

カメラペンシステムにおける射影歪みを考慮した 文書画像検索の精度向上法

近野 恵[†] 岩田 和将[†] 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

^{††} 九州大学大学院 システム情報科学研究院

^{†††} 東北大学大学院 工学研究科

E-mail: [†]{chikano,iwata}@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 本論文では、我々が開発中のカメラペンシステムで使用する、文書画像検索の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で書いた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。文書画像検索はデータベースから撮影画像の対応範囲を検索することで、筆跡を求める役割を担っており、高精度での検索が必要不可欠である。しかし現状では、撮影画像に生じる射影歪みの影響で検索精度が低下している。そこで、この問題を解決するために、検索に用いる特徴量の改良、および射影歪みを発生させた画像を生成し、データベースを拡張する手法と、撮影画像を基にクエリを拡張して検索する手法の3つを提案する。有効性を検証するため、改良した特徴量と拡張手法をそれぞれ組み合わせ、改良前の従来手法と比較実験を行った。その結果、改良した特徴量を用い、クエリを拡張した手法を用いたときに、最も精度が向上した。

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

Accuracy Improvement of Document Image Retrieval on a Camera-Pen System by Coping with Projective Distortion

Megumi CHIKANO[†], Kazumasa IWATA[†], Koichi KISE[†], Masakazu IWAMURA[†],
Seiichi UCHIDA^{††}, and Shinichiro OMACHI^{†††}

[†] Graduate School of Engineering, Osaka Prefecture University

^{††} Faculty of Information Science and Electrical Engineering, Kyushu University

^{†††} Graduate School of Engineering, Tohoku University

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

Abstract In this paper, we propose a method to improve accuracy of document image retrieval for a camera-pen system. The system is to acquire handwriting on a printed document as digital ink. In this system, document image retrieval is employed to locate the pen-tip position. Features calculated based on the foreground image are used to retrieve the document image and the pen-tip position on it. A problem of this system is that severe perspective distortion in the query image prevents us from acquiring an accurate position. To solve this problem, we improve the discrimination power of the features using a perspective invariant. In addition, we propose two expansion methods which bring database image and query image closer. Database expansion is to store geometrically distorted images in the database, and query expansion is to generate transformed images from a query image. From the experimental results, we confirm that the best combination is the query expansion with the proposed feature.

Key words document image retrieval, document image processing, LLAH, handwriting, camera-pen system

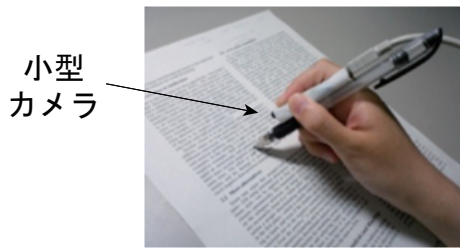


図1 カメラペン

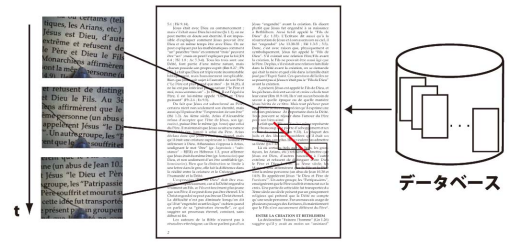


図2 筆迹復元の流れ

1. はじめに

デジタル機器が普及した現在でも、メモを取る時に紙に書き込む機会が多く存在する。この理由として、手軽に利用できることや、資料への関連情報の記入ができるということが挙げられる。そこで、紙に書き込む便利さと、デジタル化の便利さを兼ね備えたシステムがあれば、記録方法の観点でも、記録した情報の利用という観点でも便利になると考えられる。

これらの課題を満たす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、特殊な用紙とカメラを組み込んだペンを用いることで、筆迹を復元するシステムである。紙に筆記するため手軽に使用できるが、特殊な紙がなければシステムを利用できないという問題がある。そこで、印刷文書への筆迹をデジタルデータとして復元するカメラペンシステム [2] が開発されている。このシステムは、撮影画像を用いてデータベースから対応する文書の撮影範囲を検索し、そこからペン先の位置を求め、筆迹を得る仕組みである。

撮影範囲の特定には、Locally Likely Arrangement Hashing(LLAH) [3] による文書画像検索法を用いている。文書画像検索法とは、カメラでの撮影画像に対応する文書画像を、データベースから探し出す手法である。その1つのLLAHとは、印刷文字から特徴点を求め、その配置から得られる特徴量に基づいて検索を行う手法である。LLAHをこのシステムで用いる場合、2つの問題点がある。1つは、文書画像検索で用いる場合に比べ撮影範囲が狭く、識別性が低下する点である。もう1つは、筆記時のカメラの傾きにより射影歪みが生じ、検索精度が低下する点である。

