

# ペン入力タブレット用筆跡検索技術

Handwriting Retrieval Technology for Tablets with Pen Input

柴田 智行

登内 洋次郎

中居 友弘

■ SHIBATA Tomoyuki

■ TONOUCHI Yojiro

■ NAKAI Tomohiro

ペン入力可能なタブレットやスマートフォンの普及に伴い、端末に手書き入力した手書きデータをデジタル保存できるようになった。書きためた手書きデータが増えると、ユーザーがそこから知りたい情報を見つけるには時間が掛かるため、目的の情報を探す機能が必要になる。

東芝は、目的の情報を少ないメモリ量で高速に探すことができる、手書き入力による筆跡検索技術を開発した。筆跡検索は、検索のために入力した手書きデータと一致する手書きデータを、書きためた中から書き順や形状の情報を使って探す技術である。日本語や英語といった言語情報を使わない筆跡検索は、文字だけでなく記号やイラストなども同様に探すことができる。

The wide dissemination of smartphones and tablets with pen input has provided users with the capability to easily store handwritten data as digital data. Accompanying the increase in such stored data, demand has arisen for a technology to rapidly extract the necessary information from large volumes of handwritten data.

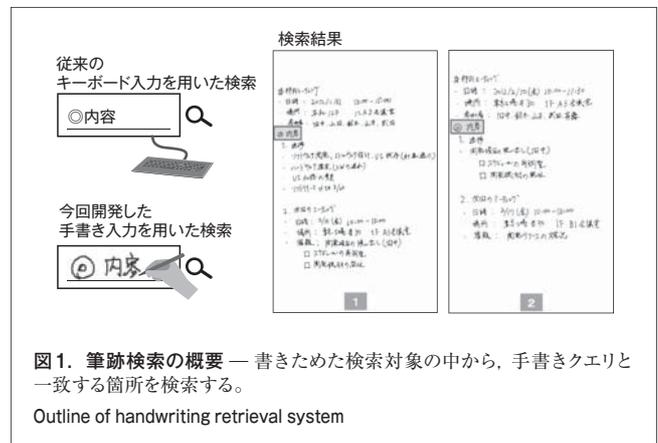
To meet this demand, Toshiba has developed a handwriting retrieval technology that can retrieve targeted handwritten data, including not only words but symbols as well, through recognition of the stroke order and shape of strokes without the need for information on the language in which the characters are written such as Japanese or English. This technology, which operates at high speed with a small amount of memory usage, has been incorporated into the AT703 REGZA tablet.

## 1 まえがき

画面上にペンで直接入力するタブレットやスマートフォンの普及により、手書きデータがデジタル保存できるようになった。東芝は、紙への書き心地に近い感覚で手書き入力できるREGZA tablet T703を2013年6月に商品化した。その差別化技術の一つとして、手書き入力による筆跡検索技術<sup>(1)</sup>を開発した。筆跡検索は、従来のキーボード入力による検索機能とは異なり、ペンで入力した手書きデータ（以下、手書きクエリと呼ぶ）を使い、書きためた手書きデータ（以下、検索対象と呼ぶ）から目的の手書きデータを見つける技術である（図1）。

手書きデータを探す手段として、文字認識を用いて検索対象と手書きクエリの両方を文字コードに変換し、文字列検索する方法がある。このような文字認識を用いる方法では、文字コードへ変換するため言語情報が必要となる。また、文字以外の手書きデータは検索できない。ペン入力は、図やイラストや記号など文字以外も自由に書き込めることが利点であり、検索機能として文字以外の手書きデータも検索対象となるため、文字認識を用いた検索技術は筆跡検索として最適ではない。

一方、文字認識を用いない検索技術では、筆画（ストローク）の形状や書き順などの情報を手書きデータの特徴として表現し、検索対象から手書きクエリと一致する手書きデータを見



つけ出す。タブレットには1ページ中に平均して1,000ストローク程度書かれ、書きためられた数百ページ全ての手書きデータについて検索することも想定される。目的の手書きデータを探すためには、検索対象と手書きクエリ間の多数あるストロークの組合せについて、ストロークの類似性を計算する必要がある。また、書き順の違いや画数の違いを考慮して検索するためには、単純な組合せの数十倍程度の計算量が更に必要になる。つまり、文字認識を使わない筆跡検索は、計算量とその計算に必要なメモリ量の多さが問題となる。

