度が高いために、広く一般的に利用されている。ただし、連続した2フレームのみを用いて動きを求めるため、複数フレームに対してこの処理を繰り返すうちに、誤差が蓄積していく可能性が考えられる。

5.3 クエリ拡張

従来手法における筆記位置特定処理の精度の低下は、画像モザイクングがトラッキング処理の結果に依存しているために生じる。したがって、画像モザイクングを省き、依存関係を無く

すことで、安定した検索結果を得られるようにする。しかし、画像モザイクングにより、1枚の画像から多くのLLAH特徴点を得ることで検索精度の向上を図っているため、単にこの処理を省くだけでは、文書画像検索の精度の低下につながる。

そこで、撮影範囲の狭い撮影画像をそのままクエリとして使用しても、ある程度の精度を維持できるように、クエリの拡張[4]を追加する。これは、クエリの候補を増やすことにより、データベースの点に対応付く可能性を上げる手法である。具体的には、撮影画像に加え、撮影画像に複数の傾きを加えたものもクエリとして用いる。

5.4 射影変換行列の推定

復元結果にずれが生じる原因として、文書上に筆跡を変換する時に、複数の射影変換行列 H_i を用いることが挙げられる。ここで、 H_i はトラッキングの基準平面から文書画像平面への射影変換行列であるため、理論的には、検索およびトラッキングの結果が正しければ、検索の度に毎回同じ値の行列が計算されると考えられる。しかし、実際に求められる行列には誤差が含まれるため、毎回同じ値にはならない。ただし、類似した値が得られると考えられる。そこで、複数の H_i を基に、変換する射影変換行列を1つに統合すれば、筆跡の途中でずれることがなくなると考えられる。また、多くの H_i を用いることができれば、外れ値を検出することが可能になるため、誤差の影響を受けにくくなると考えられる。本稿では、複数の H_i から1つの射影変換行列に統合する方法として、複数の H_i から外れ値を除いた平均値を用いる。

6. 実験

本稿では、LKトラッキングや拡張手法の、カメラペンシステムにおける有効性を検証するため比較実験を行った。使用した手法は以下の4手法である。

- (1) 従来手法: SURFトラッキング + モザイクング + 検索
- (2) 比較手法1: SURFトラッキング + 検索 (拡張有り)
- (3) 比較手法2: LKトラッキング + 検索
- (4) 提案手法: LKトラッキング + 検索 (拡張有り)

各手法の有効性を確認するため、トラッキングについては、SURFトラッキングを用いた手法(2)とLKトラッキングを用いた手法(4)、拡張については、拡張なしの手法(3)と拡張有りの手法(4)を比較した。

ここで、従来手法には、特徴点が安定して求められない問題と、ハッシュを2つ用いることにより誤差が生じる問題がある。したがって、手法(2)のSURFトラッキングには以下の様な改良を加えた。1つは、安定したSURF特徴点を取得するために、画像全体から一律に特徴点を抽出するのではなく、撮影画像をいくつかの領域に分割して、そこから特徴点を求めることである。これにより、白紙部分でも印刷文字の影響を受けることなく特徴点を取り出す。もう1つの改良として、誤差を減らすために、ハッシュ表を1つだけ使用する。これにより、全ての特徴点が1つのハッシュに保存されるため、誤対応の可能性が高まる。この問題には、探索範囲の制限で対応する。

6.1 実験条件

文書画像検索で使用するデータベースへの登録画像として、英文書画像 1,000 枚を使用した。登録画像のサイズは、5,100 × 6,600 である。

実験に使用したカメラペンは、正解画像作成に使用したタブレットのペンに、カメラと紫外線ライトを図 1 のように取り付けられたものである。データベースに登録された文書を A4 サイズの用紙に印刷し、その用紙上にカメラペンをういて筆記した。復元に使用した画像は筆記時に紙面を撮影した動画である。文書上に復元するため、この動画には白紙や余白だけ撮影されたものは含まない。撮影画像のサイズは 640 × 480、カメラのフレームレートは 30 fps である。本実験では、30 個の筆記を行った。撮影した画像にブレが生じると、特徴点を安定して抽出することができない。したがって、本実験ではモーションブラーを防ぐために通常よりゆっくり筆記した。また、ペンのアップダウンに対応していないため、筆記は一筆書きとした。