本論文では、これらの問題を解決するための手法を3つ提案する。1つは、特徴点の識別性を上げるため、特徴量を改良することである。他の2つは、クエリとデータベースの画像を近づけるため、射影変換した画像を用い拡張を行う手法である。その1つは、登録画像を射影変換し、データベースを拡張する手法 [4] である。もう1つは、撮影画像を基に複数生成したクエリを用いて検索する、クエリ拡張手法 [5] である。本論文では、改良した特徴量と2つの拡張手法をそれぞれ組み合わせた場合と、従来の特徴量を用い拡張を行わない場合の比較実験を行った。その結果、従来手法に比べ、カメラの傾きが

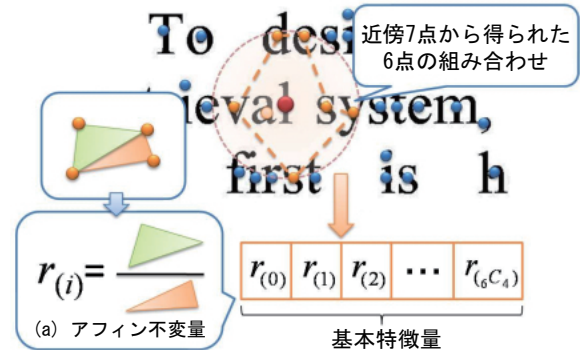


図3 基本特徴量 ($n = 7, m = 6$)

仰角 $75^\circ, 60^\circ$ のときに、文書画像検索精度、筆迹の復元精度が共に向上した。

2. カメラペンシステム

我々が開発中の、印刷文書への筆迹を復元するカメラペンシステムでは、図1のような小型カメラを取り付けたペンを用いる。このシステムにおける筆迹の復元の流れを図2に示す。撮影画像の毎フレームをクエリとし、データベースから対応する文書画像およびその撮影範囲を検索する。また、カメラはペンに固定してあるため、撮影範囲を基にペン先の位置を求めることができる。求めた複数フレームにおけるペン先の位置を、時系列に沿ってつなぎ合わせることで筆迹を得る。このシステムでは、文書上の位置を特定するため、筆迹だけでなく文書との関連性も得ることができる。

以下では、撮影範囲の特定に使用するLLAHを用いた文書画像検索法と、この手法をカメラペンシステムで用いる上での問題点について示す。

2.1 LLAHを用いた文書画像検索法

LLAHは実時間での検索が可能な文書画像検索の一手法である。この手法では、印刷文字から抽出した複数の特徴点を組み合わせ、その配置を用いて検索する。以下に文書画像検索の処理の流れを示す。

データベース構築時は、はじめに登録したい文書画像の特徴点抽出処理を行う。次に特徴量計算を行い、特徴量を基に各特徴点をデータベースに登録する。検索時は、撮影画像の特徴点を抽出し、登録処理と同じ特徴量計算手法を用いて、特徴量を計算する。そして、求めた特徴量を用いて、データベースにアクセスし、対応する文書

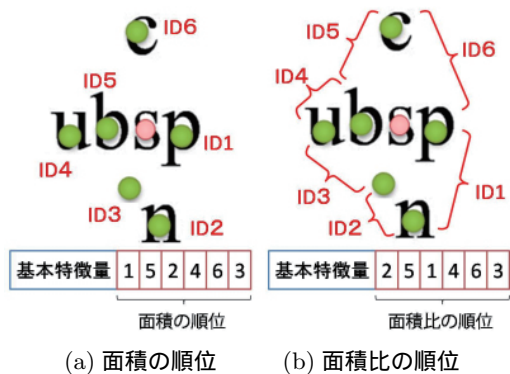


図 4 付加特微量

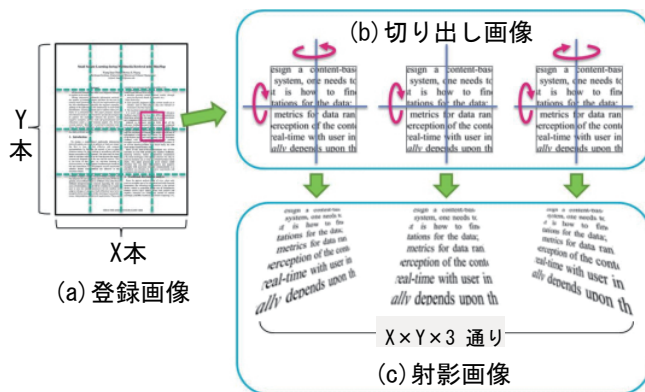


図 5 データベース拡張

画像を検索する。

図 3 に示すように、特徴点には黒画素の連結成分（一文字あるいはその一部に相当する領域）の重心を用いる。求めた各特徴点は、その近傍の特徴点を用いて、多次元の特徴量で表わされる。特徴量の計算方法を図 3 に示す。まず、登録する特徴点の近傍 n 点を求める。次に、 n 点から求められる m 点の全ての組み合わせを求める。登録する特徴点は、 ${}_n C_m$ 個の特徴ベクトルを用いて表わされる。 m 個の組み合わせが決まると、 m 点から求められる全ての 4 点の組み合わせを求める。この 4 点を用い、図 3 (a) のように、アフィン不変量である 2 つの三角形の面積比を求める。このようにして求めた ${}_m C_4$ 個の不変量を特徴ベクトルとし、 m 点の組み合わせをハッシュテーブルへ登録する。以下では、この特徴量を基本特徴量と呼ぶ。