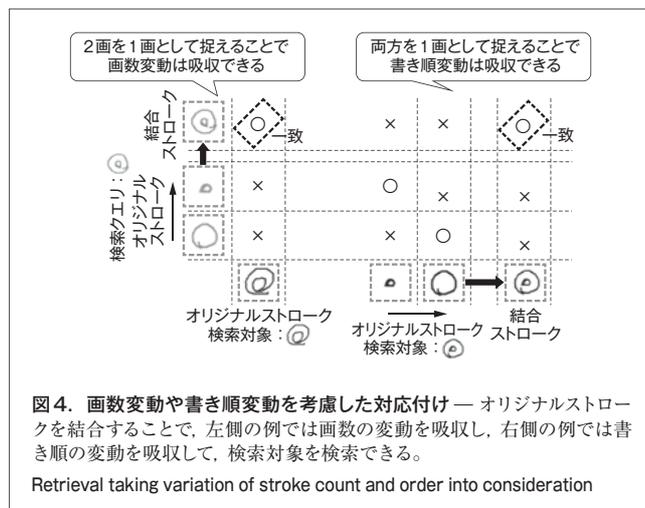
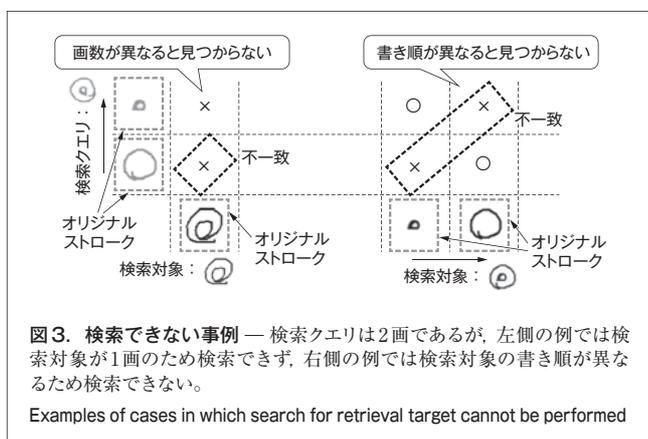
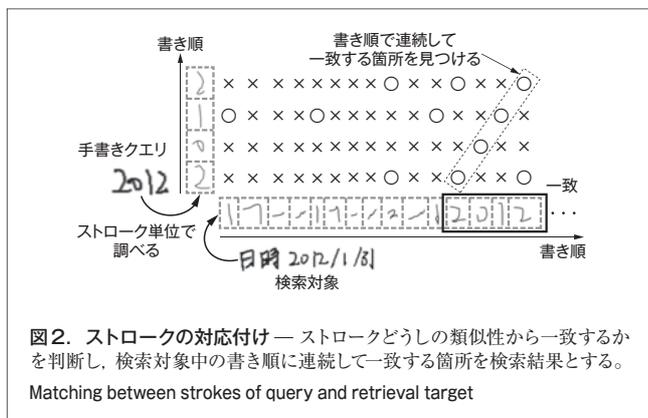
そこで、ストロークの形状を少ないメモリで表現できる特徴

記述方法と、大量の手書きデータから一致する箇所を高速に見つける検索技術を開発し、省メモリで高速な筆跡検索を実現した。ここでは、開発した筆跡検索技術の概要とストロークの特徴記述、及び評価実験の結果について述べる。

## 2 筆跡検索の概要

もっとも基本的な筆跡検索技術では、手書きクエリと検索対象のストローク間の組合せについて、ストロークが類似するかを全て判断する(図2)。そして、検索対象の中の類似するストロークが書き順に連続して手書きクエリと一致する箇所を検索結果とする。ストロークの類似性は、それぞれのストロークの形状情報を表現した特徴量の近さ(距離)として計算される。

ユーザーが検索対象と手書きクエリをまったく同じ画数かつ同じ書き順で書くことができれば、前述の基本的な検索方法で探すことができる。しかし、ユーザーが異なった書き方をすると、目的のストロークは見つからない(図3)。そこで、本来のストローク以外の可能性も考慮し、書き順で連続するオリジナルストロークを結合して、仮想的に画数を変動させた結合ストロークを生成し、類似する箇所を探す(図4)。これにより、異なる画数のストロークや、書き順が異なるストロークを対応付けることができる。



筆跡検索全体の計算量の多くは、ストロークの類似性を算出する処理である。画数変動や書き順変動に対応するには、多くの可能性を考慮する必要があるため、全ての可能性を計算するのは効率的ではない。そこで、同様の問題でよく用いられる動的計画法(DP: Dynamic Programming) マッチング<sup>(2)</sup>と呼ばれるアルゴリズムを導入した。これにより、手書きクエリと一致する可能性がないストロークとの類似性を計算しなくても検索できるため、計算量が削減される。