図 2 に示すように、手法 (1) では、5 フレーム毎に読み込んだ画像を用いて SURF トラッキングを行い、SURF トラッキングの $m = 5$ フレーム毎にモザイクングに移り、モザイクングの $n = 2$ フレーム毎に文書画像検索を行った。つまり、SURF トラッキングを 5 フレームに 1 回、モザイクングを 25 フレームに 1 回、検索を 50 フレームに 1 回行ったことになる。SURF トラッキングの度にモザイクングを行わないのは、撮影画像に写る紙面に大きな違いがないため、毎回画像を合成する必要性がないからである。他の手法では、モザイクング処理を行わないため、5 フレーム毎にトラッキングおよび検索を行った。

評価尺度として、(1) 筆跡の復元精度、(2) 処理時間、(3) メモリ使用量を用いた。以下に詳細を示す。

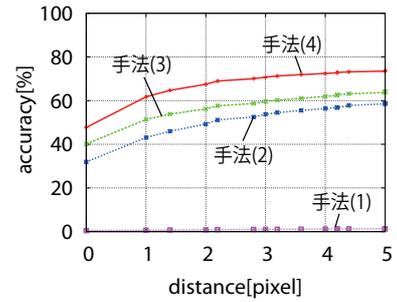
6.1.1 筆跡の復元精度

本実験では、正解画像と復元画像を重ね合わせた時の、筆跡部分のずれの距離と、精度の関係を求めた。具体的には、以下のように求めた。正解画像の筆跡部分の各画素から、復元画像の筆跡部分の最も近い画素までの距離を d とする。復元画像において、正解画像からの距離が d 以下の筆跡の画素数を求め、正解画像の筆跡の全画素数に対する割合を計算する。この割合を距離 d における精度とした。復元に使用した文書画像の大きさは、1,700 × 2,200 である。正解画像には、動画の撮影時にペンタブレットを用いて得られた筆跡を使用した。ただし、ペンタブレットにより得られた筆跡は、印刷時の紙面の傾きや余白の設定等により、実際の筆跡とはずれが生じる。そこで、ずれの補正後に評価を行った。

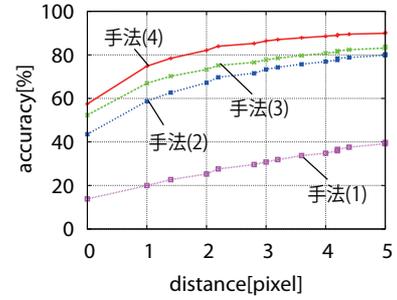
6.1.2 処理時間・メモリ使用量

処理時間は、図 5 の待機時間を除いた 1 フレームの処理に掛かる時間とした。メモリ使用量としては最大使用量を求めた。この時、クライアント側は使用する動画によりメモリ使用量が異なるため、最大値の平均とした。

実験に用いた計算機は、サーバは CPU が Opteron8378 (2.4 GHz)、メモリが 128 GB であり、クライアントは INTEL Core i7-920、メモリが 3 GB である。



(a) 30 個の筆跡の平均



(b) 精度 0% の筆跡を除いた平均

図 6 筆跡復元精度



(a) SURF トラッキング

(b) LK トラッキング

図 8 LK トラッキングの問題点

6.2 実験結果

6.2.1 筆跡の復元精度

図 6 に筆跡の復元精度を示す。横軸は復元画像の筆跡の正解画像からのずれの距離 d 、縦軸は距離 d に対する精度を示す。図 6 (a) は、実験に使用した 30 個の筆跡から求めた精度の平均、図 6 (b) は、文書画像上に復元できなかった、もしくは正解画像の筆跡位置と大幅にずれため、精度が 0 % だった場合を除いた平均値を示す。ここで、1,700 × 2,200 の画像における 1 pixel は、0.12 mm に相当する。