識別性を上げるため、基本特徴量に加え、図 4 (a) に示す特徴量を付加する [6]。この特徴量を求める際は、まず、図 4 (a) のように、登録する特徴点の近傍点 m 点に時計回りに ID を設定する。次に、近傍点の連結成分の面積が大きい順になるように、ID を並べ替える。ここで使用する連結成分の面積は、連結成分に内包された領域の面積とする。並べ替えた m 個の ID の列を付加特徴量として用いる。このように、付加特徴量に面積値ではなく、順位を用いることで、ロバスト性を高めている。

2.2 問題点

本論文では、LLAH をカメラペンシステムで用いる場合に生じる問題点のうち、2 つを取り上げる。LLAH を文書画像検索のために用いる場合と比較すると、(1) 撮影範囲が狭く、(2) 撮影画像に生じる射影歪みの影響が大きくなる点である。撮影範囲に含まれる行数が限られると、基本特徴量の計算に必要な特徴点数が十分得られず、識別性が低下するため検索に失敗する。もう 1 つの問題では、強い射影歪みを受けた撮影画像から、同じ特徴量を求めるのが難しく、検索失敗の原因となる。

3. 提案手法

上記の問題を解決するために、3 つの手法を提案する。

第 1 の問題への対応として、特徴点を改良する手法、第 2 の問題に対しては、データベース拡張とクエリ拡張を提案する。

3.1 特徴量の改良

第 1 の問題に対して、特徴点の識別性を上げるために特徴量の改良を行う。改良方法は、(1) 基本特徴量として、アフィン不変量の代わりに射影不変量を用いる、(2) 付加特徴量として、面積比の順位を用いる、の 2 つである。第 1 の改良では、A~E を登録する特徴点の近傍 5 点とし、以下の式で求められる複比を用いる [7]。

$$\frac{S(A, B, C)S(A, D, E)}{S(A, B, D)S(A, C, E)} \quad (1)$$

$S(x, y, z)$ は頂点 xyz からなる三角形の面積である。この特徴量は、アフィン不変量より 1 つ多い特徴点を必要とするが、射影変換に対して不変である。

第 2 の改良として、図 4 (b) に示すように、付加特徴量に、近傍点の連結成分の面積比の順位を用いる。これは、基本特徴量と同様、アフィン不変量である面積の順位の代わりに、射影不変な付加特徴量に改良するという手法である。改良前と同様に、近傍 m 点を使用する。まず、隣接する近傍 2 点間に ID を設定し、その 2 点の連結成分の面積比を求める。求めた m 個の面積比が、大きい順になるように並び替えた ID の列を基本特徴量に付加する。図 4 (b) の ID1 は、連結成分 “p”、“n” の面積比に対応している。

これらの不変量は、どのような撮影環境に対しても不変ではなく、ノイズなどの影響により誤差が生じる。この誤差により、不変量だけでは解決できない部分にも対応するため、以下に示す 2 つの拡張手法を提案する。

3.2 データベース拡張

データベース拡張では、データベースの特徴点を、射影歪みを受けたクエリの特徴点に近づけるために、登録文書画像を射影変換する。そして、元の登録画像と共に変換した画像から抽出した特徴点もデータベースに登