一方、筆跡検索全体のメモリ使用量の多くは、ストロークの特徴量である。少ないメモリ使用量でかつ高速な検索を実現する鍵となるストロークの特徴記述について次に述べる。

## 3 ストロークの特徴記述

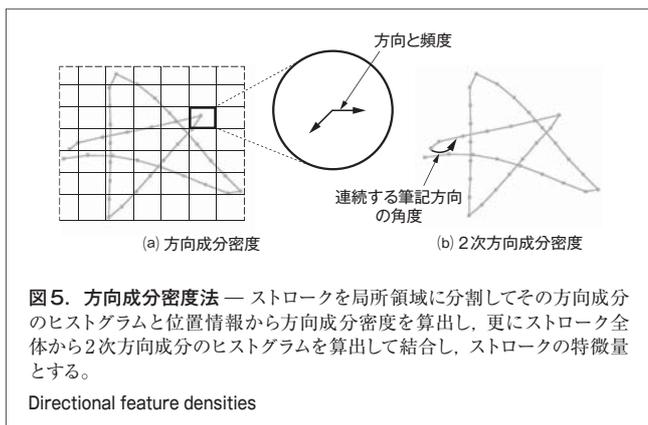
ストロークは、ペンを下ろして書き始め、再び離すことのできる線及び点である。タブレットでは、画面に接したペンの位置を2次元座標として捉え、短い時間周期で座標値を取得し、2次元座標の点列としてストロークを保存する。

ストロークは、一定の時間間隔でペンの位置を取得するため、同じ形状を書いたとしても筆記速度によって得られる点列の点数が変わる。表現する特徴が、筆記速度に依存しないように、ストロークを正規化する。例えば、取得した星形のストロークは、点の密度が角の部分で高く、直線の部分で低いが、固定の点数でストロークを正規化することにより、全てのストロークが共通の点数で点列を保存できる。

この正規化したストロークに対して、ストローク形状を表現した特徴を抽出する方法を説明する。

### 3.1 方向成分密度法

ストロークの形状を表現する特徴の一つに、方向成分密度<sup>(3)</sup>がある。方向成分密度は、文字認識などに使用される特徴量であり、算出したストロークの局所的な筆記方向の頻度分布(ヒストグラム)として表現される。局所的な筆記方向は、スト



ロークを構成する点から次の点へ向かう方向とする。終点を除く全ての点で筆記方向を離散化した方向（方向成分）のヒストグラムを、ストロークの特徴とする。空間的に分割した複数の局所領域のヒストグラムを算出し統合することで、局所的な筆記方向のヒストグラムだけでなく、書かれた位置の情報も表現する（図5(a)）。縦と横でそれぞれ7分割した局所領域から、8方向に離散化した方向成分密度を算出して統合し、 $7 \times 7 \times 8 = 392$ 次元のベクトルを求める。

更に、連続する二つの筆記方向が成す角度を離散化して2次方向成分の値とし、そのヒストグラムをストローク全体から算出する（図5(b)）。2次方向成分は36方向に量子化する。

方向成分密度と2次方向成分密度を結合し、 $392 + 36 = 428$ 次元のベクトルを、ストローク形状を表現する特徴量とする。

### 3.2 特徴量の近似

二つの特徴量の類似性を高速に計算する方法の一つに、特徴量ベクトルを低次元ベクトルへ近似する技術がある。特徴量ベクトル間の距離は、次元数に比例して計算量が増加するが、低次元ベクトルへ近似することで距離算出の計算量が抑えられる。高次元ベクトルを低次元ベクトルへ近似する方法として、線形変換行列を用いた低次元射影がある。ストロークを記述した $D$ 次元特徴量ベクトル $v$ から、 $D > d$ となる $d$ 次元ベクトル $v'$ へ、式(1)を用いて近似する。

$$v' = Mv \quad (1)$$

ここで、 $M$ は $D$ 行 $d$ 列（ $d=128$ ）の線形変換行列である。線形変換行列を算出する技術として、近似誤差の少ない当社独自の技術Random Ensemble Metrics (REMetric)<sup>(4)</sup>を用いる。

### 3.3 バイナリ特徴量

近似ベクトル $v'$ は1要素を4バイトの浮動小数点で表現しているが、メモリ使用量を抑制するために、1要素を2値（バイナリ値）である1ビットで表現する。REMetricにより低次元空間へ射影された近似ベクトルは、各要素が0を中心に分布するため、式(2)を用いてバイナリ特徴量ベクトル $b$ へ変換する。

$$b_k = \begin{cases} 0 & v'_k < 0 \text{ のとき} \\ 1 & \text{その他のとき} \end{cases}, k=1, 2, \dots, d \quad (2)$$