図 6 (a), (b) を比較すると、精度に 10 % 程度差が生じている。この理由として、文書上に復元できなかった筆跡が多いことが挙げられる。手法 (1) では、30 個の筆跡のうち文書上に復元できたのは 2 個だけである。図 6 (a), (b) における精度の差の原因は、他の手法においても同様で、精度が 0 % でない筆跡の数は、手法 (2) 22 個、(3) 23 個、(4) 25 個であった。

図 6 (b) に注目すると、手法 (4) において最も高い精度が得られたことが分かる。手法 (4) において、ずれの距離が 0 pixel、つまり正解画像と完全一致したのは約 57 %、距離が 5 pixel の時の精度は約 90 % である。5 pixel を A4 サイズに変換すると 0.6 mm と小さいため、ずれの範囲はわずかであると言える。

図 7 に復元した筆跡を示す。筆跡 2 に対しては、従来手法では文書上に復元できなかったため、SURF トラッキングによる復元結果を示す。図 7 より、SURF トラッキングを用いた手法 (1), (2) に比べ、LK トラッキングを用いた手法 (3), (4) の筆跡の方が滑らかで欠損が生じていないことが分かる。以上より、

	正解画像	(1)従来手法	(2)比較手法1	(3)比較手法2	(4)提案手法
筆跡1	valuations are reported in Sect <i>ordinary</i> Traditional discrimi	Traditional discrimi From the pattern anal	valuations are reported in Sect <i>ordinary</i> Traditional discrimi	Traditional discrimi	valuations are reported in Sect <i>ordinary</i> Traditional discrimi
筆跡2	the scatter matrix estimates can be obtai $S_y = \sum_{i=1}^N (y_i - m_y)(x_i - m_x)^T$ $S_x = \sum_{i=1}^N (x_i - m_x)(x_i - m_x)^T$ $i = 1, \dots, N_i$ denote the positive exam		the scatter matrix estimates can be obtai $S_y = \sum_{i=1}^N (y_i - m_y)(x_i - m_x)^T$ $S_x = \sum_{i=1}^N (x_i - m_x)(x_i - m_x)^T$ $i = 1, \dots, N_i$ denote the positive exam	the scatter matrix estimates can be obtai $S_y = \sum_{i=1}^N (y_i - m_y)(x_i - m_x)^T$ $S_x = \sum_{i=1}^N (x_i - m_x)(x_i - m_x)^T$ $i = 1, \dots, N_i$ denote the positive exam	the scatter matrix estimates can be obtai $S_y = \sum_{i=1}^N (y_i - m_y)(x_i - m_x)^T$ $S_x = \sum_{i=1}^N (x_i - m_x)(x_i - m_x)^T$ $i = 1, \dots, N_i$ denote the positive exam

図 7 筆跡の復元例

表 1 処理時間とメモリ使用量

手法	処理時間		メモリ	
	Server[ms]	Client[ms]	Server[GB]	Client[MB]
(1) 従来手法	7.36	131.00	3.14	414.34
(2) 比較手法 1	24.46	163.68	3.14	38.62
(3) 比較手法 2	8.61	57.25	3.14	21.50
(4) 提案手法	24.68	60.29	3.14	21.49

LK トラッキングの方が復元に適していると考えられる。

LK トラッキングを用いた手法での失敗例を図 8 に示す。LK トラッキングを用いた図 8 (b) の丸で示す部分では、ペン先の動きは捉えられているが、途中で位置にずれが生じ、筆跡が重なっている。この原因として、連続した 2 フレームの動きのみを見ていることがある。SURF を使用しなかったことで、このような問題が生じており、今後この失敗を防ぐ必要がある。

6.2.2 処理時間とメモリ使用量

表 1 に処理時間とメモリ使用量を示す。クライアント側の処理時間は、SURF トラッキングを用いた手法 (1), (2) より LK トラッキングを用いた手法 (3), (4) の方が短くなった。これは、SURF 特徴点の抽出が処理時間の大半を占めていたためである。サーバ側では、拡張なしの手法 (1), (3) に比べ、拡張を追加した手法 (2), (4) で処理時間が増加した。しかし、サーバ側の処理時間はクエリ側を下回っているため、全体の処理時間に影響を与えないと言える。