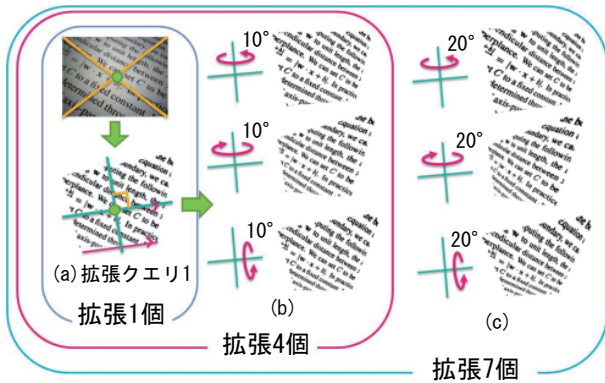


図 6 クエリ拡張

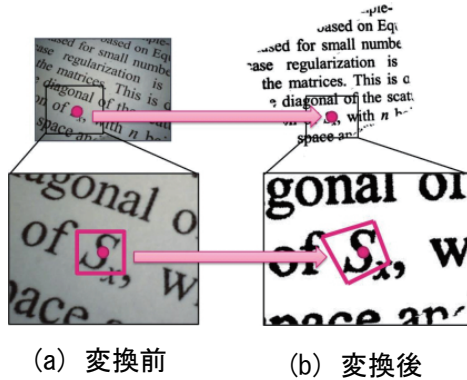


図 7 変換方法

録する。生じる射影歪みは、ペンを置く位置により異なるため、紙面の複数箇所で射影画像を生成する必要がある。しかし、射影画像を限りなく追加すると、メモリ量が増加してしまうため、射影変換する範囲を制限する。図 5 (a) のように、登録文書画像に対し縦線 X 本、横線 Y 本を等間隔に引くと、 $X \times Y$ 個の交点ができる。これらの各交点を中心として画像の一部を切り出す。また、射影歪みはカメラの向きにも因るため、複数方向に傾ける必要がある。切り出した画像に対して回転の自由度は 3 となるが、LLAH では紙面の回転に対して不変な特徴量を用いているために、切り出した画像を平面に垂直な軸で回転させる必要はない。したがって、残りの 2 自由度に対応するため、図 5 (b) のように 2 軸を設定する。ここで、多くの場合、ペンが筆記者の手前方向に傾くことや、紙面を上下逆にして筆記する状況が少ないことが考えられる。そこで、メモリ量を削減するため、上記のような使用頻度が少ない状況を除き、3 方向を考慮した画像を生成する。具体的には、図 5 (c) のように横軸周りに 1 方向、縦軸周りに 2 方向に傾けた射影画像を生成し、データベースに登録する。

3.3 クエリ拡張

クエリ拡張では、クエリの特徴点をデータベースの特徴点に近づけるために、撮影画像を射影変換して、正対に近づけたものをクエリとして検索に用いる。しかし、正確な変換パラメータを求めるのは困難である。そこで、

表 1 提案手法の組み合わせ

特徴量	拡張		
	なし	データベース	クエリ
アフィン+面積	b1(従来手法)	d1	q1-k
アフィン+面積比	b2	d2	q2-k
射影+面積比	b3	d3	q3-k

(1) 筆記時には、前後のフレーム間でペンの傾きが急激に変わる状況が少ないこと、(2) 正しく検索できれば、どのような変換を受けているか推定できるということから、前フレームでの検索結果を利用する。そのため、本論文では以下の処理を行う。まず、前のフレームでのクエリとデータベースの特徴点の対応関係を用いて、射影変換パラメータを推定する。この処理には、RANSAC [8] を用いる。次に、このパラメータを用いて、撮影画像を射影変換する。これにより、図 6 (a) のような画像が得られる。これを拡張クエリ 1 とする。さらに、前フレームと傾きが常に同じであるとは限らないため、多少の傾きの変化を補う画像を追加する。撮影画像の中心点を基に、拡張クエリ 1 に図 6 (a) のような軸を設定する。LLAH では、平面上の回転を考慮する必要がないため、この 2 軸を用いる。また、クエリの個数が増加すると、処理時間の増加が問題として考えられる。そのため、データベース拡張と同様に、使用頻度の高い縦軸周りに 2 方向、横軸周りに 1 方向に傾けた図 6 の (b)、(c) のような画像のみを生成する。

このように、クエリ拡張では複数のクエリを生成することになる。実際に、元の画像に射影変換を施してクエリの画像を新たに生成する処理を加えると、処理時間が大幅に増加し、大きな問題となる。そこで本手法では、特徴量計算に必要な特徴点の位置と、連結成分の面積の近似値のみを求めることにより、計算量を削減する。射影変換後の特徴点は図 7 のように、上記の射影変換パラメータを用いて座標変換する。一方、画像に射影変換を施さず、射影変換後の連結成分の面積を正確に求めることは困難である。したがって、連結成分を包含する矩形を使用し近似値を求める。まず、図 7 (a) の変換前の矩形を求めその面積を R とする。また、この連結成分の面積を A とする。次に、この矩形の 4 頂点を射影変換パラメータを用いて座標変換すると、図 7 (b) のように、変換後の矩形を求めることができる。ここから求められる変換後の矩形の面積を R' とすると、変換後の連結成分の面積は $A \cdot R' / R$ と近似できる。上記のように、4 頂点の変換のみで面積を求めることができるため、画像変換に比べ、大幅に計算量を削減することができる。以上のようにして求めた複数のクエリ k 個 ($k = 1, 4, 7$) を、元の撮影画像と共に検索に用いる。

この節で提案した特徴量と拡張手法の組み合わせ、全 9 通りを表 1 に示す。例えば、 $q2-k$ は、特徴量としてアフィン不変量と面積比の順位を使用し、クエリ拡張を組