$b$ のメモリ使用量は128ビット（16バイト）となり、元の特徴量と比較すると、 $1,712 (= 428 \times 4)$ バイトから16バイトへ圧縮され約1/100へと削減される。

二つのバイナリ特徴量の距離は式(3)で定義され、ハミング距離と呼ばれる。

$$\text{dist}(b_i, b_j) = W_H(b_i \text{ xor } b_j) \quad (3)$$

ここで、xorは排他的論理和とし、 $W_H(a)$ はベクトル $a$ のハミング重みを算出する（要素中の1を数える）関数とする。

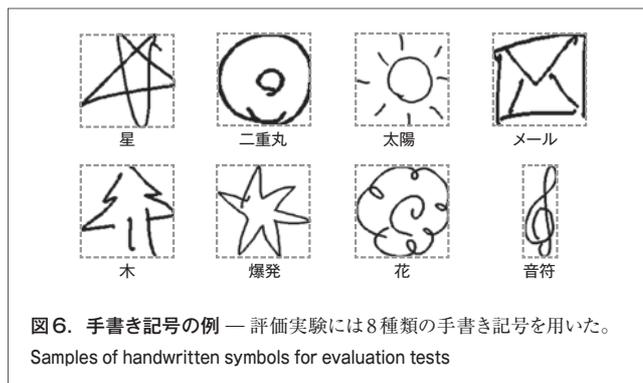
ハミング距離は、小数点精度の特徴ベクトル間の距離で用いるユークリッド距離に比べ、非常に高速に計算できる。同じ次元数のベクトル間距離では、ユークリッド距離と比較して、ハミング距離は約20倍以上高速に計算することができる。

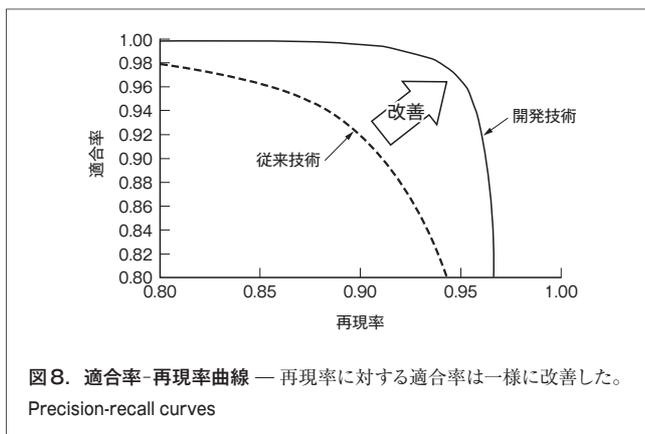
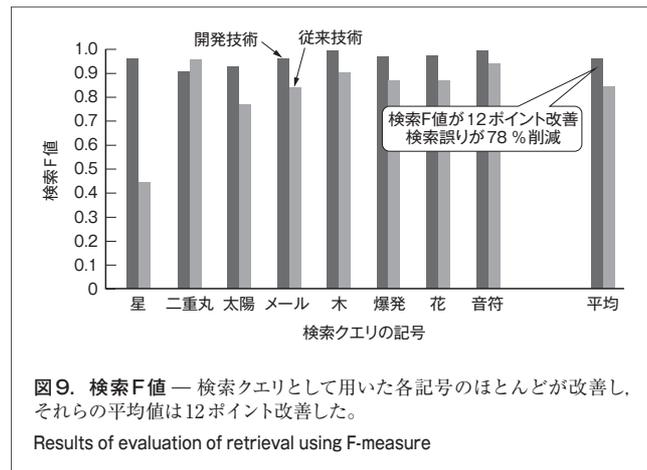
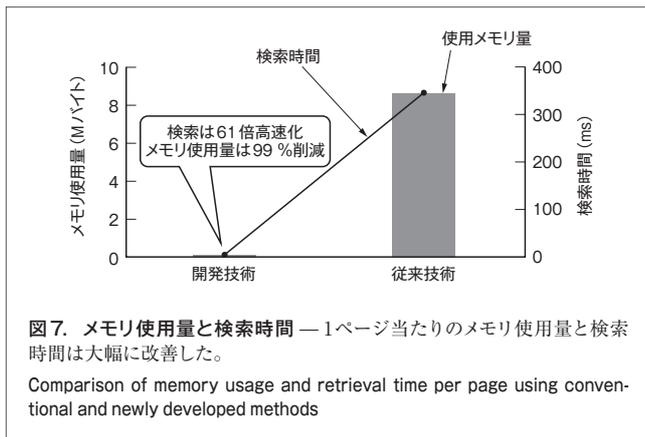
## 4 評価実験