次に、メモリ使用量について説明する。サーバ側では、拡張による大幅なメモリ使用量の増加は見られなかった。クエリ側では、LK トラッキングを用いた手法 (3), (4) でメモリ使用量が低下した。この理由は、SURF トラッキングを用いた手法では、毎フレーム抽出した SURF 特徴点を保存する必要があるが、LK トラッキングを用いる手法では、特徴点を保持し続ける必要がないことである。

6.2.3 考察

以上の結果より、トラッキングには、LK トラッキングの方が復元精度、メモリ使用量、処理時間の全てにおいて有効であると言える。また文書画像検索に拡張手法を追加することにより、精度の向上が見られた。この時、サーバ側の処理時間が増加したが、クエリ側の処理時間より短いため、システムの処理時間に影響を与えることはない。したがって、拡張手法もカメラペンシステムにおいて有効な手法であると言える。以上より、LK トラッキングと拡張手法を用いた手法が有効であると考えられる。しかし、復元精度は十分ではなく、更なる精度の向上に取り組む必要がある。特に、文書上に復元できない場合が依

然として多く、文書画像検索の精度改善が必要である。

7. まとめ

我々が開発中のカメラペンシステムでは、筆跡復元精度の低さが問題となっている。本稿では、滑らかな筆跡を復元するため LK トラッキングの導入や、検索処理における拡張手法の追加を提案した。その結果、従来手法に比べ復元精度が向上した。今後の課題として、LK トラッキングにおいて再出現を求めることや、文書画像検索の精度の更なる向上により、筆跡の復元精度を向上させることが挙げられる。また、実用的なシステム構築のためにも、精度の向上と共にメモリ使用量や処理時間の削減に取り組む必要がある。

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] <http://www.anoto.com/>.
- [2] K. Iwata, K. Kise, M. Iwamura, S. Uchida, and S. Omachi. Tracking and retrieval of pen tip positions for an intelligent camera pen. *Proc. of ICFHR2010*, 2010.
- [3] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. *Proc. of IJCAI*, 1981.
- [4] K. Kise, M. Chikano, K. Iwata, M. Iwamura, S. Uchida, and S. Omachi. Expansion of queries and databases for improving the retrieval accuracy of document portions. *Proc. of DAS2010*, 2010.
- [5] [http://www.adesso.com/home/tablets/158 cyberpad.html](http://www.adesso.com/home/tablets/158%20cyberpad.html).
- [6] <http://www.pegatech.com/>.
- [7] T. Arai, D. Aust, and S. E. Hudson. Paperlink: A technique for hyperlinking from real paper to electronic content. *Proc. of the SIGCHI conference on Human factors in computing systems*, 1997.
- [8] M. E. Munich and P. Perona. Visual input for pen-based computers. *IEEE Trans. PAMI*, Vol. 24, No. 3, pp. 313–328, 2002.
- [9] K. Yasuda, D. Muramatsu, and T. Matsumoto. Visual-based online signature verification by pen tip tracking. *Int'l Conf. on Computational Intelligence for Modelling, Control and Automation*, 2008.
- [10] J. Seok, S. Lévassieur, K. Kim, and Jin H. Kim. Tracing handwriting on paper document under video camera. *Proc. of ICFHR2008*, 2008.
- [11] S. Uchida, K. Itou, M. Iwamura, S. Omachi, and K. Kise. On a possibility of pen-tip camera for the reconstruction of handwritings. *Proc. of CBDAR2009*, 2009.
- [12] K. Kise, K. Iwata, T. Nakai, M. Iwamura, S. Uchida, and S. Omachi. Document-level positioning of a pen tip by retrieval of image fragments. *Proc. of CBDAR2009*, 2009.
- [13] 中居友弘, 黄瀬浩一, 岩村雅一. Web カメラを用いたリアルタイム文書画像検索. 信学論 D, Vol. J90-D, No. 8, pp. 2262–2265, 2007.