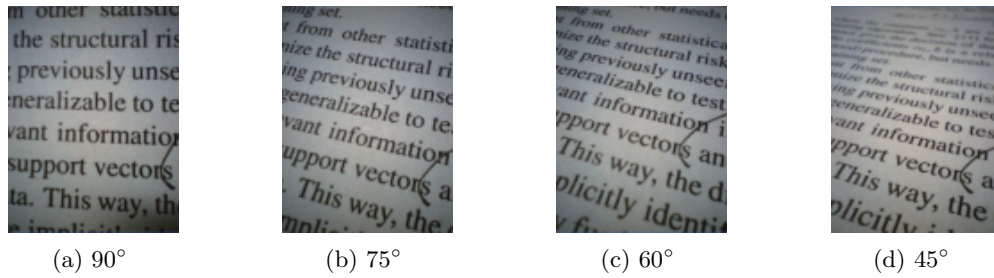


図 8 撮影画像

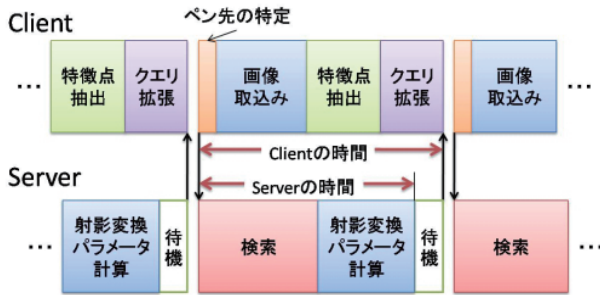


図 9 クライアントサーバモデル

み合わせる手法を示している。 k は拡張するクエリの個数である。これらの手法の中で、どの手法が最も有効であるかを以下で検証する。

4. 実験

4.1 実験条件

実験には、特徴量および拡張手法を表 1 のように組み合わせた全 9 種類を用いた。LLAH のパラメータ n, m は以下のように設定した。アフィン不変量を基本特徴量として用いる場合、 $n = 7, m = 6$ とした。つまり、1 つの特徴点に対する基本特徴量の次元数は ${}_6C_4 = 15$ となる。射影不変量を基本特徴量とする場合は、 $n = 8, m = 7$ とした。次元数は、 ${}_7C_5 = 21$ となる。

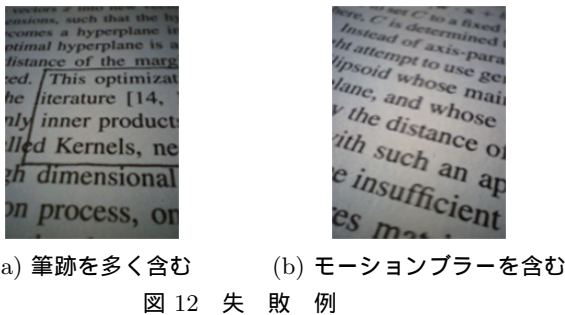
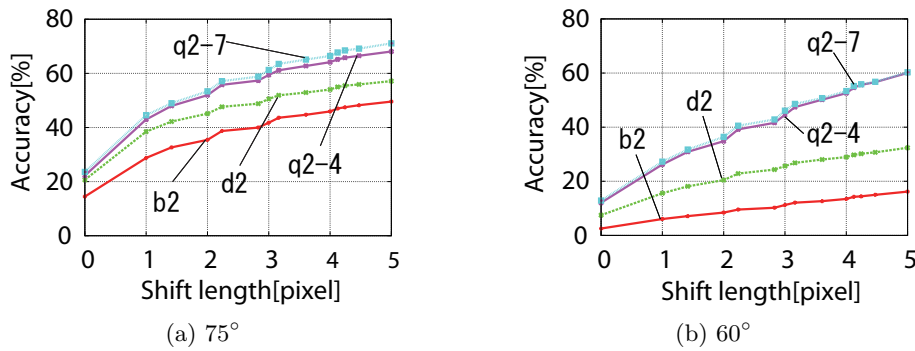
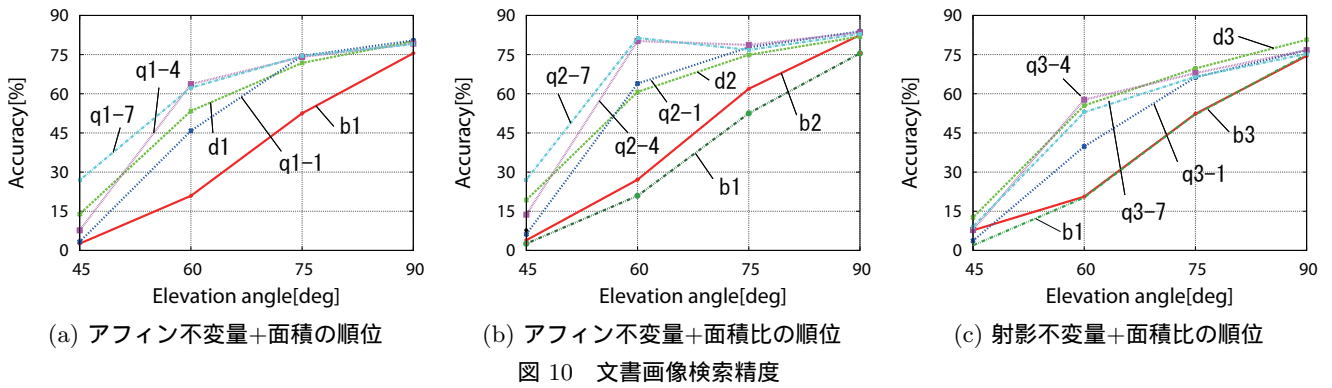
データベースへの登録画像としては、英文書画像 1,000 枚を使用した。登録画像のサイズは $5,100 \times 6,600$ である。検索に用いる撮影画像は、図 8 のように 4 種類の仰角 ($90^\circ, 75^\circ, 60^\circ, 45^\circ$) で撮影した各 300 フレームの動画を 5 通り用いた。画像のサイズは 480×640 である。データベースの拡張では、切り出す画像のサイズを 720×960 とした。これは撮影画像の 1.5 倍である。画像切り出し時のパラメータは $X = 8, Y = 7$ とし、仰角を 60° 、方位角を $0^\circ, \pm 15^\circ$ で変換した。以上により、1 枚の登録画像から 168 通り ($8 \times 7 \times 3$) の画像を生成し、登録画像 1000 枚と共にデータベースに登録した。また、ハッシュへの登録において衝突数が多い場合、衝突に関与した要素を除去した。