筆跡検索の性能を、REGZA Tablet AT703を使用して、独自に収集した手書きデータから図6に示す8種類の手書き記号を検索することで、評価した。手書きデータは、1ページ当たり平均915ストローク書かれた377ページのデータである。比較対象の従来技術では、ストロークの形状を方向成分密度で記述し、一般的なDPマッチングを用いて探索した。開発した技術では、ストロークの形状をバイナリ特徴量で記述し、高速化したDPマッチングを用いて探索した。高速化したDPマッチングは、手書きクエリと一致する可能性が低い検索対象との距離計算を打ち切ることで、少ない処理量で目的の手書きデータを探すことができる技術である。

1ページ当たりのメモリ使用量と検索時間を図7に示す。開発した技術は従来技術に比べ、検索速度が約61倍高速になり、メモリ使用量が約99%削減された。メモリ使用量と計算量は、バイナリ特徴量導入により劇的に改善した。

次に検索性能を比較した結果を図8に示す。検索性能は、





検索対象から正解をどれだけ検索できたかを測る再現率と、正しく検索されたかを測る適合率を用いて比較する。開発した技術に関して、常に再現率に対する適合率が従来技術よりも高いことから、検索性能が改善したことが確認できる。

また、再現率と適合率を一つの基準で表現でき、両技術の調和平均である検索F値を用いて、検索クエリの各記号について比較した結果を図9に示す。検索F値は、両方の評価値が高くなければ高い値を示さない指標である。検索F値は、一部の記号で劣化するが、多くは改善しており、全ての記号の平均値が12ポイント改善し、検索誤りが78%削減された。

## 5 あとがき

当社が開発した、手書きクエリと一致する手書きデータを、少ないメモリ使用量で高速に検索できる筆跡検索技術について述べた。開発した筆跡検索技術を使うことで、従来技術と比べ検索速度が約61倍速く、1/100のメモリ使用量で目的のデータを検索することができる。また検索性能も改善され、従来技術と比較して検索誤りが78%少ない、という評価結果が得られた。

筆跡検索技術は、ペンで直接クエリ入力できる手書きデータ用の検索機能として、紙への書き心地に近い感覚でペン入

力できるREGZA Tablet AT703に搭載された。同時に搭載された手書き図形認識技術<sup>(5)</sup>と併せて、紙のノートにはない、デジタルデータの利点を生かした検索や整形などの機能をユーザーに提供している。

## 文献

- (1) Shibata, T. et al. "Fast and Memory Efficient Online Handwritten Strokes Retrieval Using Binary Descriptor". Proceedings of 2nd Asian Conference on Pattern Recognition (ACPR2013). Naha, Japan, 2013-11, International Association of Pattern Recognition (IAPR). 2013, p.647 - 651.
- (2) Uchida, S. et al. "Analytical Dynamic Programming Matching". Proceedings of 12th European Conference on Computer Vision (ECCV 2012), Workshops and Demonstrations. Florence, Italy, 2012-10, Springer Berlin Heidelberg. 2012, p.92 - 101.
- (3) Kawamura, A. et al. "Online recognition of freely handwritten Japanese characters using directional feature densities". Proceedings of 11th International Conference on Pattern Recognition (ICPR1992). The Hague, The Netherlands, 1992-08, IAPR. 1992, p.183 - 186.
- (4) Kozakaya, T. et al. "Random Ensemble Metrics for Object Recognition". Proceedings of 13th International Conference on Computer Vision (ICCV 2011). Barcelona, Spain, 2011-11, IEEE. 2011, p.1959 - 1966.
- (5) 高橋梓帆美 他. "複数の手書き図形を一括変換する図形認識技術". 第16回画像の認識・理解シンポジウム (MIRU2013). 東京, 2013-07, 情報処理学会 CVIM 研究会. 2013, DS-10.



柴田 智行 SHIBATA Tomoyuki

研究開発センター インタラクティブメディアラボラトリー研究主務。画像処理及びパターン認識の研究・開発に従事。電子情報通信学会会員。

Interactive Media Lab.



登内 洋次郎 TONOUCHI Yojiro

研究開発センター インタラクティブメディアラボラトリー主任研究員。オンライン文字認識及び画像認識の研究・開発に従事。

Interactive Media Lab.



中居 友弘 NAKAI Tomohiro, D.Eng

研究開発センター インタラクティブメディアラボラトリー研究主務、博士(工学)。画像認識の研究・開発に従事。電子情報通信学会会員。

Interactive Media Lab.