クエリの拡張では、撮影画像に加え、生成したクエリを用いて検索処理を行った。拡張するクエリの個数 k は 1, 4, 7 の 3 通りとし、拡張するクエリは、 $k = 1$ では、拡張クエリ 1 (撮影画像を射影変換パラメータで変換し

たクエリ) とした。 $k = 4$ では、拡張クエリ 1 に加え、拡張クエリ 1 を仰角 10° または方位角 $\pm 10^\circ$ で傾けた 3 種類のクエリを用いた。 $k = 7$ では、 $k = 4$ のクエリ 4 種類と、拡張クエリ 1 を仰角 20° または方位角 $\pm 20^\circ$ で傾けた 3 種類のクエリを使用した。

実験結果として、(1) 文書画像検索精度、(2) 筆跡の復元精度、(3) 処理時間、(4) メモリ使用量を求めた。文書画像検索精度は、ハッシュに登録した文書 ID が正しく検索された割合とする。筆跡の復元は、文書画像検索が連続したフレームで成功してはじめて求めることができる。そのため、文書画像検索が高精度で行われることが、筆跡復元において重要な鍵を握っている。筆跡の復元精度は、このシステムで復元した筆跡画像と、解答画像を重ね合わせたときの画素の一致割合とする。具体的には、復元画像を画素単位で段階的に太くし、その画像と解答画像を重ねた時の画素の一致割合をそれぞれ求めた。これにより得られる解答画像とのずれ幅と一致割合の関係を示すことができる。解答画像には、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、印刷時に紙面が傾くことや、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、目視により、ずれの許容範囲を設定し評価を行った。本実験では、許容範囲を 2mm とした。具体的には、以下のような処理手順とした。まず、ペンタブレットによる筆跡を上下左右の方向に 2mm を限度として平行移動し、カメラペンによる筆跡と最も画素が重なる位置を求めた。次に、カメラペンによる筆跡を太くして再度、画素の重なりを求め、精度を計算した。

処理時間の説明のため、このシステムで使用しているクライアントサーバモデルを図 9 に示す。クライアント側では、撮影画像の取得、特徴点抽出、(クエリの拡張)を行う。そして、検索処理のため特徴量をサーバ側へ送る。その直後、このフレームの検索結果を受信する前に、次のフレームの処理に移る。サーバ側では、受け取った特徴量を用いて検索処理、射影変換パラメータの推定を行い、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、処理時間を短縮できる。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、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。本実験では、図 9 に示すように、両側で待機時間を除いた



1 フレームにかかる処理時間を求めた．実験に用いた計算機は，サーバ側が，CPU が Opteron2220SE(2.8GHz)，メモリが 32GB であり，クライアント側は，Intel Core 2 Duo T7700(2.40GHz)，メモリが 4GB である．

4.2 実験結果

4.2.1 文書画像検索精度

文書画像検索精度の結果を図 10 に示す．図 10 (a) は使用する特徴量をアフィン不変量と面積の順位，(b) はアフィン不変量と面積比の順位，(c) は射影不変量と面積比の順位とする．また図 10 (b)，(c) には比較対象として従来手法 (b1) の結果を示す．

図 10 (a)~(c) より，90° では従来手法と提案手法の検索精度の差はほとんどないことが分かる．この理由として，90° では傾きがないため，大部分で元の登録画像と撮影画像の特徴点に対応付き，拡張手法が及ぼした影響が少なかったためであることが挙げられる．75°，65°，45° では，従来手法に比べ，全ての拡張手法で精度が向

上したが，45° では射影歪みの影響が大きく，大幅な上昇には至らなかった．データベースを拡張した手法 (d1，d2，d3) では，拡張した画像と同じ 60° だけでなく，75° でも精度が向上した．角度は異なるものの，射影歪みを受けた部分には，元の登録画像より，拡張データベースの方が類似しており，その特徴点に対応付いたためであると考えられる．図 10(b) において，クエリを拡張した手法 (q2) では，データベース拡張以上の精度の向上が見られた．拡張クエリの個数による精度を比較すると，1 個の場合に比べ 4 個，7 個の方が精度が良くなった．このことから，射影変換パラメータを用いて生成したクエリだけでなく，それに傾きを与えたクエリも用いることが，精度向上に大きな影響を与えていることが分かる．特に，図 10(b) では 75°，65° の場合に 90° と同程度の精度が得られた．

図 10 (b)，(c) より，拡張なしで特徴量の改良のみの手法 (b1，b2，b3) を比較すると，基本特徴量にアフィン不変量，付加特徴量に射影不変量である面積比の順位を用いた b2 が最も精度が良くなった．基本特徴量として射影不変量を用いた場合より，アフィン不変量を用いた場合のほうが精度が向上したのは，特徴量計算に必要な特徴点数が少ないことが理由である．特徴量は，近傍の特徴点の配置を用いて求める．したがって，データベース画像からも，クエリ画像からも，同じ近傍特徴点を得られなければ検索できない．そのため，特徴点計算に必要な特徴点数が多いと，全く同じ近傍特徴点を得られる確率が低くなってしまい，射影不変量は精度が低下したと考えられる．

unseen data. Thus SVM is the
to test data. In addition, SVMs
tion in the training set into a
ors and use only these vectors
the distinguishing features of
identified. These features can
er observations on the support

(a) 手書き

角度	解答	b1	b2	b3	d2	q2-1	q2-4	q2-7
90°								
75°								
60°		文書画像検索失敗						

(b) 復元画像

図 13 筆跡の復元例

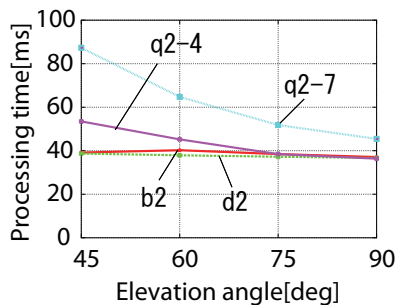
ally tested different ranges o
is used to choose best valu
is what we do in the exam
earest neighborhood of K_L
mputed using the simple Euc
 K_L points a local linear S

(a) 手書き

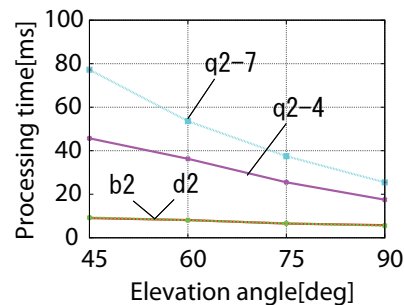
角度	解答	b1	b2	b3	d2	q2-1	q2-4	q2-7
90°								
75°								
60°								

(b) 復元画像

図 14 筆跡の復元例



(a) Client



(b) Server

図 15 処理時間

図 12 に検索に失敗した例を示す．図 12 (a) のように筆跡を含む撮影画像では，筆跡と印刷文書の文字が結合したり，筆跡から新たな特徴点が抽出されたりするため，求められる特徴点の配置がデータベースの配置と異なる．ここで，撮影画像に筆跡を多く含むと，データベースと異なる部分が多くなり，データベースに対応する特徴点数が減少し識別性が低下する．これにより，従来手法だけでなく，提案手法でも検索に失敗した．この問題を解決するために，筆跡を除去するための手法を組み込む必要がある．また，図 12 (b) のようにモーションブラーを含む画像でも，得られる特徴点の配置や面積がデータベースと大きく異なるため，提案手法を用いても改善しなかった．

4.2.2 筆跡の復元精度

75°、60°での筆跡の復元精度を図 11 に示す．図 11 は，横軸に復元した筆跡と解答画像とのずれの画素数，それに対する精度を縦軸に示している．図 11(a) より，75°では従来手法に比べ全ての手法で精度が向上した．さらに図 11(b) に示すように 60°では，拡張なしの手法がほとんど復元できなかったのに対して，拡張手法では精度が大きく向上した．クエリ拡張手法では，データベース拡張手法に比べて，精度の向上率が大きいことが分かる．

データベース拡張では，あらかじめ傾きを一定値に固定した画像を登録するため，撮影画像の傾きと拡張したデータベース画像の傾きが等しいとは限らず，検索により求められるクエリとデータベース間の対応点数が少なくなる．一方，クエリ拡張では，毎フレームの傾きを考慮したクエリを用いるため，データベースとの対応点数が増える．これにより，射影変換パラメータ推定がより正確に行えるため，精度が向上したと考えられる．このように筆跡の復元は，文書画像検索精度に依存するため，文書画像検索精度の良い手法で，復元精度も良くなっている．

図 13 (a) および図 14 (a) は実際に手書きした筆跡であり，これらを復元した画像をそれぞれ図 13 (b) および，図 14 (b) に示す．図 13 (b)，図 14 (b) より，従来手法では復元できていない部分も，拡張手法では復元できている．データベース拡張手法では，欠けている部分やノイズが多いが，クエリ拡張手法では，拡張回数が増えるにつれて正確に復元できている．さらに，図 13 (b) の 60°では，文書画像検索自体が失敗したため，従来手法では全く復元できなかったのに対して，クエリ拡張 (q2-4, q2-7) では 90°の筆跡と同程度の復元ができた．ただし，復元は十分ではなく，筆跡に生じているブレや

ノイズ、欠けている部分などに関して、改善していく必要がある。

4.2.3 処理時間

クライアントとサーバの処理時間を図 15 に示す。データベース拡張手法の処理時間は、両側とも従来手法と同等であった。クエリ拡張手法では、クライアントで傾きが大きくなるにつれて、処理時間が増加した。図 15 (a) に示すように、 60° でクエリ拡張個数が 7 個のとき、クライアントの処理時間が増加した。ただし、拡張個数 4 個の場合は、増加時間は 6ms 程度であり、実時間処理に大きな影響を及ぼさないと考えられる。一方、図 15 (b) に示すように、サーバではクエリの拡張数が増えるほど検索に要する時間が増えるため、処理時間が増加した。クエリを 4 個拡張する手法で、 60° における処理時間は、従来手法の約 5 倍に増加していることが分かる。しかし、サーバの処理時間 (36.29ms) はクライアントの処理時間 (45.20ms) を下回っており、カメラペンシステムでの処理時間には影響を与えない。

4.2.4 メモリ使用量

表 2 に示すように、データベース拡張手法により、メモリ使用量は大幅に増加した。これは、データベースを大規模化する上で大きな問題となる。クエリ拡張手法では、メモリ使用量は従来手法と同じであった。基本特徴量に射影不変量を使用した手法では、アフィン不変量を使用した手法に比べ、メモリ使用量が大きくなった。この理由として、特徴量の次元数が多いことが挙げられる。

4.2.5 考察

基本特徴量に射影不変量を用いる場合は、特徴量計算により多くの特徴点数を必要とするため、撮影範囲が狭い場合には精度、メモリ使用量共に、アフィン不変量の方が適していると考えられる。基本特徴量をアフィン不変量とした場合、付加特徴量には、射影不変量である面積比の順位を用いる方が精度が良い結果が得られた。データベース拡張は、処理時間を従来手法と同等に保てたものの、メモリ使用量の増大が問題として残った。それに対し、クエリ拡張手法は、メモリ使用量が従来手法と同じであった。拡張するクエリの個数により処理時間は増加したが、実時間処理を維持できるようにクエリの拡張個数を抑えても、精度を大幅に改善できるため有効であるといえる。これらの組み合わせを考慮すると、特徴量にアフィン不変量と面積比の順位を用い、4 個のクエリを拡張する手法 (q2-4) が良いと考えられる。ただし、カメラペンシステムの用途として、筆跡の復元精度は十分ではなく、更なる精度の向上が必要となる。

5. まとめ

LLAH を用いたカメラペンシステムには、撮影範囲が狭く識別性が低下する問題と、ペンが傾くことにより撮影画像に射影歪みが生じ、検索精度が低下する問題がある。本論文では、これらの問題を解決するために、3 つ

表 2 メモリ使用量

手法	メモリ使用量 [GB]
b1, b2, q1, q2	3.62
b3, q3	4.33
d1	10.00
d2	10.45
d3	11.72

の手法を提案した。1 つは前者の問題に対し、特徴点の識別性を上げるために、特徴量を改良する手法である。他の 2 つは、射影歪みに対し、射影変換画像を用いデータベースおよびクエリを拡張する手法である。改良した特徴量と拡張手法をそれぞれ組み合わせた結果、従来手法に比べ、文書画像検索精度、筆跡の復元精度共に向上した。特に、特徴量にアフィン不変量と面積比の順位を用い、クエリを拡張する手法では最も精度の向上がみられた。ただし、筆跡の復元精度は十分用途を満たすものではなく、改善の余地がある。

最も重要な今後の課題は、筆跡の復元精度を向上させることである。具体的には、岩田らによるカメラペンシステム [9] との統合、撮影画像に含まれる筆跡への対処などを考えている。

謝辞 本研究の一部は日本学術振興会科学研究費補助金基盤研究 (B)(19300062, 20300049) ならびに挑戦的萌芽研究 (21650026) の補助による。

文 献

- [1] <http://www.anoto.co.jp/>.
- [2] K. Iwata, K. Kise, T. Nakai, M. Iwamura, S. Uchida and S. Omachi: "Capturing digital ink as retrieving fragments of document images", Proceedings of the 10th International Conference on Document Analysis and Recognition (ICDAR2009), pp. 1236-1240 (2009).
- [3] 中居, 黄瀬, 岩村: "Web カメラを用いたリアルタイム文書画像検索", 電子情報通信学会論文誌 D, **J90-D**, 8, pp. 2262-2265 (2007).
- [4] 近野, 岩田, 黄瀬, 岩村, 内田, 大町: "カメラペンシステムの検索精度向上のための一手法", 平成 21 年電気関係学会関西支部連合大会講演論文集 (2009).
- [5] 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 Document Analysis Systems 2010 (2010). [to appear].
- [6] 中居, 黄瀬, 岩村: "Web カメラを用いた多言語文書画像のリアルタイム検索システム", 電子情報通信学会技術研究報告, **108**, 432, pp. 115-120 (2009).
- [7] 中居, 黄瀬, 岩村: "特徴点の局所的配置に基づくデジタルカメラを用いた高速文書画像検索", 電子情報通信学会論文誌 D, **J89-D**, 9, pp. 2045-2054 (2006).
- [8] M. A. Fischler and R. C. Bolles: "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", Comm. of the ACM, pp. 381-395 (1981).
- [9] 岩田, 黄瀬, 岩村, 内田, 大町: "特徴追跡と文書画像検索を用いたカメラペンシステム", 電子情報通信学会技術研究報告, **109**, 418, pp. 39-44 (2010).